ORIGINAL ARTICLE

Are individuals with higher psychopathic traits better learners at lying? Behavioural and neural evidence

R Shao[1,2] and TMC Lee[1,2,3,4]

High psychopathy is characterized by untruthfulness and manipulativeness. However, existing evidence on higher propensity or capacity to lie among non-incarcerated high-psychopathic individuals is equivocal. Of particular importance, no research has investigated whether greater psychopathic tendency is associated with better 'trainability' of lying. An understanding of whether the neurobehavioral processes of lying are modifiable through practice offers significant theoretical and practical implications. By employing a longitudinal design involving university students with varying degrees of psychopathic traits, we successfully demonstrate that the performance speed of lying about face familiarity significantly improved following two sessions of practice, which occurred only among those with higher, but not lower, levels of psychopathic traits. Furthermore, this behavioural improvement associated with higher psychopathic tendency was predicted by a reduction in lying-related neural signals and by functional connectivity changes in the frontoparietal and cerebellum networks. Our findings provide novel and pivotal evidence suggesting that psychopathic traits are the key modulating factors of the plasticity of both behavioural and neural processes underpinning lying. These findings broadly support conceptualization of high-functioning individuals with higher psychopathic traits as having preserved, or arguably superior, functioning in neural networks implicated in cognitive executive processing, but deficiencies in affective neural processes, from a neuroplasticity perspective.

INTRODUCTION

It is widely recognized that individuals with high levels of psychopathy manifest greater tendencies to lie,[1] a disposition which is a core aspect of Machiavellian traits in the Psychopathic Personality Inventory-Revised (PPI-R).[2] Although psychopathy is present in ~ 15–20% of criminal offenders,[3] it also exists in the general population and is sometimes conceptualized as a continuous personality trait.[4,5] While the majority of psychopathy research has utilized offender populations, some recruited non-incarcerated community samples, while others involved university students with varying levels of psychopathic traits.[6] In this context, it is important to note that trait, behavioural or neurobiological differences may exist among those three types of 'high-psychopathy' individuals.[6–8] Here, we mainly focused on university students with higher psychopathic traits who may also be considered as having relatively high functioning levels.

A body of research indicates higher psychopathy is associated with reduced functioning of paralimbic circuitries including the ventromedial prefrontal cortex (VMPFC), the hippocampus and the amygdala, possibly accounting for these individuals' aberrant affective and social processing,[6,9–12] such as deficient identification of negative facial and vocal emotions observed in incarcerated[13,14] and non-offender[15] psychopaths, as well as reduced physiological reactivity to negative stimuli in students scoring higher in psychopathy.[16] VMPFC dysfunctioning in high-psychopathy individuals may also disrupt normal associations between actions and negative consequences caused to other people, thus hampering those individuals' acquisition of social

moral norm,[10] such as observed in university students with higher psychopathic traits.[12]

On the other hand, psychopathic tendencies such as deficient affective processing may not be associated with damage to the dorsal and lateral prefrontal or parietal cortices.[9,17] Consistent with this, community samples with psychopathic characteristics exhibit preserved prefrontal grey matter volume[18–20] and enhanced prefrontal connectivities,[21,22] as well as superior set-shifting abilities and intact IQ.[7] Also, limited evidence indicates that among student samples, higher psychopathic tendencies may be associated with intact or even superior executive functions, such as selective attention/inhibition,[23,24] processing speed/mental flexibility[25] or in general.[26] However, due to the small number of studies, heterogeneity in psychometric and task measures and existence of negative findings, the association between psychopathic traits and executive functions needs to be considered tentative.[8]

Lying is a ubiquitous social phenomenon common to daily life. Past research has uncovered significant insight into the cognitive and neural processes underpinning lying, specifically how an individual mentally suppresses the 'true' information while deriving and conveying counterfactual alternatives.[27–29] The common belief is that lie-telling recruits additional cognitive processes (for example, inhibition, working memory manipulation, attention switching processes) compared to truth-telling, accounting for the longer reactions times (RTs) needed to lie.[30–33] In addition, lie-telling likely invokes greater affective conflict.[27,34] Consistent with the cognitive theory of lying,[35] the inferior frontal

[1]Laboratory of Neuropsychology, Department of Psychology, The University of Hong Kong, Hong Kong, Hong Kong; [2]Laboratory of Cognitive Affective Neuroscience, Department of Psychology, The University of Hong Kong, Hong Kong, Hong Kong; [3]The State Key Laboratory of Brain and Cognitive Sciences, The University of Hong Kong, Hong Kong and [4]Institute of Clinical Neuropsychology, Department of Psychology, The University of Hong Kong, Hong Kong, Hong Kong. Correspondence: Professor TMC Lee, Laboratory of Neuropsychology, The University of Hong Kong, Rm 656, Jockey Club Tower, Pokfulam Road, Hong Kong, Hong Kong.
E-mail: tmclee@hku.hk

2

gyrus (IFG)–insula circuitry heavily involved in inhibitory control and action selection,[30,36,37] the dorsolateral frontal and parietal cortices consistently associated with complex working memory manipulation, attention and cognitive control,[30,36,38] and the anterior cingulate cortex (ACC) linked with conflict/interference resolution, performance monitoring and dynamic social motivation processing along with its key roles in value-based choice-making,[36–40] are all recruited during lying.[27,30,32,33,41,42–45] Also, the cerebellum involved in learning about stimulus–response associations and in cognitive and emotion control processes,[38,46–48] and the VMPFC implicated in affective reactions and conflict elicited by moral transgression,[49] are recruited during lying.[30,32,41,50] Thus, neurobiological evidence indicates that lying involves a cascade of coordinated cognitive-affective processes.

Only a few behavioural studies have assessed the trainability of lying, and inconsistent findings have been reported.[51–54] Research on the training effect on lying-related neural processes is even more limited,[51] except for evidence indicating that cerebellar activity is reduced after repeated stimulus–response associations are learned.[47,55,56] Understanding whether and how lying processes may be 'rewired' will critically extend the existing approach, which treats lying as a constant process across time. As discussed, students with higher psychopathic traits may be characterized by normal or even superior cognitive capacities coupled with aberrant socio-affective processing, based on which one may predict that those individuals are more capable at lying. However, existing evidence on superior ability to lie among non-incarcerated community and student high-psychopathy samples is equivocal.[31,42] Critically, no research has investigated whether high-functioning individuals showing higher psychopathic tendencies are more 'trainable' in lying, reflected by greater improvement in lying performance following practice, compared to those scoring lower in psychopathy.

This study employed a longitudinal design involving participants with varying levels of psychopathy. All participants performed a directed lie task on face familiarity judgement,[33] which was modified to delineate the cue and response-related lying signals before and after training. We aimed to differentiate the response preparatory stage that involves complex working memory manipulations, from the subsequent response selection stage, which mostly involves inhibition and memory retrieval processes.[30,48] For example, existing evidence indicates important roles of the dorsolateral prefrontal cortex and the parietal cortex in cognitive preparatory processes,[57,58] whereas the ACC may be principally involved in response monitoring.[57] The role of the VMPFC in signalling affective responses to the lie-signalling cue, and that of the cerebellum in mediating the practice effect on lying response preparation and initiation,[47,48] were also investigated. Given community and student samples with higher psychopathic traits show reduced activities in prefrontal and paralimbic neural circuitries when engaging in socially defiant behaviours,[11,12] and that increased cognitive-motor performance was found to be associated with decreased activities in fronto-cerebellar circuitries,[47,59] we formed the following a priori hypotheses: (1) higher psychopathy traits would be associated with greater improvements in lying speed upon training; (2) individuals scoring higher in psychopathy would show larger decrease in neural activities following training, during both cue-processing (the dorsolateral frontal cortex, parietal cortices, VMPFC and cerebellum) and response-making phases (the VLPFC–insula, ACC and cerebellum); and (3) individual participants' training effects on behavioural and neural changes during lying would be correlated.

MATERIALS AND METHODS

Participants
Ethical approval was granted by The University of Hong Kong. All participants gave written informed consent for participation. Recruitment

consisted of two stages. In Stage 1, 120 Chinese participants were recruited from the University of Hong Kong. Each potential participant needed to team up with at least five same-gender friends whom the participant had known for more than 6 months. Such requirements were for obtaining photos of personally familiar faces for each participant. Stage 1 participants were screened on the Chinese version of the PPI-R,[2] which is a 154-item self-report questionnaire that can assess psychopathic traits among nonclinical, non-forensic populations,[4,5] and showed good internal consistency and construct validity.[2] Given considerable controversies exist for the sub-factor structure of the PPI-R,[60,61] and previous claim that psychopathy could be an emergenic disorder arising from combination of its sub-facets,[4] we classified participants based on their total PPI-R scores (formed by summing raw scores of all eight subscales, see Supplementary Information). In the current sample, the Cronbach's Alpha for the total score was high (0.858). We obtained a high-PPI group scoring top 25% of the total sample (raw score ⩾ 318), and a low-PPI group scoring bottom 25% (raw score ⩽ 279). These participants entered Stage 2 screening. All participants were free of past or present major physical or psychological illnesses. In total, 23 low-PPI and 29 high-PPI participants were included (Table 1). Two participants were subsequently excluded from behavioural analyses and two others were excluded from imaging analyses, leaving 50 participants for the behavioural analyses (25 females, 22 low-PPI) and 48 participants for the imaging analyses (24 females, 22 low-PPI). Further details about participant characteristics and exclusions are included in Supplementary Information.

Task and materials
During Stage 1, a digital photo was taken for each participant in a well-lit room. The photos of 60 age- and ethnicity-matched unfamiliar faces (30 females and 30 males) were downloaded from online sources. The unfamiliar faces also matched closely with the familiar faces on pictorial aspects and rated by independent judges to ensure neutrality and non-familiarity (Supplementary Information).

Each participant performed a modified directed lie paradigm (DLT).[33,62] In each trial, based on the cue received, the participant had to make either an honest or dishonest response on whether they 'know' or 'do not know' the face being presented (Figure 1a). Each task trial began with a 2-s pictorial cue denoting whether an honest or dishonest response was required. Following a variable delay, a photo containing either a familiar or an unfamiliar face of the same gender as the participant was presented, and the participant needed to respond as quickly and accurately as possible by pressing one of two keys. There was no time limit for responding, but RTs that were too long (> 5 s or 4 s.d.'s from the session mean) were deemed as outliers and excluded. The DLT contained 30 honest and 30 dishonest trials delivered in pseudo-random sequences in the initial and testing sessions. Half of the honest and dishonest trials presented familiar faces, and the other half presented unfamiliar faces. During the two training sessions, the task contained 120 trials, formed by doubling-up the initial task.

The initial and testing sessions also contained 16 trials of a visuo-spatial response control task (CT) that also required working memory maintenance, selective attention and acquisition of stimulus–response associations, but not lying (Figure 1b). This task provided an explicit baseline against which DLT neural activations could be evaluated. The trial timeline of the CT was identical to that of the DLT, except that the possible cues were either an upward-pointing or downward-pointing arrow. On the response screen, the words 'Known' and 'Unknown' were presented one above the other (balanced across trials). Participants needed to respond in consistency with the word whose relative spatial position matched the direction of the arrow. CT trials were randomly interspersed with the DLT trials. Further details about the task paradigms are included in Supplementary Information.

Procedure
Each participant underwent 3 study days spanning no longer than 6 days. On day 1, participants first rated the familiarity level and emotion intensity of all the familiar and unfamiliar faces (presented in pseudo-random orders). Participants then received the instructions and practice on the DLT and CT, before completing the initial DLT and CT inside the functional magnetic resonance imaging scanner. Following task completion and a 20-min break, the participant completed the training 1 DLT outside the scanner. On day 2, the participant completed the practice task and the training 2 DLT outside the scanner. On day 3, the participant firstly rated

Table 1. Participants' (1) demographic information, (2) PPI-R total and factor scores, (3) initial-to-testing time interval and (4) behavioural task performance

	Total	Low-PPI	High-PPI	Comparison
Age	20.25 ± 1.58	20.04 ± 1.43	20.41 ± 1.70	$P = 0.407$
Gender (F/M)	26/26	12/11	14/15	$P = 0.782$
PPI-R-total	306.23 ± 38.72	267.48 ± 17.37	336.97 ± 16.78	$\boldsymbol{P < 0.001}$
PPI-R-FD	113.23 ± 20.67	96.26 ± 15.20	126.69 ± 13.19	$\boldsymbol{P < 0.001}$
PPI-R-ScI	158.44 ± 21.84	140.04 ± 12.99	173.03 ± 15.41	$\boldsymbol{P < 0.001}$
PPI-R-CH	34.56 ± 5.63	31.17 ± 5.07	37.24 ± 4.56	$\boldsymbol{P < 0.001}$
Study Interval	3.02 ± 1.11	2.78 ± 1.00	3.21 ± 1.18	$P = 0.174$
Error: Initial	1.54 ± 1.72	1.27 ± 1.96	1.75 ± 1.51	$P = 0.334$
Error: Testing	1.80 ± 1.70	1.71 ± 2.12	1.88 ± 1.32	$P = 0.728$
Total Excluded: Initial	2.68 ± 1.83	2.18 ± 2.08	3.07 ± 1.54	$P = 0.089$
Total excluded: Testing	3.36 ± 2.23	2.96 ± 2.40	3.67 ± 2.08	$P = 0.269$

Abbreviations: PPI-R, Psychopathic Personality Inventory-Revised; PPI-R-CH: the 'coldheartedness' factor of the PPI-R; PPI-R-FD, the 'fearless dominance' factor of the PPI-R; PPI-R-ScI, the 'self-centred impulsivity' factor of the PPI-R. The data are shown separately for the total participant sample ($n = 52$) and for the low-PPI and high-PPI groups. For the behavioural measures (frequency of error and excluded trials), the data are shown for the remaining 50 participants after excluding one male high-PPI participant due to abnormal picture familiarity ratings and one female low-PPI participant for excessive excluded trials in training session 2 (see main text). Both the mean and standard deviation (s.d.) are shown. Statistical values for group comparisons were generated by independent-samples t-tests (for continuous variables) or Mann–Whitney U-tests (for gender). P-values for significant group differences are highlighted in bold. Age is in years. For gender, the first value denotes the number of females, and the second value denotes the number of males. Study interval is in days.

the familiarity levels of the familiar and unfamiliar faces again, and then completed the practice task and testing DLT and CT inside the functional magnetic resonance imaging scanner. Participants were then debriefed, thanked and paid HKD 600 as remuneration.

Behavioural analysis

Analyses of participants' error frequency and outlier rate are included in Supplementary Information. Participants' RTs to the pictures were subjected to a random-intercept linear regression model implemented in MLwiN.[63] To improve the normality of the data, the RTs were transformed using the natural logarithm (ln). The hierarchical regression model had two levels (level 1 = individual trial, level 2 = participant). The first stage of the analysis assessed the main effects of PPI group, gender, session (initial vs testing), response type (honest vs dishonest) and picture type (familiar vs unfamiliar). The second stage assessed all the two-way interactive effects of the Stage 1 variables, with a focus on the effects of response type × PPI and response type × session. The third stage assessed all the three-way interactive effects, with a focus on the effects of response type × session × PPI. To assess whether changes in RT on the DLT were due to changes in RT on the CT, we additionally entered participants' between-session CT RT change into the model. The statistical thresholds for the behavioural analyses were set as $P < 0.05$, two-tailed.

Imaging acquisition and analysis

For both the initial and testing sessions, a total of seven-hundred sixty $3.5 \times 3.3 \times 3.5$ mm^3 T2*-weighted echo-planar images (slice number/TR/TE/flip angle = 39/2000 ms/30 ms/90°, matrix = 68×70, FOV = 230×230 mm^2) were collected. Anatomical T1-weighted images were acquired (155 sagittal slices, TR/TE/flip angle = 6.9 ms/3.2 ms/8°, matrix = 240x240, FOV = $240 \times 240 \times 155$ mm^3; voxel size = $1 \times 1 \times 1$ mm^3) for coregistration with the EPI. Image processing was carried out using SPM12 software. Preprocessing included correction for slice timing and head motion, normalization to Montreal Neurological Institute space and smoothing using 6-mm kernel (see Supplementary Information for more details).

All events of interest were modelled as impulse functions and convolved with a canonical haemodynamic response function for modelling of blood oxygen level-dependent signals. At the participant level, for each of the initial and testing sessions, a general linear model included within-subjects factor of response type (honest vs dishonest) for cue-elicited activity, and factors of response type and familiarity (familiar vs unfamiliar) for face-elicited signals. Cues and face stimuli for the CT trials, outlier trials and error trials were separately modelled. Six motion regressors and two run-specific mean regressors were additionally incorporated. The contrasts of *a priori* interest pertaining to the main effect of response type (honest > dishonest, dishonest > honest) were then computed along with other

contrasts assessing the effect of face familiarity (see Supplementary Information and Supplementary Table T4).

The same contrast estimates for the initial and testing sessions were then averaged and contrasted in a voxel-wise manner. The resultant images were subjected to a two-sample t-test to compare the effect of response type and training in the high-PPI and low-PPI groups. Gender was entered as a covariate of no interest. Whole-brain activation threshold was set at $P < 0.005$ at the voxel level and $P < 0.05$ at the cluster level, corresponding to a cluster size of $\geqslant 94$ voxels for the main analyses and $\geqslant 81$ voxels for the generalized psychophysiological interaction (gPPI) analysis (see below), based on Monte-Carlo simulations performed using Alphasim implemented in REST v. 1.8.[64] Please see Supplementary Information for validation of such correction technique.

At the group level, we primarily focused on *a priori* regions of interest (ROIs). For cue-elicited activities, the ROIs included the parietal lobules, the dorsolateral frontal cortex, the VMPFC and the cerebellum. For picture-elicited activities, the ROIs included the ventrolateral prefrontal cortex/insula (VLPFC–insula), the ACC and the cerebellum. Anatomical masks were constructed using WFU_Pickatlas software based on Talairach Daemon atlas. All masks were bilateral (please refer to Supplementary Information for further details about ROI construction). ROI-based small volume correction tests were then conducted with similar Alphasim-based cluster-thresholding procedures as whole-brain analysis, while setting the search spaces to those of the individual ROI masks. These procedures generated size thresholds for clusters that fell in the predefined ROIs, as described above. To further characterize the signals in those ROIs, we extracted parameter estimates (betas) from the significant regions within the ROIs to the effect of response type, response type × PPI group and response type × session × PPI group, and subjected those values to follow-up analysis of variance using SPSS v. 20 (IBM, Armonk, NY, USA). The analysis of variance incorporated the within-subject factors of session, response type and familiarity (for picture-elicited activities) and the between-subject factors of PPI group and gender. Moreover, correlation analyses were conducted between the neural parameter estimates and behavioural β-values under the same effect. Statistical thresholds for the ROI analyses were set as $P < 0.05$, two-tailed. For all ROI analyses, we conducted multiple-testing correction for the number of ROIs, using the Holm–Bonferroni sequential correction.[65] Corrected P-values (P_{corr}) are thus reported for each analysis. As all our ROIs were predefined and theory-driven, those results which were significant but did not survive multiple-testing correction were also reported, but were discussed with caution.

Finally, we conducted gPPI analyses[66] based on 6-mm spherical seeds centred at the locus of maxima of significant clusters to the response type × PPI and response type × session effects that fell in the *a priori* ROIs, to elucidate functional connectivity patterns during making dishonest vs honest responses, which might be different between initial and training
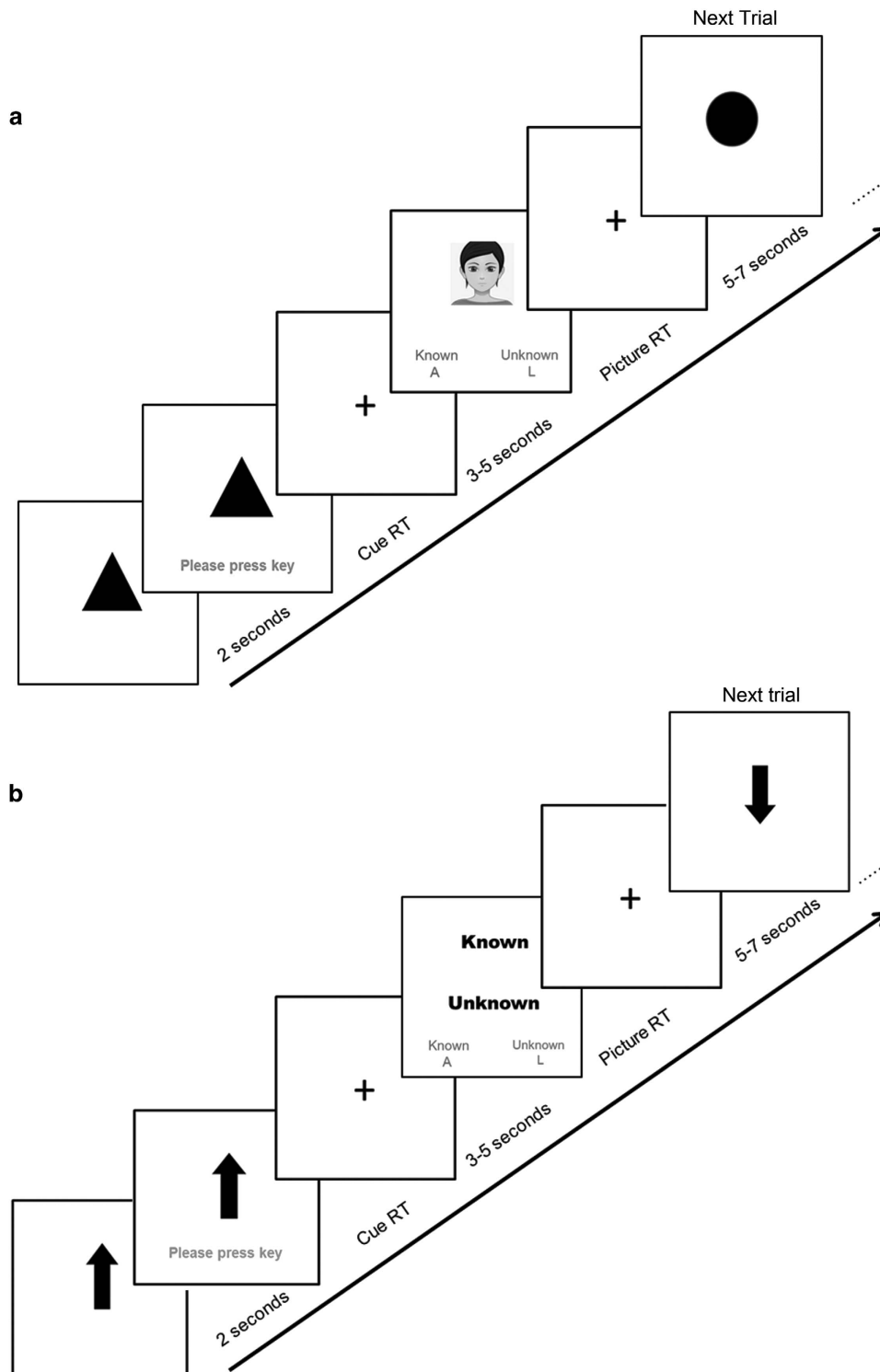
Figure 1. Trial structure of the DLT (**a**) and the CT (**b**). In the DLT, on each trial, participants were first presented with a circle/triangle-shaped cue, indicating whether an honest or dishonest response was required. Following 2 s and upon the appearance of the text prompt, the participants pressed a key to continue. Following another delay of 2 s (for the practice and training tasks) or 3–5 s, Poisson distributed (for the initial and testing tasks), a familiar or unfamiliar face was presented, and participants needed to make either an honest or dishonest response as fast and accurately as possible, by indicating whether the face belonged to a known or unknown person. The trial ended immediately following the participants' responses, and the new trial commenced following no delay (for the practice and training tasks) or a jittered ITI of 5-7 s (for the initial and testing tasks). In the CT, the timeline was identical, except the cue was either an upward-pointing or downward-pointing arrow. The response screen displayed the words 'Known' and 'Unknown' one above the other, whose relative on-screen positions were equal-balanced across the 16 CT trials. Participants needed to make a response consistent with the word in the position that matched the direction of the arrow. Faces presented in the practice tasks did not appear in the actual tasks. Please note an avatar image rather than the actual face photo used in the task is included in this figure for protecting personal information. CT, control task; DLT, directed lie paradigm; ITI, inter-trial-interval; RT, reaction time.
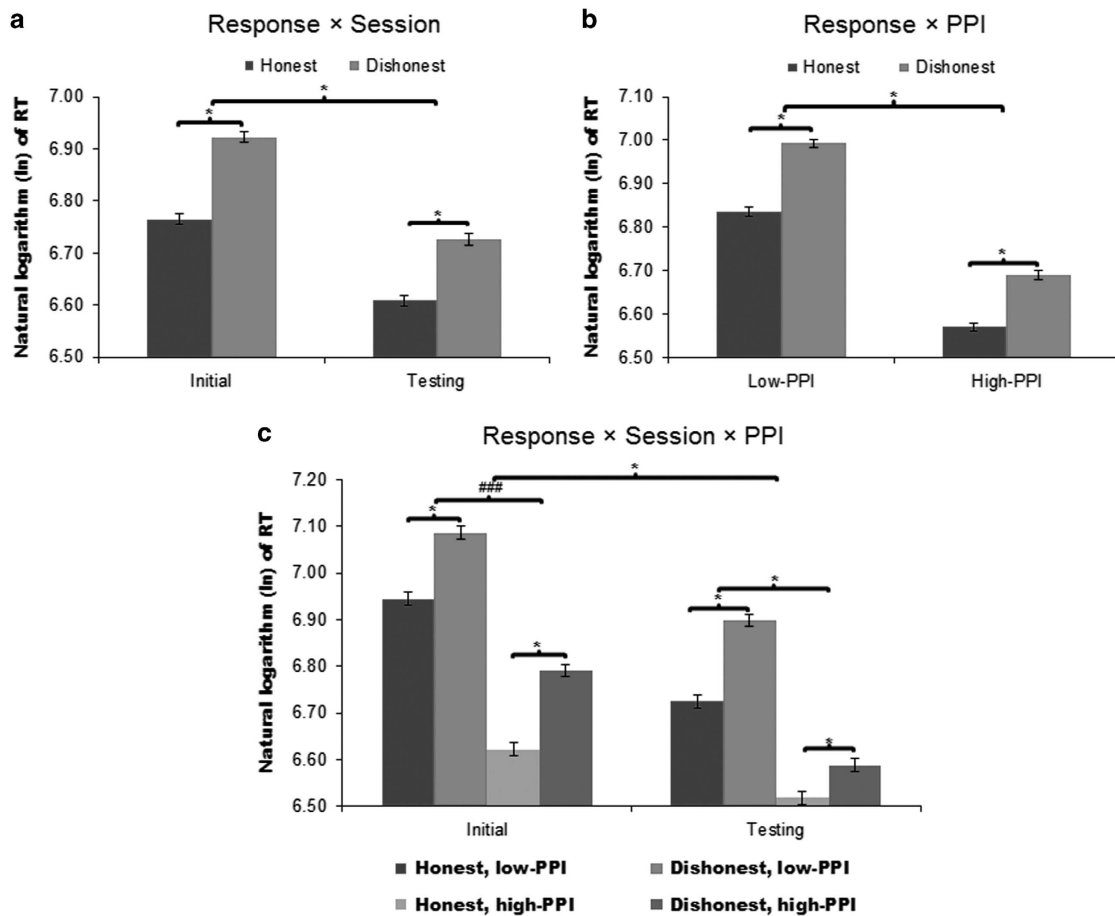
Figure 2. Participants' reaction times (RTs) to the face stimuli. Face RTs were significantly modulated by the effects of (**a**) response type × session (**b**) response type × PPI group and (**c**) response type × session × PPI group. The RTs were transformed using the natural logarithm (ln) to improve normality. The untransformed RTs (in milliseconds) can be calculated using the formulae $RT_{original} \approx 2.718 \times RT^{transformed}$. The error bars represents ± 1 standard error of the mean. * Indicates significant main or interactive effects at $P < 0.05$. ### Indicates insignificant effects.

sessions for low-PPI and high-PPI participants. We also conducted correlation analyses between gPPI connectivity estimates and participants' behavioural RTs.

Figure construction and adjustment were carried out using Powerpoint 2013, and Figure layout arrangement, dimension/resolution setting and formatting were conducted using Adobe Photoshop 12.0 (Adobe Systems, San Jose, CA, USA).

RESULTS

Behavioural analysis

Participants' face RTs were longer for dishonest vs honest responses ($\beta = 0.135$, $z = 16.88$, $P < 0.001$), an effect that was further modulated by the session ($\beta = -0.04$, $z = -2.5$, $P = 0.012$), such that RT increase when making dishonest vs honest responses was larger during the initial session ($\beta = 0.257$, $z = 14.31$, $P < 0.001$) than during the testing session ($\beta = 0.192$, $z = 10.54$, $P < 0.001$) (Figure 2a). The effect of response type was also modulated by PPI group ($\beta = -0.042$, $z = -2.63$, $P = 0.009$), such that the increase in RTs when making dishonest vs honest responses was larger in the low-PPI ($\beta = 0.263$, $z = 13.83$, $P < 0.001$) compared to high-PPI participants ($\beta = 0.195$, $z = 11.28$, $P < 0.001$) (Figure 2b). Furthermore, the three-way interactive effect of response type × session × PPI group was significant ($\beta = -0.131$, $z = -4.09$, $P < 0.001$). Specifically, the response type × session effect was significant only in the high-PPI group ($\beta = -0.084$, $z = -4.88$, $P < 0.001$) but

not in the low-PPI group ($\beta = 0.025$, $z = 1.31$, $P = 0.189$) (Figure 2c). Thus, only high-PPI, but not low-PPI, participants showed reduction of lying-related RT gain following training. See Supplementary Information and Supplementary Figure S1 for further behavioural analyses and discussions.

Imaging analysis

Small volume correction analyses. Although we primarily focus on results of the ROI-based small volume correction analyses in this section, whole-brain findings are detailed in Supplementary Table T1 (for cue-elicited activities) and in Supplementary Table T2 (for face response-elicited activities).

Averaging across sessions and PPI groups, dishonest vs honest cues elicited significant parietal lobule cluster(s) centred at (18, −60, 60) and (42, −36, 33), which however did not survive Holm–Bonferroni correction ($P_{corr} = 0.14$ and 0.08, respectively). No significant activations were observed for the reverse contrast. Between-group comparison revealed that low-PPI participants showed greater activations for the dishonest > honest cue contrast than high-PPI participants in the parietal cortices (36, −69, 48) and VMPFC (0, 51, −3), which also did not survive Holm–Bonferroni correction ($P_{corr} = 0.12$ and 0.12, respectively). When comparing testing vs initial session, neural activities for the dishonest > honest cue contrast showed significant reductions in the dorsolateral frontal cortex (51, 15, 30; −42, −3, 51)

6

($P_{corr} = 0.043$ and < 0.001) and the cerebellum ($-27, -60, -27; 27, -48, -21$) ($P_{corr} = 0.048$ and 0.03). Lastly, the decrease in activity for the dishonest > honest cue contrast after training was more pronounced in the high-PPI vs low-PPI participants in the dorsolateral frontal cortex ($33, 51, 24; -27, 51, 33$) ($P_{corr} = 0.006$ and 0.012), the VMPFC ($0, 57, -9; 12, 45, -3$) ($P_{corr} = 0.006$ and 0.032) and the cerebellum ($12, -57, -15$) ($P_{corr} < 0.001$).

The main effect of response type (dishonest > honest) generated significant activations to face stimuli in the VLPFC–insula circuitry ($-48, 18, 3; 33, 21, 6$) ($P_{corr} < 0.001$), the ACC ($15, 42, 6$) ($P_{corr} < 0.001$) and the cerebellum ($-33, -60, -30; 33, -69, -24$) ($P_{corr} = 0.002$ and < 0.001). The reverse contrast (honest > dishonest) generated no significant activation. No significant cluster was found for the PPI group × response type effect. Comparing testing against initial session, neural activities for the dishonest > honest response contrast decreased following training in the VLPFC–insula circuitry ($-48, 24, -9; 60, 21, 9$) ($P_{corr} < 0.001$) and the ACC ($0, 39, 24$), the latter of which did not survive Holm–Bonferroni correction ($P_{corr} = 0.09$). Such decreases were more pronounced in high-PPI participants than in low-PPI participants in the ACC ($-6, 36, 12; 12, 36, 9$) ($P_{corr} < 0.001$ and $P_{corr} = 0.001$) and the cerebellum ($-6, -51, -6; 12, -60, -30$) ($P_{corr} < 0.001$).

Additional analyses concerning the effect of face familiarity and the neural responses during performing the CT are included in Supplementary Tables T3 and T4.

Although participants also showed decreases in temporo-parietal, cerebellar and frontal activities for control cues and pictures following training, those decreases were less prominent and involved partially non-overlapping regions as the activity decrease during the DLT, particularly in the lateral PFC. Furthermore, little PPI group effect or PPI group × session effect was observed for the CT neural patterns (Supplementary Information).

ROI and correlation analyses. Although analyses on parameter estimates extracted from significant ROI clusters should not be considered as independent from the whole-brain analyses, they were conducted to characterize interactive effects and to elucidate the behavioural significance of the neural effects. We first looked at cue-elicited activities. Two pieces of important findings emerged. First, within the parietal lobule cluster where cue-related signals were sensitive to the response type × PPI group effect (low-PPI(dishonest > honest) > high-PPI(dishonest > honest)), averaged signals to the dishonest > honest contrast positively predicted participants' lie-related RT increase ($r = 0.302$, $P_{corr} = 0.037$) (Figure 3a). Second, for all three ROI clusters (the cerebellum, dorsolateral frontal cortex and VMPFC) where cue-related signals were significantly modulated by the response type × session × PPI group effect (low-PPI (testing(dishonest > honest) > initial(dishonest > honest)) > high-PPI (testing(dishonest > honest) > initial(dishonest > honest))), signals to the testing (dishonest > honest) > initial(dishonest > honest) effect positively predicted increase in lie-related RT gain following training ($r = 0.396$, 0.378 and 0.351, $P_{corr} = 0.02$, 0.024 and 0.045, respectively) (Figures 3b–d). Overall, highly consistent response patterns were observed across the four ROIs: (1) only low-PPI, but not high-PPI, participants showed significant neural signals to the dishonest > honest cue contrast (parietal cortices and VMPFC); and (2) different or opposite patterns of response × session effects were observed between the low-PPI and high-PPI participants—while high-PPI participants exhibited reduced activities for the dishonest > honest cue effect following training, low-PPI participants showed either no effect (cerebellum) or the reverse interaction pattern (dorsolateral frontal cortex and VMPFC) (Figures 3b–d).

During the face-responding phase, we also noted two important pieces of findings. First, the increased signals in the VLPFC circuitry to the dishonest > honest response were further modulated by session × PPI group effect (low-PPI(testing(dishonest > honest) > initial (dishonest > honest)) > high-PPI(testing (dishonest > honest) > initial(dishonest > honest))) (F (1, 44) = 4.725, $P_{corr} = 0.035$) (Figure 4a). Signals in this region to the testing (dishonest > honest) > initial(dishonest > honest) effect positively predicted increase in participants' lie-related RT gain following training, which nevertheless did not survive Holm–Bonferroni correction ($r = 0.296$, $P_{corr} = 0.082$). Second, similar significant positive correlation was observed for the cerebellum, which showed signals modulated by the response type × session × PPI group effect ($r = 0.329$, $P_{corr} = 0.045$) (Figure 4b). Similar to the cue phase, highly consistent patterns were observed across the three ROIs. Specifically, we observed different or even opposite patterns of the response × session effect in the low-PPI vs high-PPI groups—while high-PPI participants exhibited much larger increases in activity to dishonest vs honest responses during initial phase than during testing, low-PPI participants showed either no significant response type × session effect (VLPFC–insula circuitry) or the reverse interaction pattern (ACC and cerebellum) (Figures 4a and b). Further control analyses confirmed that the above ROI results were largely independent of between-session differences in face familiarity ratings or in CT activities (Supplementary Information).

Exploratory gPPI analyses. Seed regions for the gPPI analyses were constructed for four ROIs that showed significant activities for the response type × session effect during the cue presentation and face responding phases (Supplementary Table T5). In addition, a region of the parietal cortices that showed significant cue-related activities for the response type × PPI group effect was also included (Supplementary Table T5). We established the theoretical validity of the gPPI tests by only examining functional connectivities of the seed regions within *a priori* ROIs and on task event contrasts of *a priori* interest (that is, those involving the effects of PPI group and PPI group × session). That is, the total number of gPPI tests was reduced by only examining gPPI contrasts encompassing the same event contrasts that generated the seed regions, plus the modulation of the PPI group. Notwithstanding, given the relatively high number of gPPI tests conducted, our gPPI results should be considered exploratory as multiple-testing correction was not carried out for these analyses.

Upon cue presentation, the parietal region that exhibited more positive activities for the low-PPI(dishonest > honest) > high-PPI (dishonest > honest) contrast showed more positive functional connectivities with the bilateral thalamus, the right IFG and the right insula during the reverse contrast of high-PPI(dishonest > honest) > low-PPI(dishonest > honest). Moreover, the functional connectivity patterns of the region of the cerebellum revealed by the event contrast (initial(dishonest > honest) > testing (dishonest > honest)) with the left dorsolateral prefrontal cortex, right dorsomedial prefrontal cortex, bilateral middle/posterior cingulate cortex and right visual areas during the reverse contrast (testing (dishonest > honest) > initial(dishonest > honest)) were more positive in the high-PPI relative to the low-PPI group. Furthermore, the cerebellar functional connectivity to this event contrast with all four regions correlated negatively with increase in participants' lie-related RT gain following training ($r \leqslant -0.32$, $P \leqslant 0.027$) (Supplementary Figure S2). No significant functional connectivity results were observed for the face responding phase, at a whole-brain level.

DISCUSSION

The key findings of the current investigation were that student samples with higher and lower psychopathic tendencies were markedly different in how their lying processes changed following two-session training. Although students with higher psychopathic tendency exhibited clear improvement in lying speed along with
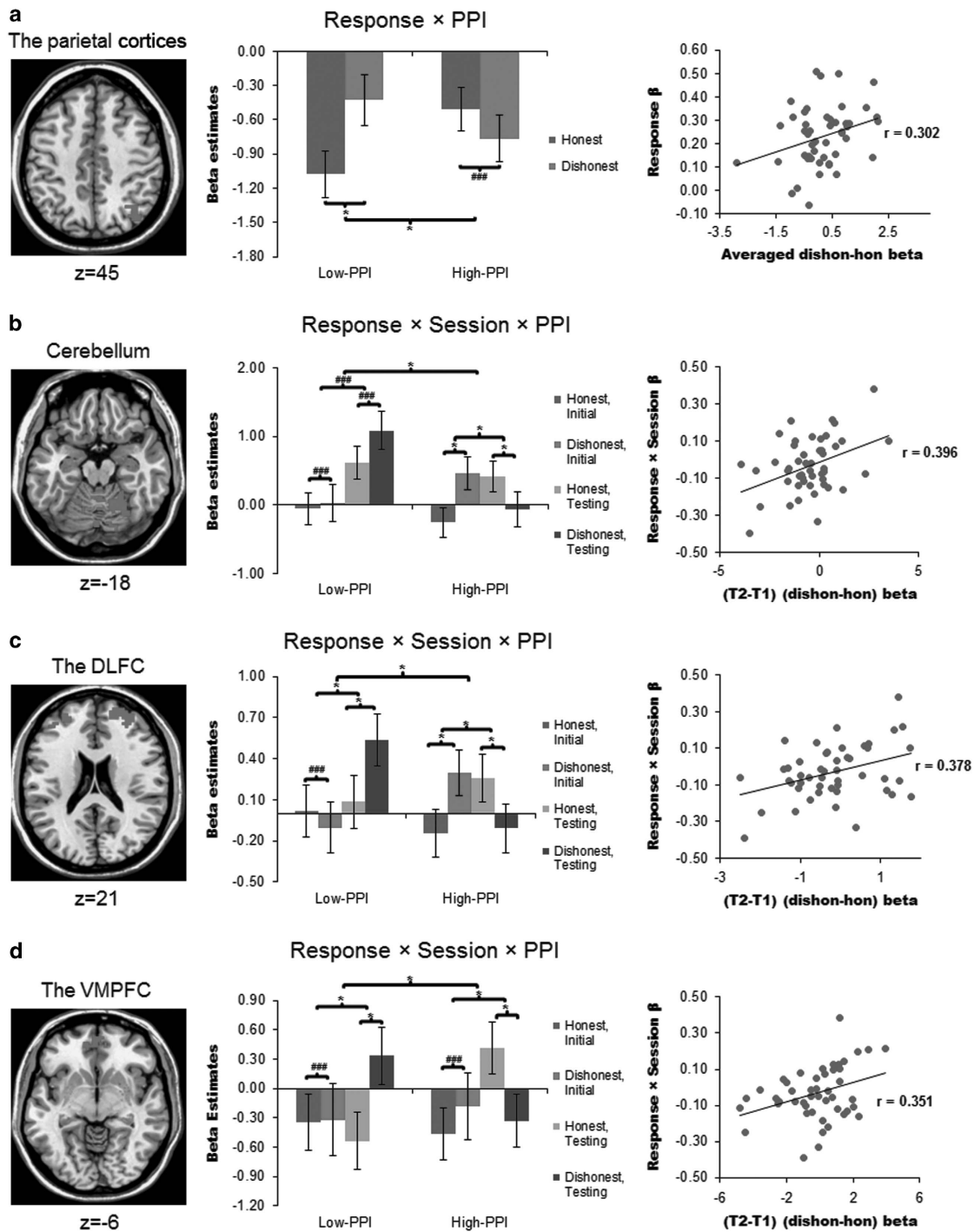
Figure 3. Extracted parameter estimates (β-values) of significant clusters during cue processing for 48 participants (22 low-PPI). Signals were extracted from four *a priori* ROIs: the parietal cortices (**a**), the cerebellum (**b**), the dorsolateral frontal cortex (DLFC) (**c**) and the ventromedial prefrontal cortex (VMPFC) (**d**). Signals during processing honest or lying cues are displayed after subtracting signals during processing control task (CT) cues. This is to display signals of interest against an explicit CT baseline rather than against a more ambiguous 'implicit' baseline, but does not alter the statistical results. The significant clusters in each ROI are overlaid on a standard anatomical template (ch2). MNI z coordinates are provided below the axial slices. Within each scatter plot, the x axis denotes the beta values of neural signals, whereas the y axis denotes standardized β-values on behavioural RTs. *Indicates statistically significant effects at $P < 0.05$, after Holm–Bonferroni correction for the number of ROIs. Parietal cortices: response × PPI $P_{corr} = 0.001$, if low-PPI, response effect $P_{corr} = 0.003$; cerebellum: response × session × PPI $P_{corr} < 0.001$; if high-PPI, response × session $P_{corr} < 0.001$, response effect at T1 $P_{corr} < 0.001$, response effect at T2 $P_{corr} = 0.026$; dorsolateral frontal cortex: response × session × PPI $P_{corr} < 0.001$; if low-PPI, response × session $P_{corr} = 0.005$, response effect at T2 $P_{corr} = 0.003$, if high-PPI, response × session $P_{corr} < 0.001$, response effect at T1 $P_{corr} = 0.02$, response effect at T2 $P_{corr} = 0.026$; VMPFC: response × session × PPI $P_{corr} < 0.001$, if low-PPI, response × session $P_{corr} = 0.02$, response effect at T2 $P_{corr} < 0.001$, if high-PPI, response × session $P_{corr} = 0.02$, response effect at T2 $P_{corr} = 0.008$. ###Indicates insignificant effects. Error bars represent ± 1 standard error of the mean. dishon, dishonest cue; hon, honest cue; MNI, Montreal Neurological Institute; T1, initial session; T2, testing session. Please refer to text for details.
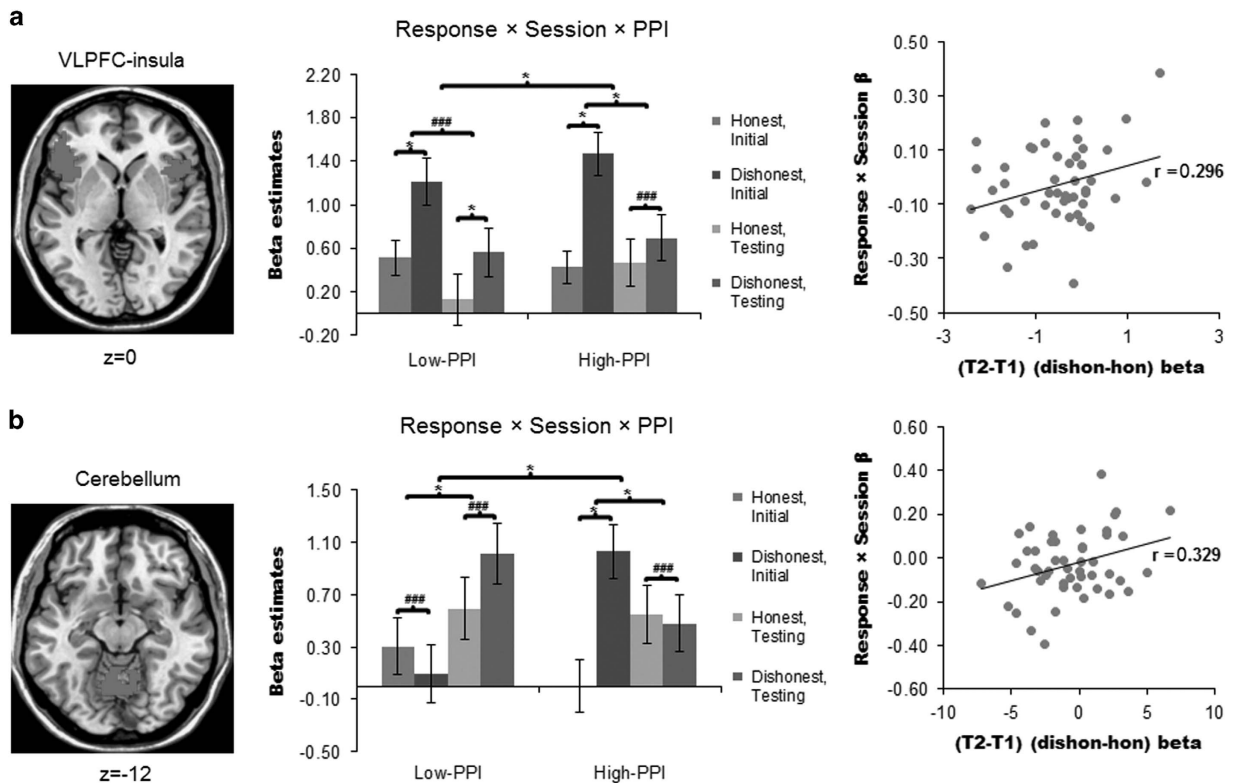
Figure 4. Extracted parameter estimates (beta values) of significant clusters during responding to face stimuli for 48 participants (22 low-PPI). Signals were extracted from two *a priori* regions of interest (ROIs): the ventrolateral prefrontal cortex (VLPFC)–insula circuitry (**a**) and the cerebellum (**b**). Signals acquired while participants made honest or dishonest responses are displayed after subtracting the signals acquired during control task (CT) responses. This is to display signals of interest against an explicit CT baseline rather than against a more ambiguous 'implicit' baseline, but does not alter the statistical results. The significant clusters in each ROI are overlaid on a standard anatomical template (ch2). MNI z coordinates are provided below the axial slices. Within each scatter plot, the x axis denotes the beta values of the neural signals, whereas the y axis denotes standardized β-values on behavioural reaction times (RTs). *Indicates statistically significant effects at $P < 0.05$, after Holm–Bonferroni correction for the number of ROIs. VLPFC–insula: response × session × PPI $P_{corr} = 0.035$, if low-PPI, response effect at T1 $P_{corr} = 0.003$, response effect at T2 $P_{corr} = 0.013$; if high-PPI, response × session $P_{corr} < 0.001$, response effect at T1 $P_{corr} < 0.001$; Cerebellum: response × session × PPI $P_{corr} = 0.012$, if low-PPI, response × session $P_{corr} = 0.03$, if high-PPI, response × session $P_{corr} < 0.001$, response effect at T1 $P_{corr} < 0.001$. ###Indicates insignificant effects. Error bars represent ± 1 standard error of the mean. hon, honest response; dishon, dishonest response; MNI, Montreal Neurological Institute; T1, initial session; T2, testing session. Please refer to text for details.

reduced lie-related neural activities but increased functional connectivity changes in frontoparietal and cerebellum networks, those with lower psychopathic tendency did not improve lying speed, and showed different neural changes following training. The finding that lying about face familiarity was at least partly trainable among individuals with higher psychopathic tendency was observed despite previous research suggesting that face familiarity generally involves largely automatic perceptual and memory processes.[62,67] These novel findings are generally consistent with the conceptualization of high-functioning individuals scoring higher in psychopathy as exhibiting preserved, or tentatively superior, functioning of brain networks implicated in cognitive processing, but reduced reactivity of neural circuitries linked with emotional processing,[6,9] possibly within a framework of uncoordinated cognitive and affective systems.[68]

Our finding that improvements in lying speed following training occurred exclusively in the high-PPI group reconciles the existing inconsistent literature on whether lying performance can be improved through training in the general population,[51–54] and on whether non-criminal university and community samples with higher psychopathic traits are more capable at lying,[31,42] and indicates that psychopathy is a key factor determining training effects on lying. Notably, previous studies on the trainability of lying show large diversity in methodology, including training

length (range 145–540 trials), pre-to-post-training latency (range several hours to 1 week), within- or between-subjects design, and the topic of lying (word recognition or autobiographic information).[51–54] Here we adopted a within-subjects design that is arguably more suited for assessing training effect, and showed that even at a latency spanning 6 days and relatively limited training of 240 trials, the high-PPI group showed clear evidence of RT improvement on lying about face familiarity, which has a social nature and may involve largely automatic processes.[62,67] Further, our findings indicate that it could be the superior ability to learn to lie faster, rather than a naturally higher capacity to lie, is pertinent to the lie-related characteristics of psychopathy, particularly among university students who may not have prevaricated extensively before. On the other hand, while it was clear that our high-PPI group showed reduced lie-related RT gain following training, meaning that their lying act became less detectable based on RT, it is worth noting that their RTs were still significantly longer for lying than for truth-telling after training.

Lying requires overriding default honest responses and switching to a stimulus-response pattern that is consistent with counterfactual responses.[27–29,69] In accordance, we observed elevated dorsolateral frontal and cerebellar activities during processing lying cues prior to training, reinforcing the notion that participants engaged in greater cognitive and emotional

processing upon receiving the cue to lie. However, we obtained some evidence indicating that lie-related increases in parietal and VMPFC activities were observed only in the low-PPI group. Such findings, albeit tentative, concur with existing evidence indicating superior functioning of the parietal cortex,[7] but compromised VMPFC functioning in non-incarcerated community individuals with higher psychopathic traits,[6,11] along with an emotional bias towards social norm noncompliance in students scoring higher in psychopathy.[12] Moreover, the positive association between lying-related parietal activities and RT suggests the parietal regions may mediate individual differences in lying speed. Thus, the reduced parietal activities in the high-PPI group might indicate higher processing efficiency for lying,[45] possibly supported by the stronger parietal-IFG functional connectivity during lying in those participants.

Our findings clearly indicate reductions in activities during the processing of lying cues in the dorsolateral frontal and cerebellar regions following training, consistent with previous evidence that showed decreased activity in those areas along with increased performance following cognitive-motor training.[47,55,56,59,70] Crucially, in both the dorsolateral and ventromedial frontal regions, along with the cerebellum, lying-related activity reductions were exclusively observed among the high-PPI group. Indeed, in the dorsolateral frontal cortex and VMPFC, the low-PPI group exhibited the opposite pattern of lying-related activity increase following training. The reduced VMPFC activities following training in the high-PPI group may be linked with blunted negative affective reactions to lying, complementing existing cross-sectional evidence implicating reduced emotional bias in students scoring higher in psychopathy,[12] and further indicates that such characteristics may be plastic and emerge only following repeated exposure to lying cues. This reduced affective conflict in the prospect to lie could also lead to reduced dorsolateral frontal and cerebellar activities needed for affective regulatory processes.[12,46] Moreover, activities in other social affective networks, including the posterior cingulate cortex and dorsomedial prefrontal cortex, were also disproportionally decreased in the high-PPI group when processing lying cues (Supplementary Table T1).[9,44] Overall, the results are broadly consistent with the notion that through training, individuals with higher psychopathic traits acquired a performance strategy of switching between two stimulus–response patterns based on the cue content. As these patterns became better learned, task performance improved and neural activities in the dorsolateral frontal cortex and cerebellum continued to decrease.[55] This notion was collaterally supported by the increased prefrontal–cerebellum functional connectivity during lying following training in the high-PPI group, which may further promote cognitive flexibility and predict greater improvements in lying speed.[71] In contrast, the increased lying-related dorsolateral frontal, VMPFC and cerebellar activities following training in the low-PPI group may suggest greater cognitive recruitment, as well as increasingly negative affective reactions to lying cues that demand higher regulatory efforts. Such patterns of neural changes may contribute to the failure of those individuals to improve lying speed.

Participants' neural patterns during face responding were markedly similar to those during cue processing. Following training, across the ventrolateral frontal cortices, the ACC and the cerebellum, activities during lying relative to truth-telling decreased unanimously among the high-PPI, but not low-PPI, group. Furthermore, the close association between lying-related activity decrease in the cerebellum, and to an extent in the VLPFC, and lying-related RT decreases, suggest that the different training effects on the lying performances of low and high-PPI groups were partly explained by these distinct neural changes.[33] During lying, the 'true' information and response need to be suppressed and reversed.[33] Thus, lying requires a series of attention, working memory, inhibition and conflict resolution processes, accounting

for the widespread neural activations particularly in the prefrontal cortex and the cerebellum. As described above, we speculate that participants with higher psychopathic traits acquired bimodal stimulus-response patterns for the honest and lying responses following training, and were able to flexibly switch between them according to the cue. We further speculate that once the 'correct' response pattern had been activated, implementing such a pattern following face presentation enabled those participants to respond accurately and quickly with relatively low cognitive effort. Such notion is collaterally supported by the findings that the high-PPI group showed greater lying-related decrease in visual and parahippocampal regions following training (Supplementary Table T2), suggesting reduced face stimuli processing once (relatively) autonomous stimulus–response patterns had been activated.[72] On the contrary, the low-PPI group seemed to rely on marked recruitment of cognitive resources when lying, resulting in slower lying speeds.[51]

Some might argue that the DLT does not capture the more ecologically valid lying processes as participants merely followed explicit instructions for giving counterfactual responses, rather than producing spontaneous lies.[44] Although the DLT may not fully capture the affective components of spontaneous lying, we believe this task captures the essential cognitive components widely considered to be integral to lying, including inhibition, working memory manipulation, attention switching and conflict resolution.[29,73] Our finding that the differential behavioural and neural changes observed in the high-PPI and low-PPI groups following training were specific to the DLT, but not to the CT, adds strength to such notion. Moreover, the distinct neural pattern changes elicited by the DLT and the CT in frontoparietal and cerebellar regions are difficult to be explained by difference in task visual stimuli alone. Our findings are also unlikely to be explained by motivational factors, given the psychopathic tendency in community or student samples is not generally associated with altered cognitive task motivation,[4,6] that it is difficult to see how motivational effect on task performance may differ between response type and session, and how motivational factors could substantially explain observed activity difference in lateral fronto-parietal and fronto-cerebellar regions, which are more regarded as 'cool' cognitive networks.[8] Future studies may further increase the task's ecological validity, such as allowing participants to make their own decisions on whether and when to lie.[44] We also did not include explicit measures for cognitive functioning or affective reactivity, which would aid the interpretation of the behavioural and neural results.[73] Future research may additionally incorporate tasks that assess different cognitive and affective domains. Moreover, our study involved performing the same task throughout, and future research should investigate whether the practice effect would transfer to different tasks or stimuli.[54] Last but not the least, our participants were all university students, who may have different functioning levels, traits and neurobehavioural characteristics from forensic or even community samples. Future research is needed for generalizing our findings to other types of high-psychopathy individuals.

In conclusion, to our knowledge, this is a first study assessing the influence of practice on behavioural and neural processes while lying about face familiarity in university students with varying levels of psychopathic traits. We uncovered marked differences in frontal, parietal and cerebellar neural changes following practice between the low-PPI and high-PPI groups, which predicted their differential improvements in lying speed. These findings support behavioural and neural plasticity of lying processes specifically among individuals reporting greater psychopathic tendency, and are consistent with the conceptualization of psychopathy as a trait linked with preserved, or tentatively superior, functioning of brain networks implicated in cognitive processing, but reduced reactivity of neural circuitries connected with affective processing. The translational clinical implication is

10

that certain phenotypes of psychopathy, such as higher potency to lie, may be the consequence of joint influence of innate mechanisms and life experiences, and early-life behavioural interventions may be effective in altering behavioural manifestation of psychopathy.[74] Future research should further improve the validity of the lie task and incorporate independent cognitive and affective measures, as well as extending our findings to forensic and community populations.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

REFERENCES

1 Hare RD. *The Hare Psychopathy Checklist-Revised (PCL-R)*. 2nd edn., Multi-Health Systems: Toronto, Ontario, Canada, 2003.
2 Lilienfeld SO, Widows MR. *Psychopathic Personality Inventory-Revised: Professional Manual*. Psychological Assessment Resources: Lutz, FL, 2005.
3 Hare RD. *The Hare Psychopathy Checklist—Revised*. Multi-Health Systems: Toronto, Ontario, Canada, 1991.
4 Benning SD, Patrick CJ, Hicks BM, Blonigen DM, Krueger RF. Factor structure of the psychopathic personality inventory: validity and implications for clinical assessment. *Psychol Assess* 2003; **15**: 340–350.
5 Ross SR, Benning SD, Patrick CJ, Thompson A, Thurston A. Factors of the Psychopathic Personality Inventory: criterion-related validity and relationship to the BIS/BAS and five-factor models of personality. *Assessment* 2009; **16**: 71–87.
6 Gao Y, Raine A. Successful and unsuccessful psychopaths: a neurobiological model. *Behav Sci Law* 2010; **28**: 194–210.
7 Ishikawa SS, Raine A, Lencz T, Bihrle S, Lacasse L. Autonomic stress reactivity and executive functions in successful and unsuccessful criminal psychopaths from the community. *J Abnorm Psychol* 2001; **110**: 423–432.
8 Maes JHR, Brazil IA. No clear evidence for a positive association between the interpersonal-affective aspects of psychopathy and executive functioning. *Psychiatry Res* 2013; **210**: 1265–1274.
9 Kiehl KA. A cognitive neuroscience perspective on psychopathy: evidence for paralimbic system dysfunction. *Psychiatry Res* 2006; **142**: 107–128.
10 Blair RJR. The amygdala and ventromedial prefrontal cortex in morality and psychopathy. *Trends Cogn Sci* 2007; **11**: 387–392.
11 Lotze M, Veit R, Anders S, Birbaumer N. Evidence for a different role of the ventral and dorsal medial prefrontal cortex for social reactive aggression: an interactive fMRI study. *Neuroimage* 2007; **34**: 470–478.
12 Rilling JK, Glenn AL, Jairam MR, Pagnoni G, Goldsmith DR, Elfenbein HA *et al*. Neural correlates of social cooperation and non-cooperation as a function of psychopathy. *Biol Psychiatry* 2007; **61**: 1260–1271.
13 Blair RJR. Neuro-cognitive models of acquired sociopathy and developmental psychopathy. In: Glicksohn J (ed). *The Neurobiology of Criminal Behavior*. Kluwer Academic Publishers: Dordrecht, The Netherlands, 2002, pp 157–186.
14 Kosson DS, Suchy Y, Mayer AR, Libby J. Facial affect recognition in criminal psychopaths. *Emotion* 2002; **24**: 398–411.
15 Iria C, Barbosa F. Perception of facial expressions of fear: comparative research with criminal and non-criminal psychopaths. *J Forens Psychiatry Psychol* 2009; **20**: 66–73.
16 Osumi T, Shimazaki H, Imai A, Sugiura Y, Ohira H. Psychopathic traits and cardiovascular responses to emotional stimuli. *Personal Individ Differ* 2007; **42**: 1391–1402.
17 Hornak J, Bramham J, Rolls ET, Morris RG, O'Doherty J, Bullock PR *et al*. Changes in emotion after circumscribed surgical lesions of the orbitofrontal and cingulate cortices. *Brain* 2003; **126**: 1691–1712.
18 Yang Y, Raine A, Lencz T, Bihrle S, LaCasse L, Colletti P. Volume reduction in prefrontal gray matter in unsuccessful criminal psychopaths. *Biol Psychiatry* 2005; **57**: 1103–1108.
19 Yang Y, Raine A. Functional neuroanatomy of psychopathy. *Psychiatry* 2008; **7**: 133–136.
20 Yang Y, Raine A, Colletti P, Toga AW, Narr KL. Morphological alterations in the prefrontal cortex and the amygdala in unsuccessful psychopaths. *J Abnorm Psychol* 2010; **119**: 546–554.
21 Yang Y, Raine A, Lencz T, Bihrle S, Lacasse L, Colletti P. Prefrontal white matter in pathological liars. *Br J Psychiatry* 2005b; **187**: 320–325.
22 Yang Y, Raine A, Narr KL, Lencz T, LaCasse L, Colleti P *et al*. Localisation of increased prefrontal white matter in pathological liars. *Br J Psychiatry* 2007; **190**: 174–175.
23 Carlson SR, Thái S. ERPs on a continuous performance task and self-reported psychopathic traits: P3 and CNV augmentation are associated with fearless dominance. *Biol Psychol* 2010; **85**: 318–330.
24 Sadeh N, Verona E. Psychopathic personality traits associated with abnormal selective attention and impaired cognitive control. *Neuropsychology* 2008; **22**: 669–680.
25 Mahmut M, Homewood J, Stevenson RJ. The characteristics of non-criminals with high psychopathy traits: are they similar to criminal psychopaths? *J Res Personal* 2008; **42**: 679–692.
26 Sellbom M, Verona E. Neuropsychological correlates of psychopathic traits in a nonincarcerated sample. *J Res Personal* 2007; **41**: 276–294.
27 Langleben DD, Schroeder L, Maldjian JA, Gur RC, McDonald S, Ragland JD *et al*. Brain activity during simulated deception: an event-related functional magnetic resonance study. *Neuroimage* 2002; **15**: 727–732.
28 Lee TMC, Liu HL, Tan LH, Chan CCH, Mahankali S, Feng CM *et al*. Lie detection by functional magnetic resonance imaging. *Hum Brain Mapp* 2002; **15**: 157–164.
29 Spence SA, Hunter MD, Farrow TFD, Green RD, Leung DH, Hughes CJ *et al*. A cognitive neurobiological account of deception: evidence from functional neuroimaging. *Philos Trans R Soc Lond B Biol Sci* 2004; **359**: 1755–1762.
30 Abe N, Suzuki M, Mori E, Itoh M, Fujii T. Deceiving others: distinct neural responses of the prefrontal cortex and amygdala in simple fabrication and deception with social interactions. *J Cogn Neurosci* 2007; **19**: 287–295.
31 Fullam RS, McKie S, Dolan MC. Psychopathic traits and deception: functional magnetic resonance imaging study. *Br J Psychiatry* 2009; **194**: 229–235.
32 Ito A, Abe N, Fujii T, Ueno A, Koseki Y, Hashimoto R *et al*. The role of the dorsolateral prefrontal cortex in deception when remembering neutral and emotional events. *Neurosci Res* 2011; **69**: 121–128.
33 Sun D, Lee TMC, Chan CCH. Unfolding the spatial and temporal neural processing of lying about face familiarity. *Cereb Cortex* 2015; **25**: 927–936.
34 Lee TMC, Lee TMY, Raine A, Chan CCH. Lying about the valence of affective pictures: An fMRI study. *PLoS ONE* 2010; **5**: e12291.
35 Vendemia J, Buzan RF, Simon-Dack SL. Reaction time of motor responses in two-stimulus paradigms involving deception and congruity with varying levels of difficulty. *Behav Neurol* 2005a; **16**: 25–36.
36 Blasi G, Goldberg TE, Weickert T, Das S, Kohn P, Zoltick B *et al*. Brain regions underlying response inhibition and interference monitoring and suppression. *Eur J Neurosci* 2006; **23**: 1658–1664.
37 Christ SE, Van Essen DC, Watson JM, Brubaker LE, McDermott KB. The contributions of prefrontal cortex and executive control to deception: evidence from activation likelihood estimate meta-analyses. *Cereb Cortex* 2009; **19**: 1557–1566.
38 Lie CH, Specht K, Marshall JC, Fink GR. Using fMRI to decompose the neural processes underlying the Wisconsin Card Sorting Test. *Neuroimage* 2006; **30**: 1038–1049.
39 Apps MJ, Rushworth MS, Chang SC. The anterior cingulate gyrus and social cognition: tracking the motivation of others. *Neuron* 2016; **90**: 692–707.
40 Kolling N, Behrens T, Wittmann MK, Rushworth M. Multiple signals in anterior cingulate cortex. *Curr Opin Neurobiol* 2016; **37**: 36–43.
41 Ganis G, Kosslyn SM, Stose S, Thompson WL, Yurgelun-Todd DA. Neural correlates of different types of deception: an fMRI investigation. *Cereb Cortex* 2003; **13**: 830–836.
42 Nunez JM, Casey BJ, Egner T, Hare T, Hirsch J. Intentional false responding shares neural substrates with response conflict and cognitive control. *Neuroimage* 2005; **25**: 267–277.
43 Phan KL, Magalhaes A, Ziemlewicz TJ, Fitzgerald DA, Green C, Smith W. Neural correlates of telling lies: a functional magnetic resonance imaging study at 4 Tesla. *Acad Radiol* 2005; **12**: 164–172.
44 Sip KE, Roepstorff A, McGregor W, Frith CD. Detecting deception: the scope and limits. *Trends Cogn Sci* 2008; **12**: 48–53.
45 Jiang W, Liu H, Liao J, Ma X, Rong P, Tang Y *et al*. A functional MRI study of deception among offenders with antisocial personality disorders. *Neuroscience* 2013; **244**: 90–98.
46 Schutter DJ, van Honk J. The cerebellum in emotion regulation: a repetitive transcranial magnetic stimulation study. *Cerebellum* 2009; **8**: 28–34.
47 Balsters JH, Ramnani N. Cerebellar plasticity and the automation of first-order rules. *J Neurosci* 2011; **31**: 2305–2312.

48 Balsters JH, Whelan CD, Robertson IH, Ramnani N. Cerebellum and cognition: evidence for the encoding of higher order rules. *Cereb Cortex* 2013; **23**: 1433–1443.

49 Ohira H, Nomura M, Ichikawa N, Isowa T, Lidaka T, Sato A *et al.* Association of neural and physiological responses during voluntary emotion suppression. *Neuroimage* 2006; **29**: 721–733.

50 Hayashi A, Abe N, Ueno A, Shigemune Y, Mori E, Tashiro M *et al.* Neural correlates of forgiveness for moral transgressions involving deception. *Brain Res* 2010; **1332**: 90–99.

51 Johnson R, Barnhardt J, Zhu J. Differential effects of practice on the executive processes used for truthful and deceptive responses: an event-related brain potential study. *Cogn Brain Res* 2005; **24**: 386–404.

52 Vendemia JM, Buzan RF, Green EP. Practice effects, workload, and reaction time in deception. *Am J Psychol* 2005b; **5**: 413–429.

53 Hu X, Chen H, Fu G. A repeated lie becomes a truth? The effect of intentional control and training on deception. *Front Psychol* 2012; **3**: 488.

54 Van Bockstaele B, Verschuere B, Moens T, Suchotzki K, Debey E, Spruyt A. Learning to lie: effects of practice on the cognitive cost of lying. *Front Psychol* 2012; **3**: 526.

55 Toni I, Ramnani N, Josephs O, Ashburner J, Passingham RE. Learning arbitrary visuomotor associations: temporal dynamic of brain activity. *Neuroimage* 2001; **14**: 1048–1057.

56 Bischoff-Grethe A, Goedert KM, Willingham DT, Grafton ST. Neural substrates of response-based sequence learning using fMRI. *J Cogn Neurosci* 2004; **16**: 127–138.

57 MacDonald AW, Cohen JD, Stenger VA, Carter CS. Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science* 2000; **288**: 1835–1838.

58 Sohn MH, Ursu S, Anderson JR, Stenger VA, Carter CS. The role of prefrontal cortex and posterior parietal cortex in task switching. *Proc Natl Acad Sci USA* 2000; **97**: 13448–13453.

59 Garavan H, Kelley D, Rosen A, Rao SM, Stein EA. Practice-related functional activation changes in a working memory task. *Microsc Res Tech* 2000; **51**: 54–63.

60 Hughes MA, Stout JC, Dolan MC. Concurrent Validity of the Psychopathic Personality Inventory–Revised and the Psychopathy Checklist: Screening Version in an Australian offender sample. *Crim Justice Behav* 2013; **40**: 802–813.

61 Neumann CS, Uzieblo K, Crombez G, Hare RD. Understanding the Psychopathic Personality Inventory (PPI) in terms of the unidimensionality, orthogonality, and construct validity of PPI-I and-II. *Personal Disord* 2013; **4**: 77–79.

62 Lee TMC, Leung MK, Lee TMY, Raine A, Chan CCH. I want to lie about not knowing you, but my precuneus refuses to cooperate. *Sci Rep* 2013; **3**: 1636.

63 Rasbash J, Charlton C, Browne WJ, Healy M, Cameron B. *MLwiN Version 21*. Centre for Multilevel Modelling, University of Bristol, 2009.

64 Song X, Dong Z, Long X, Li S, Zuo X, Zhu C *et al.* REST: a toolkit for resting-state functional magnetic resonance imaging data processing. *PLoS ONE* 2011; **6**: e25031.

65 Holm S. A simple sequential rejective method procedure. *Scand J Stat* 1979; **6**: 65–70.

66 McLaren DG, Ries ML, Xu G, Johnson SC. A generalized form of context-dependent psychophysiological interactions (gPPI): a comparison to standard approaches. *Neuroimage* 2012; **61**: 1277–1286.

67 Natu V, O'Toole AJ. The neural processing of familiar and unfamiliar faces: a review and synopsis. *Br J Psychol* 2011; **102**: 726–747.

68 Hamilton RKB, Racer KH, Newman JP. Impaired integration in psychopathy: a unified theory of psychopathic dysfunction. *Psychol Rev* 2015; **122**: 770–791.

69 Lee TMC, Liu HL, Chan CCH, Ng YB, Fox PT, Gao JH. Neural correlates of feigned memory impairment. *NeuroImage* 2005; **28**: 305–313.

70 Klingberg T. Training and plasticity of working memory. *Trends Cogn Sci* 2010; **14**: 317–324.

71 Luna B, Padmanabhan A, O'Hearn K. What has fMRI told us about the development of cognitive control through adolescence? *Brain Cognit* 2010; **72**: 101–113.

72 Haxby JV, Hoffman EA, Gobbini MI. The distributed human neural system for face perception. *Trends Cogn Sci* 2000; **4**: 223–233.

73 Vendemia J. fMRI as a method of detection of deception: a review of experiences. *Eur Polygraph* 2014; **8**: 5–21.

74 Hawes DJ, Price MJ, Dadds MR. Callous-unemotional traits and the treatment of conduct problems in childhood and adolescence: a comprehensive review. *Clin Child Fam Psychol Rev* 2014; **17**: 248–267.

Supplementary Information accompanies the paper on the Translational Psychiatry website (http://www.nature.com/tp)