Effects of school-based mindfulness training on emotion processing and well-being in adolescents
Sanger, Kevanne; Thierry, Guillaume; Dorjee, Dusana

Developmental Science

DOI:
10.1111/desc.12646

E-pub ahead of print: 22/01/2018

Publisher's PDF, also known as Version of record

Cyswllt i'r cyhoeddiad / Link to publication

Dyfnyiad o’r fersiwn a gyhoeddwyd / Citation for published version (APA):

Hawliau Cyffredinol / General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Effects of school-based mindfulness training on emotion processing and well-being in adolescents: evidence from event-related potentials

Kevanne Louise Sanger1,2 | Guillaume Thierry2 | Dusana Dorjee2

1Centre for Mindfulness Research and Practice, Bangor University, Bangor, Gwynedd, UK
2School of Psychology, Bangor University, Bangor, Gwynedd, UK

Correspondence
Dusana Dorjee, Bangor University, School of Psychology, Brigantia Building, Penrallt Road, Bangor, Gwynedd LL57 2AS, UK.
Email: d.dorjee@bangor.ac.uk

Funding information
Economic and Social Research Council, Grant/ Award Number: ES/J500197/1

Abstract
In a non-randomized controlled study, we investigated the efficacy of a school-based mindfulness curriculum delivered by schoolteachers to older secondary school students (16–18 years). We measured changes in emotion processing indexed by P3b event-related potential (ERP) modulations in an affective oddball task using static human faces. ERPs were recorded to happy and sad face oddballs presented in a stimulus stream of frequent faces with neutral expression, before and after 8 weeks of mindfulness training. Whilst the mean amplitude of the P3b, an ERP component typically elicited by infrequent oddballs, decreased between testing sessions in the control group, it remained unchanged in the training group. Significant increases in self-reported well-being and fewer doctor visits for mental health support were also reported in the training group as compared to controls. The observed habituation to emotional stimuli in controls thus contrasted with maintained sensitivity in mindfulness-trained students. These results suggest that in-school mindfulness training for adolescents has scope for increasing awareness of socially relevant emotional stimuli, irrespective of valence, and thus may decrease vulnerability to depression.

RESEARCH HIGHLIGHTS
- Mindfulness training was associated with maintained P3b mean amplitudes to facial target stimuli, indicating sustained sensitivity to socially relevant, affective stimuli.
- Trained students reported higher well-being despite mindfulness course engagement being correlated with greater stress awareness.
- Self-reported changes in empathy correlated significantly with changes in P3b to emotional faces across groups.

1 | INTRODUCTION
Adolescence is a time of stress vulnerability, with high academic demands and social pressures. It is also considered a sensitive point for depression onset, with the reward system of the brain (ventral striatum) maturing before the prefrontal regions (PFC, ACC) that balance decision-making and regulate behaviour (Ernst, Pine, & Hardin, 2006). According to the Rescorla-Wagner learning model, the dopaminergic response to reward means that when a distal goal is not achieved, dopamine expression is suppressed, which can lead to prolonged...
supporting this model (Bylsma, Morris, & Rottenberg, 2008) shows that MDD groups exhibit dampened responses to both positive and negative stimuli compared to controls. In the adolescent literature, Blom et al. (2015) reported that depressed adolescents showed significantly reduced anterior/middle insular cortex activation when viewing sad as opposed to happy faces, as compared to controls, an effect considered a developmental signature of depression (see Smith, Steinberg, & Chein, 2013).

It has been argued that mindfulness encourages openness to both positive and negative experience whilst minimizing reactivity or rumination. It can promote activation, and neural connectivity in regions associated with social understanding, bodily awareness, and empathy—the insular cortex (Farb, Segal, & Anderson, 2012) and the TPJ (Hölzel et al., 2011). Birnie, Specsa, and Carlson (2010) also showed that mindfulness training can increase social connectedness and empathy while decreasing personal distress. Mindfulness training may therefore enhance the processing of emotional expression, a mechanism by which individuals can then connect with others. This was assessed after MBCT training with formerly depressed adults by De Raedt et al. (2012), who found that mindfulness training resulted in more balanced receptiveness to positive and negative facial stimuli, in comparison to controls who also had a history of depression. Such an effect in adolescents who may also be considered vulnerable to depression could lead to enhancing social connectedness and potentially buffer them against mental illness (Donald & Dower, 2002; Music, 2014).

Event-related brain potentials (ERPs) have several advantages for neurodevelopmental research in an education context: their temporal resolution is high, the recording system is portable thus allowing for testing in schools, and the method is cost-effective in comparison to fMRI (Sanger & Dorjee, 2015). To our knowledge, however, no neuroscientific research has yet investigated changes in emotion processing with mindfulness training in adolescents, using either fMRI or ERPs. One appropriate ERP marker for detecting such changes is the P300 (particularly the P3b), which indexes task-related information processing (Sutton, Braren, Zubin, & John, 1965). The P3b has been associated with several brain regions, but evidence suggests a key link with the anterior cingulate cortex (ACC) and TPJ (Kok, 2001; Polich, 2007). The ACC is involved in top-down regulation of attention and emotion (Bush, Luu, & Posner, 2000), and both brain regions are sensitive to modulation by mindfulness-based training (Cahn & Polich, 2006; Hölzel et al., 2007; Hölzel et al., 2011). Moreover, the P3b has previously been used as an index of choice in cognitive empathy tasks (e.g., Han, Fan, & Rao, 2008; Ikezawa, Corbena, & Wexler, 2014; Meng et al., 2012). The P3b is also modulated by change in emotive facial processing in adults with depression (Cavanagh & Geisler, 2006). Specifically, P3b amplitude elicited by happy faces was reduced in MDD participants, indicating a limited ability to process positive facial cues. Given the negative correlation between mindfulness and depression documented in previous studies (Deng et al., 2014; Desrosiers et al., 2013), it can be hypothesized that mindfulness training would increase P3b amplitudes to happy faces. Consistent with the predictions of the ECI model of depression and the results
of MBCT effects in adults vulnerable to depression (De Raedt et al., 2012), P3b responses to sad faces may equally be modulated, while self-report measures would show increased well-being. These predictions would corroborate the pattern of enhanced affective processing, without negative mood induction, which has been observed in previous research (Birnie et al., 2010; Dorjee, Lally, Darral-Rew, & Thierry, 2015).

The purpose of this study was to investigate longitudinal changes in emotion processing in older secondary school students (16–18 years) after mindfulness training delivered as a module in Personal, Social, and Health Education (PSHE). The evaluations included self-report questionnaires and P3b modulations recorded in a computerized emotional oddball task in response to affective faces (happy and sad face oddballs—10% of stimuli each type, presented amongst frequent neutral faces—80% of stimuli). The task was based on a previous study by Cavanagh and Geisler (2006) that found reduced P3b response to happy face oddballs and delayed P3b amplitudes to fearful face oddballs in depressed compared to non-depressed university students. In line with previous findings (Huppert & Johnson, 2010; Metz et al., 2013), we expected that mindfulness training would improve well-being and mental health in trained adolescents compared to controls. Critically, we expected that P3b responses elicited by emotional stimuli would be significantly increased in the intervention group. This would suggest that mindfulness-based practice enhances the processing of socially relevant stimuli and healthy emotional exposure, a pattern that contrasts with neurocognitive responses to emotional stimuli in people with depression (Blom et al., 2015; Cavanagh & Geisler, 2006).

2 | METHODS

2.1 | Participants

The study was approved by the Ethics Committee in the School of Psychology at Bangor University, prior to participant recruitment. Participants were recruited from four schools in North Wales (UK), two for the training group and two wait-list controls. The schools were selected based on socioeconomic status and academic attainment and matched on these criteria. All four schools also expressed interest in training their teachers in mindfulness and providing a mindfulness programme to their students as part of regular school curricula. The first two schools to volunteer for participation in the study were assigned to the training group. Sixth form students (16–18 years) from all schools were recruited after presentations describing the study, and sign-up sheets were placed in sixth-form common rooms. Participants volunteered for questionnaires plus ERP recordings during an experimental task, or only questionnaires. Those participating in the ERP part of the study were allocated a time-slot in January–February (pre-training) and in April–June (post-training). Participation was open to the entire sixth form for control students, and all those enrolled on the mindfulness course for the training group.

The total sample population was N = 48 (21 training group), however class attendance records led to the exclusion of one training group participant who only attended one mindfulness session. In order to compare self-report and ERP measures, analysis was run only on participants who completed all assessments at both time points. This led to a final sample of 40 students (19 training group, mean [M] age 16.8, standard deviation [SD] 0.6), as several participants had opted to only complete self-report measures. Pre-test differences were assessed using independent t tests for age, doctor visits, and sickness absences. A chi-square analysis was also run for gender, past training relating to stress relief, cognitive skills, and mindfulness experience. No significant pre-test differences were identified (all ps > .05). However, age and gender were marginally significant (ps = .06) due to more upper sixth-form Year 13 participants volunteering in the control group (training group M = 16.6, SD = 0.6 / control group M = 17.0, SD = 0.6) and more boys volunteering in the training group. Nevertheless, the difference between 16- and 17-year-olds in developmental terms is minimal (Palluel, Nougier, & Olivier, 2010; Waxes & Morton, 2011).

However, to ensure that the primary mindfulness effects were not impacted, ERP analysis was also run as a hierarchical multiple regression, additionally controlling for age and gender. All other pre-test differences were non-significant (all ps > .1).

2.2 | Measures

The Five-Facet Mindfulness Questionnaire (FFMQ; Baer, Smith, Hopkins, Krietemeyer, & Toney, 2006) assessed whether mindfulness levels increased after training. It has 39 items and five subscales measured on a 5-point Likert scale: Observing, "I pay attention to sensations, such as the wind in my hair or sun on my face"; Describing, "I'm good at finding words to describe my feelings"; Acting with Awareness, "I am easily distracted"; Non-Judging, "I tell myself I shouldn't be feeling the way I'm feeling"; and Non-Reacting, "I watch my feelings without getting lost in them". The final score can be calculated as FFMQ-Total for each individual, or by subscale, with a higher score indicating a more mindful disposition. The internal consistencies (Cronbach's α) for these facets are good, reported as 0.83 for FFMQ-O, 0.91 for FFMQ-D, 0.87 for FFMQ-AwA, 0.87 for FFMQ-NJ, and 0.75 for FFMQ-NR (Baer et al., 2006).

The Perceived Stress Scale (PSS; Cohen, Kamarck, & Mermelstein, 1983) evaluated ability to cope with stress, and has good internal consistency (Cronbach's α = 0.85; Cohen et al., 1983). The PSS has been used previously in mindfulness research, showing reductions in scores in adults (Baer, Carmody, & Hunsinger, 2012) and adolescents (Biegel et al., 2009) after mindfulness training. It consists of 10 items scored on a 5-point Likert scale, where a higher score indicates more perceived stress during the last month.

The World Health Organization, Well-Being Index 5-item version (WHO-5; WHO Collaborating Centre in Mental Health, 1998) can provide a valid measure of depression levels in adolescents (Blom, Bech, Högberg, Larsson, & Serlachius, 2012), and asks how participants have felt over the past 2 weeks. It uses a 5-point Likert scale, and phrases questions positively, for example, "I have felt active and vigorous". It has good internal reliability, with a Cronbach's α of 0.84 (Bech, Olsen, Kjoller, & Rasmussen, 2003).
The Toronto Empathy Questionnaire (TEQ; Spreng, McKinnon, Mar, & Levine, 2009) is a brief self-report measure created on the basis of a factor-analysis of longer empathy questionnaires. The measure has been successfully used with adolescents (Barry, Kauten, & Lui, 2014; Brewer & Kerslake, 2015). The TEQ conceptualizes empathy as primarily an emotional process, for example, “When someone else is feeling excited, I tend to get excited too”. It includes 16 questions, eight of which are reverse coded, and scores are measured on a 5-point Likert scale. The internal reliability with adolescents is good, with Spreng et al. (2009) reporting Cronbach’s alpha ranging from 0.85 to 0.87 across three studies.

An acceptability measure was designed for the study, asking mindfulness-trained students to rate their satisfaction with the mindfulness course and report on home practice. Course enjoyment was measured on a 7-point Likert scale from 1 (not at all) to 7 (very much) and home practice was measured on a 4-point Likert scale. The internal reliability with adolescents is good, with Spreng et al. (2009) reporting Cronbach’s alpha ranging from 0.85 to 0.87 across three studies.

Participants also provided data pre- and post-training regarding the number of sickness absences, the frequency of visits to their local doctor (GP), and the reason for GP visits. Students were free to refuse any questions.

### 2.3 Emotional oddball paradigm

Happy, sad, and neutral faces from the Karolinska database (Goeleven, De Raedt, Leyman, & Verschueren, 2008) were used in the oddball task. Happy (10%) and sad (10%) faces were the target stimuli, including 15 male and 18 female model face pairs (33 oddball targets per emotion), which participants were instructed to respond to by pressing the spacebar on a computer keyboard, using their dominant hand. The frequent (80%) non-target faces were two repeated images of one male and one female model with a neutral expression. Data from the Goeleven et al. (2008) validation paper were used to balance stimuli, with analyses of variance (ANOVA) showing no difference in arousal levels between target emotions (F(1, 62) = 3.0, p = .09), between genders (F(1, 62) < 0.1, p = .91), and no significant interaction between emotion and gender (F(1, 62) = 0.97, p = .33). Mean correct emotion identification scores for selected stimuli were > 75%, based on data from Goeleven et al. (2008).

Face stimuli were presented in the centre of the computer monitor for 900 ms, with an inter-stimulus interval of 750 ms. Participants performed the task in three blocks of 110 trials with the same proportion of the oddball and neutral stimuli in each block, with faces displayed randomly within each block. Block presentation was counterbalanced across participants to control for order effects. The task took 9 minutes to complete, plus breaks between blocks.

### 2.4 Mindfulness-based programme

An age-appropriate mindfulness-based curriculum called ‘b Foundations’ (MiSP; http://mindfulnessinschools.org/) was delivered over eight 50-minute sessions plus an initial orientation session. The course was taught by participants’ regular schoolteachers within the PSHE curriculum. For a full description of the implementation model please see Sanger and Dorjee (2016).

### 2.5 Procedures

This study followed a non-randomized, pre-post design with wait-list control group. Participants were tested individually during school hours, scheduled within independent study periods, using a portable EEG system. Quiet testing spaces were provided on school premises. At pre-test all procedures were explained to participants, and informed consent was obtained before the start of testing. Participants were asked to come to their testing sessions with clean, dry hair and not to use any hair products or conditioner. During the EEG set-up period students filled in self-report measures. If these were not completed during set-up, students took them away in a sealed envelope and were asked to return them on the next school day.

EEG signal was recorded with 36 Ag/AgCl electrodes placed according to the 10–20 standard system, using the right mastoid as the reference site and was re-referenced offline to the algebraic mean of the right and left mastoid (A1), as is recommended in Luck (2005) and standardly used in practice (e.g., Dorjee et al., 2015; Savill & Thierry, 2011). FPz was used as the system’s ground. EEG data were recorded with Neuroscan NuAmps amplifiers with a sampling rate of 1 kHz. Two electrodes placed above and below the right eye monitored ocular movements. In addition, electrodes attached on both forearms recorded heart rate variability; the results of these analyses will be reported elsewhere. Electrode impedance was kept below 7 kΩ. The EEG signal was filtered online with a bandpass filter between 0.01 and 200 Hz, and additional filtering was applied offline using a zero phase shift low-pass filter with a cut-off frequency of 30 Hz, and a 48 dB/Oct slope. ERP data were manually cleaned, rejecting motor and irregular ocular artefacts. An algorithm in Neuroscan Edit software was then employed to regress out eye-blink artefacts, and mathematically remove artefacts using the Gratton, Coles, and Donchin (1983) method. The data were epoched into 1000 ms segments (beginning 100 ms before stimulus onset), and baseline corrected relative to pre-stimulus activity. Averages for each condition and participant, as well as grand averages across participants for each condition and group, were then computed, considering only correct trials, that is, trials in which happy or sad target faces had been correctly detected. ERPs elicited by neutral faces (standard stimuli) were computed from the last neutral face that preceded an emotional target stimulus. This method of standard ERP calculation ensures that standard stimuli were maximally habituated (i.e., maximally standard in nature) and, by the same token, that trial numbers were comparable between oddball and standard conditions. Mean false alarms to standards during this task were minimal, 0.88 (SD 1.1) at pre-test, and 0.85 (SD 1.2) at post-test. Therefore false alarms were not excluded after acquiring only standards that preceded a target trial, as the potential influence of any false alarms that may have been included would be negligible.
**FIGURE 1** Graphs a, b, e, and f show linear derivations (CP2, CP4, P2, P4) of group average waveforms for each of the three conditions. Topographies c and g show P3b amplitude distribution for happy and sad targets respectively. Graph d represents the P3b mean amplitude change averaged across the three conditions with significant decrease in the control group ($p < .004$).

### 2.6 Data analysis

Pre-post questionnaire and oddball task performance measures were analysed using a mixed factorial analysis of variance (ANOVA), with time (pre, post) and group (training, control) as factors. Significant interactions were further investigated using paired samples $t$ tests. Outliers above and below two standard deviations (SD) of the group mean were removed prior to analysis. ERP analysis was conducted over a cluster of centroparietal electrodes identified in a preliminary laterality analysis (CP2, CP4, P2, and P4). Similarly to previous studies investigating P3b responses to emotional stimuli (Keil et al., 2002; Schupp, Junghofer, Weike, & Hamm, 2003; Schupp et al., 2006; Olofsson, Nordin, Sequeira, & Polich, 2008), the distribution was right-lateralized (Figure 1). An omnibus ANOVA with factors of time, group and condition was then conducted on P3b mean amplitudes. Significant interactions were further investigated using paired samples $t$ tests. Greenhouse-Geisser adjustment for degrees of freedom was applied where applicable. Due to the conflicting evidence in the school-based literature regarding the impact of mindfulness training engagement on outcomes (Foret et al., 2012; Huppert & Johnson, 2010), we ran correlation analyses to assess any moderating effects of course attendance, satisfaction, and home practice in the training group. In addition, potential relationships between differential P3b amplitudes (deviant minus standard) and empathy scores across groups were studied. This approach to correlations with self-report measures was also used by Ikezawa et al. (2014) and Fan and Han (2008). Lastly, a hierarchical multiple regression using post-training average P3b mean amplitudes as the dependent variable for all task conditions, controlling for pre-training mean amplitudes and group, then age, then gender, examined whether the marginal group differences in age and gender impacted on the main ERP findings.

### 3 RESULTS

#### 3.1 Five-Facet Mindfulness Questionnaire

The FFMQ had an acceptable level of internal reliability (Cronbach’s $\alpha$) within this sample at pre-test, 0.82 for FFMQ-Total, 0.76 for FFMQ-O, 0.86 for FFMQ-D, 0.83 for FFMQ-AwA, 0.87 for FFMQ-NJ, and 0.68 for FFMQ-NR. Mixed factorial ANOVA results for FFMQ-Total did not reveal any significant effects (all $p > .1$). Subscales were also analysed, but no significant effects were found (all $p > .1$).

#### 3.2 Perceived-Stress Scale

Pre-test reliability for the PSS was good ($\alpha = 0.83$). Two participants did not provide responses on this measure, reducing the sample to $n = 38$. There was a marginal main effect of time ($F(1, 36) = 3.6, p = .07$) but no significant group effect ($F(1, 36) = .2, p = .66$) or interaction ($F(1, 36) = 1.5, p = .23$). A significant positive correlation was found between PSS change scores and self-reported enjoyment of the mindfulness course, suggesting that participants in the training group who reported more enjoyment of the course actually increased in perceived stress ($r = .65, p = .004$). Interestingly, post-hoc analysis found that perceived stress change pre-post did not correlate with change in well-being scores in the training group ($r = -.08, p = .78$), while a
TABLE 1 Pre to post self-report measure changes in training and control groups (* < .05; ** < .01; ~ = trend towards significance in follow-up t test analysis)

<table>
<thead>
<tr>
<th>Questionnaire means (SD)</th>
<th>Pre-training group</th>
<th>Post-training group</th>
<th>Pre-control group</th>
<th>Post-control group</th>
</tr>
</thead>
<tbody>
<tr>
<td>FFMQ-Total</td>
<td>119.9 (15.9)</td>
<td>120.4 (12.7)</td>
<td>121.1 (17.1)</td>
<td>122.3 (20.6)</td>
</tr>
<tr>
<td>FFMQ-O</td>
<td>24.4 (6.3)</td>
<td>24.1 (5.3)</td>
<td>24.6 (5.2)</td>
<td>23.6 (5.3)</td>
</tr>
<tr>
<td>FFMQ-D</td>
<td>26.1 (5.8)</td>
<td>24.6 (4.3)</td>
<td>26.0 (6.3)</td>
<td>26.4 (5.7)</td>
</tr>
<tr>
<td>FFMQ-AwA</td>
<td>24.1 (5.4)</td>
<td>23.3 (4.2)</td>
<td>24.2 (6.9)</td>
<td>23.8 (6.5)</td>
</tr>
<tr>
<td>FFMQ-NJ</td>
<td>25.8 (6.4)</td>
<td>27.5 (5.9)</td>
<td>26.0 (6.6)</td>
<td>27.7 (6.5)</td>
</tr>
<tr>
<td>FFMQ-NR</td>
<td>19.6 (3.4)</td>
<td>20.7 (3.5)</td>
<td>20.4 (4.4)</td>
<td>20.9 (4.5)</td>
</tr>
<tr>
<td>PSS</td>
<td>20.1 (6.7)</td>
<td>19.8 (4.9)</td>
<td>22.5 (6.7)</td>
<td>19.7 (6.0)</td>
</tr>
<tr>
<td>WHO-5</td>
<td>48.0 (16.7)</td>
<td>64.8 (20.4)</td>
<td>55.6 (19.6)</td>
<td>61.1 (16.9)</td>
</tr>
<tr>
<td>TEQ</td>
<td>43.1 (10.8)</td>
<td>39.8 (8.7)</td>
<td>50.5 (7.1)</td>
<td>48.1 (7.0)</td>
</tr>
<tr>
<td>Health-Absenteeism</td>
<td>3.5 (3.7)</td>
<td>3.4 (3.8)</td>
<td>4.9 (5.5)</td>
<td>5.9 (7.5)</td>
</tr>
<tr>
<td>Health-GP Visits</td>
<td>1.1 (2.4)</td>
<td>0.5 (0.7)</td>
<td>0.3 (0.8)</td>
<td>0.5 (0.7)</td>
</tr>
<tr>
<td>Health-Psychological GP Visits</td>
<td>0.4 (1.0)</td>
<td>0.1 (0.2)</td>
<td>0 (0)</td>
<td>0.2 (0.6)</td>
</tr>
</tbody>
</table>

3.3 | WHO-5 well-being index

The WHO-5 was reliable within this sample (α = 0.79). Five participants did not contribute to this measure, reducing the sample to n = 35. The mixed factorial ANOVA revealed a significant main effect of time (F(1, 33) = 12.9, p = .001, η² = .26), with all participants’ well-being increasing by post-test. There was a non-significant effect of group (F(1, 33) < .1, p = .81), and marginally significant time by group interaction (F(1, 33) = 3.1, p = .08, η² = .06). As the interaction was in the predicted direction, follow-up paired samples t tests were conducted, showing the trend to be due to a significant increase in self-reported well-being in the training group (t(15) = -4.3, p = .001, d = 1.1), whilst no significant changes were observed in the control group (t(18) = -1.2, p = .24). The between-group comparison at the post-test was not significant (p > .1)

3.4 | TEQ empathy questionnaire

The reliability of this measure at pre-test was good (α = 0.84). Four participants did not provide responses, reducing the sample to n = 36. The ANOVA showed a significant main effect of time (F(1, 34) = 10.1, p = .003, η² = .23), with participants’ empathy reducing post-test, and a significant main effect of group (F(1, 34) = 8.8, p = .006, η² = .21), indicating that the training group were less empathetic on average. The interaction effect was non-significant (F(1, 34) = .3, p = .61). However, changes in this measure correlated positively with course attendance (r = .66, p = .006) and marginally with home practice (r = .49, p = .06), suggesting that those who attended more of the mindfulness course and practised more displayed larger increases in empathy.

3.5 | Health measures

ANOVAs revealed no change in absenteeism over time (F(1, 38) = .6, p = .45), between group (F(1, 38) = 1.3, p = .25) or an interaction (F(1, 38) = .9, p = .35). GP visits were also not affected by time (F(1, 38) = .6, p = .44) or group differences (F(1, 38) = 1.2, p = .28), the time by group interaction was marginally significant (F(1, 38) = 3.0, p = .09, η² = .07). To investigate the possibility of differential effects for visits due to physical and mental health reasons, GP visits were further broken down accordingly (e.g., asthma and stress, respectively). For mental health-related visits only, the ANOVA showed no change over time (F(1, 38) = .3, p = .58) or group (F(1, 38) = .7, p = .42), but there was a significant time by group interaction (F(1, 38) = 5.0, p = .03, η² = .12). However, follow-up paired samples t tests revealed only trends towards significance, with some reduction in the training group (t(18) = 1.7, p = .11, d = .39) and a non-significant increase in GP visits over time in the control group (t(20) = −1.5, p = .16, d = .32). Nevertheless the effect sizes suggest a small but meaningful association between mindfulness practice and health, and, given the short time-scale this was measured over (8 weeks), it merits consideration. On visual inspection there appeared to be pre-test group differences on GP visits (see Table 1). However, independent t-tests confirmed that these were only marginally significant at pre-test for general and specifically mental health visits (p > .05). This was most likely due to the non-randomized nature of participant recruitment.

3.6 | Task performance

Mixed factorial ANOVAs assessed oddball task performance. There was a significant main effect of time on accuracy to happy target faces (F(1, 38) = 4.8, p = .04, η² = .11), but no significant main effect of group (F(1, 38) = .1, p = .88), and no significant interaction (F(1, 38) = .5, p = .48), suggesting that all participants became more accurate by post-test. A marginally significant positive correlation
was found between mindfulness course enjoyment and improved target accuracy to happy faces ($r = .40, p = .07$). No significant effects of accuracy were found for sad target faces (all $p$s $> .1$). In terms of response time (RT), all target effects were non-significant ($p$s $> .1$) except a marginal group effect for sad targets ($F(1, 38) = 3.29, p = .08, \eta^2 = .08$), suggesting a small increase in RT across groups. False alarms to standard stimuli were minimal, $0.88$ ($SD = 1.1$) at pre-test and $0.85$ ($SD = 1.2$) at post-test over the 267 standard trials. The ANOVA showed no significant main effects or interactions (all $p$s $> .1$).

### 3.7 | P3b ERP analysis

A 2 (time: pre, post) $\times$ 3 (condition: happy, sad, standard) $\times$ 2 (group: training, control) ANOVA assessed any variation in trial numbers included in ERP analysis across conditions and groups. There were no significant main effects of time or group ($p$s $> .1$), but there was a significant main effect of condition ($F(2, 76) = 2503.7, p < .001, \eta^2 = .95$). This was further explored using simple contrasts showing that this difference was significant between standard and oddball conditions ($p < .001$), and marginally significant between positive and negative emotional targets ($p = .06$). Thus, the rejection rates were significantly higher for standards than oddballs and marginally higher for sad than happy faces. As shown in Table 2, the difference between the happy and sad face oddballs only amounted to about one in 32 trials. There were no rejection rate differences for any of the conditions between groups at either baseline or post-test.

The time window for the P3b analyses (420–520 ms) was determined by inspection of the global field power calculated across the scalp, which provided a synthetic view of key change points of power with time. First, the distribution of the P3b was assessed by an analysis of the pre-test mean amplitudes using combined averages for the two oddball conditions and groups, thereby reducing the number of ANOVA factors and thus minimizing the likelihood of spurious effects (Luck & Gaspelin, 2017). The resulting 2 $\times$ 4 ANOVA with factors of laterality (left, right) and electrode (left: CP1, CP3, P1, P3; right: CP2, CP4, P2, P4) revealed a significant main effect of laterality ($F(1, 39) = 15.11, p < .001, \eta^2 = .27$) with greater mean amplitudes on the right ($left M = 6.36 \mu V, right M = 7.71 \mu V, Figure 1$). There was also a significant main effect of electrode ($F(1.70, 66.17) = 8.01, p < .002, \eta^2 = .17$) and a significant laterality-by-electrode interaction ($F(2.26, 88.19) = 6.28, p < .003, \eta^2 = .14$). Therefore, further analyses including time and group factors focused on the linear derivation of the right-sided cluster of four electrodes (CP2, CP4, P2, P4), thus removing the electrode factor from further analyses as recommended by Luck and Gaspelin (2017).

A 2 $\times$ 2 $\times$ 3 ANOVA with factors of time (pre, post), group (training, control) and condition (happy, sad, neutral) assessed possible changes across time in mean amplitudes of the P3b (Figure 1). There was a significant main effect of time ($F(1, 38) = 5.88, p < .03, \eta^2 = .13$) which indicated a significant decrease in mean amplitudes from pre ($M = 5.68 \mu V$) to post-test ($M = 4.93 \mu V$). There was no significant main effect of group ($p > .1$). The main effect of condition was significant ($F(1.63, 61.89) = 88.74, p < .001, \eta^2 = .70$)

### Table 2

<table>
<thead>
<tr>
<th></th>
<th>Pre-happy</th>
<th>Pre-sad</th>
<th>Pre-standard neutral</th>
<th>Post-happy</th>
<th>Post-sad</th>
<th>Post-standard neutral</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Training group</strong></td>
<td>32.1 (.8)</td>
<td>31.6 (1.6)</td>
<td>52.5 (3.5)</td>
<td>31.2 (1.9)</td>
<td>31.1 (1.9)</td>
<td>50.5 (4.0)</td>
</tr>
<tr>
<td><strong>Control group</strong></td>
<td>32.0 (1.5)</td>
<td>31.4 (2.0)</td>
<td>51.4 (4.4)</td>
<td>31.6 (1.3)</td>
<td>31.2 (1.9)</td>
<td>51.8 (3.3)</td>
</tr>
</tbody>
</table>

**FIGURE 2** Correlation plots show the significant positive relationship between P3b mean amplitude changes (post-training minus pre-training in both groups combined to happy (a; $p = .03$) and sad (b; $p = .05$) target stimuli with neutral condition subtracted) and change (post minus pre) in TEQ empathy questionnaire scores.
with sad faces eliciting the largest mean amplitudes (M = 8.07 μV), followed by happy faces (M = 6.75 μV) and neutral faces (M = 1.11 μV). Importantly, the time-by-group interaction was also significant (F(1, 38) = 4.73, p < .04, η² = 0.11). Follow-up analyses revealed that the mean amplitudes across conditions significantly decreased in the control group over time (t(20) = 3.32, p < .004, d = .34), whereas the training group did not show any significant change from pre to post-test (p > .1). None of the remaining interactions were significant (all ps > .1). The ANOVA for latency revealed no significant effects (all ps > .1).

To further investigate the relationship between P3b mean amplitude modulations (post- versus pre-training) and changes in measures of empathy (post minus pre TEQ scores), mean P3b amplitudes to neutral faces were subtracted from the mean P3b amplitudes to happy and sad face oddballs, respectively. This correlation was positive for both happy and sad faces (happy: r = .37, p = .03; sad: r = .33, p = .05; Figure 2).

4 | DISCUSSION

We studied emotional face processing before and after a school-based mindfulness programme in adolescents. Mindfulness training was associated with maintained P3b mean amplitudes to correctly identified happy and sad target faces in an oddball task involving neutral faces as standard stimuli, whereas P3b amplitude was overall reduced over time in wait-list controls. There were no between-group effects on response time or accuracy, but mindfulness course satisfaction positively correlated with improved accuracy for happy faces. Training-based improvements were also noted in self-reported well-being and health measures, with significant increases on the WHO-5 well-being index and trends towards fewer mental health-related GP visits. Although mindfulness training did not impact TEQ scores overall, course attendance and home practice both positively correlated with empathy change. Overall, this study demonstrated that a mindfulness-based PSHE module delivered by internal schoolteachers can positively impact students’ processing of faces and well-being.

Modulation in P3b responses to happy and sad faces also correlated with changes in empathy. This suggests that mindfulness practice may help sustain attentional focus on socially relevant affective stimuli, overriding typical stimulus habituation (Geisler & Polich, 1994; Ravden & Polich, 1998), and potentially indexing heightened empathy. The finding of stable P3b amplitudes elicited by face stimuli, along with self-reported well-being increases, also aligns with our initial predictions and is consistent with previous research, showing that adolescents with MDD tend to show reduced levels of neural activity when processing sad faces (Blom et al., 2015). In addition, research supporting the ECI model of depression showed that many adults with MDD have a general impairment in processing affective stimuli manifesting through dampened responses (Bylmsa et al., 2008). In contrast, mindfulness training has been shown to enhance brain activity and mood in adults (Bostanov, Keune, Kotchoubey, & Hautzinger, 2012; Williams et al., 2014). This study was the first to demonstrate the modulation of depression-related ERP markers of emotion processing in adolescents. Tentatively, we can interpret the findings as suggesting that mindfulness practice, which encourages curiosity and exposure to emotion without judgement or reactivity, can help maintain attention on socially relevant stimuli, in comparison to habituation patterns observed in controls. We note that attention shifts indexed by P3b amplitude were observed irrespective of emotional valence and thus could be underpinned by more generic attention modulation resulting from mindfulness training.

The self-report findings support previous research on school-based mindfulness training, showing a significant increase in well-being (Huppert & Johnson, 2010; Metz et al., 2013). The inclusion of health-related data adds to insights from previous school-based interventions, suggesting that mindfulness training may reduce adolescents’ needs to seek mental health advice. Marginal decreases in GP visits for psychological reasons (e.g., stress, trouble sleeping) were found in the training group, as control participants reported slight increases. This divergent pattern of GP visits was supported by a small to moderate effect size, which is important to examine given the limited sample size and timeframe. The timing of data collection may be relevant here, as students were preparing for summer exams and the potential for stress and anxiety would have been high. Thus, mindfulness practice may have had a buffering effect on psychological well-being, manifesting in less need to seek help during a challenging period.

No changes in perceived stress were found between groups although previous school-based and adult intervention studies have found reductions (Baer et al., 2012; Biegel et al., 2009; Metz et al., 2013; Shapiro, Brown, & Biegel, 2007), but the current data showed a positive correlation between perceived stress change scores and mindfulness course satisfaction. Post-hoc, we explored the relationship between PSS and WHO-5 fluctuations. Whilst self-reported change on these measures was strongly correlated in controls, there was no such relationship in training participants. Mindfulness is the practice of attending to the present moment with curiosity and acceptance, allowing people to become aware of their experience without reactivity (Kumar, Feldman, & Hayes, 2008; Goldin & Gross, 2010). These data would suggest that mindfulness practice enhanced students’ affective awareness without impairing well-being, an emotion regulation strategy previously found in adults with anxiety problems learning mindfulness (Goldin & Gross, 2010).

Moreover, no pre-post change was seen in empathy scores, but within the training group both course attendance and home practice positively correlated with increased empathy. We speculate that greater engagement with mindfulness practice could be required for self-reported empathy change. Interestingly, P3b mean amplitude to targets, which were significantly modulated by mindfulness training, also correlated with empathy scores. It may be that ERP measures operate as a more sensitive measure of changes in social cognition and it is only with prolonged, or potentially more experienced guidance in, mindfulness practice that self-reported changes in empathy would be found.

The lack of an active control group is a limitation of this study, as the experience of a new curriculum in school could have impacted students. However, the PSHE curriculum in Wales for 16+ students does...
cover morality, well-being, spirituality, and metacognition (Dept for Children, Education, Life-long Learning and Skills, 2008) and students are taught new topics throughout the school year. A schoolteacher specializing in a different subject teaches PSHE, and it has an atypical set-up involving active discussion and group work in a more relaxed environment than typical lessons. So, although novelty was not controlled for, in many other respects the standard PSHE curriculum can be considered a suitable active control for a mindfulness-based course.

It should also be acknowledged that in this study it was the first time that the teachers who were involved had delivered a mindfulness-based curriculum, and after only 9 months of personal practice. While all teachers were considered ready by an experienced mindfulness trainer (in terms of readiness to deliver the programme), they remained relatively inexperienced. This limitation of our study can also be construed as reflecting the conditions of implementation to be expected in the real world.

The positive correlation between mindfulness course satisfaction and improved target accuracy to happy faces could have resulted from increased performance motivation. However, three other results go against motivation bias. First, there was no difference in RT and target accuracy between groups at pre- or post-test. Second, course satisfaction and perceived stress correlated significantly. Third, the marginal decrease over time in GP visits was a naturalistic behavioural effect, not a performance-based outcome, and is unlikely to have been impacted by motivation bias. Nevertheless, future studies should take motivational bias into consideration and include manipulations that can reduce its effects, such as an active control group with a second intervention.

The limited sample size for self-report measures must be considered. Although using the same sample for self-report and ERP measures ensures a more accurate comparison between assessments, it constrains the likelihood of finding significant differences between groups. However, effect sizes arguably better reflect the impact of mindfulness training on adolescents, as these are not so constrained by sample size (Cumming, 2012). Another limitation was the lack of follow-up results in this study, which is a common problem in neuroscientific intervention studies, for logistical and cost-related reasons. We hope that future work will build on these initial findings and explore the long-term effects of mindfulness training in adolescents. Of particular interest could be improvements in self-reported well-being and empathy after continued practice, as well as sustained changes in ERP indices of emotional processing predictive of anxiety and depression. These should provide more conclusive evidence of possible long-term effects of mindfulness training delivered within a school context. Future investigations should map the trajectory of how empathy might develop with prolonged mindfulness practice, or how learning from teachers with more experience might reflect on outcomes for their students.

4.1 Conclusion

To our knowledge, this non-randomized controlled study was the first neurocognitive investigation of longitudinal modulation in face processing in adolescents resulting from mindfulness training delivered by schoolteachers as part of a regular school curriculum. The results show that mindfulness training can maintain participants’ attention and associated exposure to socially relevant stimuli, while improving self-reported well-being. Mindfulness practice may also have a buffering effect on stress-related illness, as suggested by the small modulations in GP visits over 8 weeks. Overall, these findings suggest that mindfulness training delivered as part of the school curriculum might be effective in improving the well-being of older school-aged students during a period of heightened stress and depression vulnerability. The study also highlights the potential of neuroscientific methods in contributing to our understanding of mindfulness effects in education.

REFERENCES


How to cite this article: Sanger KL, Thierry G, Dorjee D. Effects of school-based mindfulness training on emotion processing and well-being in adolescents: evidence from event-related potentials. *Dev Sci*. 2018;e12646. https://doi.org/10.1111/desc.12646