**ORIGINAL ARTICLE**

# Deficient auditory emotion processing but intact emotional multisensory integration in alexithymia

Zhihao Wang[1,2] | Mai Chen[3] | Katharina S. Goerlich[2] | André Aleman[1,2] | Pengfei Xu[4,7,8] | Yuejia Luo[1,4,5,6,7]

[1]Shenzhen Key Laboratory of Affective and Social Neuroscience, Magnetic Resonance Imaging, Center for Brain Disorders and Cognitive Sciences, Shenzhen University, Shenzhen, China

[2]Department of Biomedical Sciences of Cells & Systems, Section Cognitive Neuroscience, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

[3]School of Psychology, Shenzhen University, Shenzhen, China

[4]State Key Laboratory of Cognitive and Learning, Faculty of Psychology, Beijing Normal University, Beijing, China

[5]Department of Psychology, Southern Medical University, Guangzhou, China

[6]The Research Center of Brain Science and Visual Cognition , Medical School, Kunming University of Science and Technology, Kunming, China

[7]Center for Neuroimaging, Shenzhen Institute of Neuroscience, Shenzhen, China

[8]Guangdong-Hong Kong-Macao Greater Bay Area Research Institute for Neuroscience and Neurotechnologies, Kwun Tong, Hong Kong, China

**Correspondence**
Pengfei Xu, Faculty of Psychology, Beijing Normal University, No. 19, Xinjiekouwai St, Haidian District, Beijing 100875, China.
Email: pengfeixu.px@gmail.com

Yuejia Luo, Shenzhen Key Laboratory of Affective and Social Neuroscience, Shenzhen University, 3688 Nanhai Ave., Nanshan District, Shenzhen 518060, China.
Email: luoyj@szu.edu.cn

**Abstract**

Alexithymia has been associated with emotion recognition deficits in both auditory and visual domains. Although emotions are inherently multimodal in daily life, little is known regarding abnormalities of emotional multisensory integration (eMSI) in relation to alexithymia. Here, we employed an emotional Stroop-like audiovisual task while recording event-related potentials (ERPs) in individuals with high alexithymia levels (HA) and low alexithymia levels (LA). During the task, participants had to indicate whether a voice was spoken in a sad or angry prosody while ignoring the simultaneously presented static face which could be either emotionally congruent or incongruent to the human voice. We found that HA performed worse and showed higher P2 amplitudes than LA independent of emotion congruency. Furthermore, difficulties in identifying and describing feelings were positively correlated with the P2 component, and P2 correlated negatively with behavioral performance. Bayesian statistics showed no group differences in eMSI and classical integration-related ERP components (N1 and N2). Although individuals with alexithymia indeed showed deficits in auditory emotion recognition as indexed by decreased performance and higher P2 amplitudes, the present findings suggest an intact capacity to integrate emotional information from multiple channels in alexithymia. Our work provides

**PSYCHOPHYSIOLOGY** SPR

valuable insights into the relationship between alexithymia and neuropsychological mechanisms of emotional multisensory integration.

## 1 | INTRODUCTION

Alexithymia, a subclinical personality trait, is characterized as an impaired ability to identify, describe, and regulate one's emotions (Luminet et al., 2018). Ample evidence has demonstrated that alexithymia is not only associated with difficulties in identifying visual emotional stimuli (Van der Velde et al., 2013), especially sad and angry facial expressions (Kano et al., 2003; Lee et al., 2011), but high alexithymia levels are also related to a blunted sensitivity to the emotional qualities of speech (Goerlich-Dobre et al., 2014). Indeed, individuals with alexithymia may need more attentional resources to process emotion-laden contents (Franz et al., 2004). Emotions, however, are inherently multimodal in daily life (Tang et al., 2016). It has been shown that emotional multisensory integration (eMSI), the ability to bundle emotional information synergistically from different modalities, plays a key role in emotion recognition (Klasen et al., 2012). However, it is currently unclear whether individuals with alexithymia have deficits in eMSI. This is of particular importance given its ecological relevance, as humans are constantly exposed to competing, complex audiovisual emotional information in social interaction contexts. Indeed, it has been shown that a reliable interpretation of others' affective states requires the integration of multimodal stimuli into a single, coherent percept (see De Gelder & Bertelson, 2003, for a review).

There have been two eMSI-related studies in alexithymia. Delle-Vigne et al. (2014) observed a larger N1 component in alexithymia during an emotional oddball audiovisual task, suggesting that individual with alexithymia need more attentional resources to process emotionally congruent audiovisual cues. However, the impact of emotion recognition from visual clues or from the voice on eMSI in alexithymia was not investigated. Another study showed a negative correlation of alexithymia levels with the N400 for emotionally incongruent compared with congruent audiovisual stimuli during an affective priming task, indicating reduced sensitivity to affective mismatches in audiovisual information with higher levels of alexithymia (Goerlich et al., 2011). This study did not test clinically relevant alexithymia levels (only two individuals

had alexithymia scores exceeding the clinical threshold). Therefore, it remains uncertain whether the observed effects can be attributed to abnormal eMSI in clinically-relevant alexithymia. Previous studies suggested that eMSI is impaired in patients with affective disorders (Müller et al., 2013), even in subclinical populations (Campanella et al., 2010). Given that alexithymia has long been considered as a risk factor for affective disorders (highly correlated with anxiety and depression; Hendryx et al., 1991; Li et al., 2015), it is reasonable to speculate that eMSI may be reduced in alexithymia (hypointegration hypothesis). However, the opposite may also be hypothesized: stronger integration as a compensation for reduced emotion perception ability (hyper-integration hypothesis). That is, given that alexithymia seems to be associated with a reduced intensity perception of emotional visual and auditory stimuli (for a review, see Donges & Suslow, 2017), individuals with high alexithymia levels may show better eMSI than a control group with average alexithymia levels. If so, eMSI would function as a compensation for deficits in visual or auditory emotion processing.

Event-related potential (ERP) techniques, with high temporal resolution, have been used to shed light on the time course of eMSI (Klasen et al., 2012; Wang et al., 2012). It has been shown that there are two stages of eMSI processing, including multisensory facilitation and multisensory inhibition (Campanella & Belin, 2007; Wang et al., 2012). Multisensory facilitation, a sensory process during which the emotional attention triggered by one domain is amplified by another modality to increase efficiency (Klasen et al., 2012), occurs when bimodal stimuli have a close spatial and temporal coincidence (Stein & Stanford, 2008) or when the bimodal stimuli are congruent, leading to priming effects (Chen et al., 2016; Föcker & Röder, 2019). Multisensory facilitation can be reflected by a larger auditory N1 component, a negative deflection over frontal-central regions (Luck, 2014; Luck et al., 2000). By contrast, multisensory inhibition occurs when simultaneously presenting bimodal stimuli that are incongruent with each other, for example in their affective or semantic connotation (Wang et al., 2012). This incongruency presents a conflict, causing a need to allocate more attentional

resources to disambiguate and select a correct response, resulting in prolonged reaction times (Donohue et al., 2013). At the neural level, increases in the frontal-central N2 have been repeatedly reported for emotionally incongruent audiovisual stimuli, signaling conflict monitoring and disambiguation processes in eMSI (Zinchenko et al., 2017). Taken together, abnormally high or low amplitude of the classic eMSI-related components N1 and N2 can be regarded as indices of abnormal neuropsychological processes based on the hyper- or hypo-integration hypotheses.

The aim of the current study was to examine eMSI ability in relation to alexithymia by applying ERP techniques in a Stroop-like visual-auditory task. During the task, participants were asked to identify whether someone's voice speaking in a sad or angry prosody while ignoring the simultaneously presented static face which could be either emotionally congruent or incongruent to the human voice. According to the hypo- or hyper-integration hypothesis, we hypothesized that individuals with high alexithymia levels (HA) would show abnormal integrative performance and differences in early ERP components during emotional multisensory integration compared with individuals with low alexithymia levels (LA). Specifically, regarding the hypothesis of hypo-integration, one would expect worse integrative performance in HA versus LA. At the neural level, this would be reflected in smaller differences in N1 and N2 between congruent and incongruent conditions in HA as compared with LA. With regard to the hypothesis of hyper-integration, better integrative performance would be predicted in HA versus LA. At the neural level, this would be reflected in larger differences in N1 and N2 between congruent and incongruent conditions in HA as compared with LA. In addition to eMSI-relevant hypotheses, we hypothesized that HA would show deficits in auditory emotion recognition as indexed by the P2 (replicating previous research; Franz et al., 2004).

## 2 | METHOD

### 2.1 | Participants

Forty-eight healthy adults from a pool of 543 (368 females; age: 17–38 years, mean $\pm$ *SD*: 20.03 $\pm$ 2.11 years) students at Shenzhen University participated in the experiment. Each participant in the pool completed the Chinese version of the 20-item Toronto Alexithymia Scale (TAS-20; Bagby et al., 1994a; Zhu et al., 2007). In light of the international cutoff to assess clinical-relevant alexithymia in TAS-20 (Taylor et al., 1988), individuals with TAS-20 scores higher or equal to 61 (14.9% of the pool) were identified as high alexithymia (HA) whereas those with TAS-20 scores lower or equal to 51 (54.6% of the pool) were classified as low alexithymia (LA; Table 1). The sample size of 48 participants (Table 1) was determined based on the medium effect size by G*power (version: 3.1; Faul et al., 2007). Twenty-three participants per group were needed to detect a reliable effect (Cohen $f = 0.25$, $\alpha = .05$, $1 - \beta = .9$, ANOVA Repeated-measures, within-between interaction; Faul et al., 2007). All participants had normal or corrected-to-normal vision and normal hearing and reported no psychiatric illness in present or past. This study was approved by the Ethics Committee of Shenzhen University and informed written consents were obtained from all participants.

### 2.2 | Stimulus materials

The vocal stimuli were 64 pseudo-sentences uttered in sad and angry prosody selected from the Database of Chinese Vocal Emotions (DoCVE; Liu & Pell, 2012). These angry (32 items) and sad (32 items) pseudo-utterances were recorded by four native speakers (two females; eight items per speaker).

**TABLE 1** Demographics and questionnaire scores

| | HA (25; 12 females) | | LA (23; 12 females) | | | | Cronbach's alpha |
|---|---|---|---|---|---|---|---|
| | Mean (*SD*) | [Min, max] | Mean (*SD*) | [Min, max] | *t* | *p* | |
| Age | 19.96 (1.46) | [18, 24] | 19.78 (1.81) | [17, 24] | 0.376 | .709 | |
| TAS-20 | 66.08 (2.96) | [62, 73] | 36.09 (5.01) | [22, 43] | 24.992 | <.001 | 0.856 |
| DIF | 23.56 (2.38) | [20, 27] | 11.09 (2.68) | [7, 16] | 17.077 | <.001 | 0.916 |
| DDF | 18.52 (1.98) | [14, 22] | 8.74 (1.91) | [5, 13] | 17.373 | <.001 | 0.534 |
| EOT | 24.00 (3.01) | [19, 32] | 16.26 (3.31) | [10, 24] | 8.485 | <.001 | 0.369 |
| BDI | 12.92 (6.36) | [3, 27] | 4.00 (4.68) | [0, 17] | 5.492 | <.001 | 0.870 |
| BAI | 33.92 (6.36) | [24, 48] | 25.00 (4.68) | [21, 38] | 5.492 | <.001 | 0.857 |
| AQ | 124.44 (10.21) | [100, 142] | 109.39 (8.42) | [89, 120] | 5.543 | <.001 | 0.673 |

Abbreviations: AQ, Autism Spectrum Quotient; BAI, beck anxiety inventory; BDI, beck depression inventory; DDF, difficulty describing feelings; DIF, difficulty identifying feelings; EOT, externally oriented thinking; HA, individuals with alexithymia; LA, individuals without alexithymia; TAS-20, twenty-item Toronto Alexithymia Scale.

**PSYCHOPHYSIOLOGY** SPR

In light of previous studies (Liu et al., 2015), each item was cut down to 800 ms using the Praat software (Boersma & Weenink, 2021) to avoid ceiling effects and to control the amount of expressed information across stimuli. The original reflected the whole utterance. Although the whole utterances selected from this dataset had matched (Table S1), it is uncertain how cutting to 800 ms would affect them. Therefore, we conducted a behavioral follow-up. To check whether emotional properties of the 800 ms voice datasets were matched as those of the original categorizations, a follow-up rating was conducted on the auditory stimuli in each group (see Supporting Information for details). Although angry and sad utterances used in the current study were not strictly matched in emotional properties of intensity and valence (Table 2), the absence of significant group differences in these parameters suggests that alexithymia-related effects were not confounded by basic emotional characteristics (for details, see Supporting Information). The amplitude values (maximum amplitude, minimum amplitude) were normalized in relation to the average minimum amplitude value of all neutral utterances by that speaker, to correct for individual differences in a speaker's intensity by subtracting the normalized minimum amplitude values from the normalized maximum amplitude values (Liu & Pell, 2012). Thus, all vocal stimuli had approximate equal intensity. Note that pseudo-sentences are composed of real function words, but the sentence content is meaningless to avoid influences of the linguistic-semantic content during speech. The vocal stimuli were expressed in stereo mode. The visual stimuli of 64 grey-scaled emotional faces were selected from the Chinese Facial Affective Picture System (CFAPS; Gong et al., 2011). These emotional faces consisted of 32 angry (16 females) and 32 sad (16 females) faces. We also balanced the emotional characteristics between angry and sad faces (Table S1). Specifically, no significant differences were found between angry and sad faces in valence, arousal, and attractiveness (Table S1), suggesting that angry and sad facial stimuli matched in these emotional dimensions. The visual stimuli of faces subtended a visual angle of 3° × 4°.

## 2.3 | Self-report questionnaires

The TAS-20 measures three facets of alexithymia: (a) difficulty identifying feelings (DIF; 7 items); (b) difficulty describing feelings (DDF; 5 items), and (c) externally oriented thinking (EOT; 8 items). In a self-reporting manner, each item is rated on a five-point Likert scale ranging from 1 ("strongly disagree") to 5 ("strongly agree"), with five items being negatively scored. For analysis, the negatively keyed items are reverted and item scores for each respective dimension are summed up. The total score is calculated as the sum of all items. High scores represent high levels of alexithymia (Bagby et al., 1994; Bagby, Taylor, et al., 1994; Taylor et al., 2003). Importantly, the Chinese version of the TAS-20 has been established with acceptable reliability and validity (Zhu et al., 2007). To control for the potential confounding effect of depression, anxiety, and autism (Bird & Cook, 2013; Li et al., 2015; Van der Velde et al., 2013), participants also completed the Beck Anxiety Inventory (BAI; Beck et al., 1988), the Beck Depression Inventory (BDI; Beck, 1967), and the Autism Spectrum Quotient (AQ; Baron-Cohen et al., 2001).

## 2.4 | Task and procedure

We adopted an emotional Stroop-like visual-auditory task (Figure 1). In each trial, a fixation cross was first shown on the screen center with a random duration (800~1,500 ms). Subsequently, the face-voice pairs were presented simultaneously. The voice lasted for 800 ms maximally while the face was displayed for 2,500 ms maximally. Once they had identified the emotion (angry or sad) from the voice, participants pressed button "f" if the voice was angry and "j" if the voice was sad. The assignment of the two response buttons was counterbalanced across participants. While viewing the static face all the time (2,500 ms maximally), presentation of the voice stimuli stopped once a button was pressed (800 ms maximally). Finally, a blank screen lasted for 1,000 ms, and then the next trial began.

**TABLE 2** Follow-up ratings for voice stimuli with angry and sad emotions in each group

| Group | Categorization | Recognition rate | Intensity | Valence |
|---|---|---|---|---|
| HA | Angry | 0.67 ± 0.19 | 4.04 ± 0.60 | 2.01 ± 0.46 |
|  | Sad | 0.68 ± 0.18 | 3.65 ± 0.65 | 2.18 ± 0.60 |
| MA | Angry | 0.66 ± 0.18 | 4.10 ± 0.37 | 2.08 ± 0.48 |
|  | Sad | 0.64 ± 0.22 | 3.71 ± 0.49 | 2.19 ± 0.38 |
| LA | Angry | 0.72 ± 0.22 | 4.21 ± 0.31 | 1.86 ± 0.38 |
|  | Sad | 0.75 ± 0.19 | 3.74 ± 0.44 | 2.05 ± 0.43 |

*Note::* Descriptive data are presented as mean (standard deviation).

Abbreviations: HA, individuals with high level of alexithymia; LA, individuals with low level of alexithymia; MA, individuals with middle levels of alexithymia.
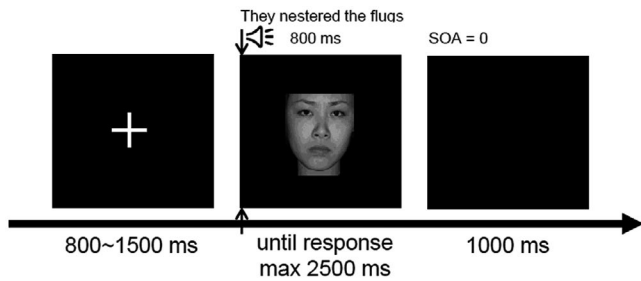
**FIGURE 1** Trial design. In each trial, the face-voice pair expressing congruent or incongruent emotions (angry–angry, angry–sad, sad–angry, and sad–sad pairs) was presented simultaneously. Faces were presented with a duration of 2,500 ms maximally, accompanied simultaneously by the voice stimuli (pseudowords) lasting for 800 ms maximally. Participants pressed a button as soon as they had identified the emotion from the voice (self-paced, maximally trial duration 2,500 ms). Finally, a blank screen lasted for 1,000 ms and then the next trial began. The upward and downward arrows mark the beginning of the auditory and visual stimuli. SOA, stimulus onset asynchrony

There were four same-gender face-voice pairs (sad–sad, sad–angry, angry–sad, angry–angry). After dividing four pairs into two conditions: congruent (Con; sad–sad, angry–angry) and incongruent (Incon; sad–angry, angry–sad), each condition included 128 trials. Trial presentation was randomized. Participants were instructed to identify the vocal expression (sad or angry) as accurate and quick as possible while ignoring the facial expression. All experimental procedures were presented using the E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA). In addition to the approximate equal intensity of all vocal stimuli (mentioned above), the vocal stimuli were presented binaurally by an insert headphone with equal sound intensity across participants (half of sound intensity of our computer). All participants reported to hear the vocal stimuli clearly. No participant reported having felt uncomfortable at any part of the experiment. The current study adopted the voice task instead of the facial task, because it has been demonstrated that the eMSI effect on Chinese participants is more sensitive to the voice task (Wang et al., 2012). Specifically, during a timed target detection task (visual target and auditory target), the significant integration effect was found from the auditory target whereas no significant integration effect from the visual target was observed (Wang et al., 2012).

Participants completed several practice rounds until achieving at least 80% of accuracy before the formal experiment started. There were four blocks with 64 trials per block. Participants could take self-paced breaks between blocks. After finishing the task, participants rated the intensity of all stimuli (face and voice) on a nine-point Likert scale.

## 2.5 | Electrophysiology (EEG) recording and preprocessing

We recorded EEG data from a 64-electrode scalp cap according to the international 10–20 system (Brain Products, Munich, Germany), with the reference to the channel FCz. The electrooculogram (EOG; vertical) was recorded with electrodes placed below the right eye. Electrodes impedances of EEG and EOG were maintained <5 kΩ. All electrodes were amplified using a 0.01 online high-pass filter and continuously sampled at 1,000 Hz per channel for offline analysis.

EEG data were preprocessed with EEGLAB 14.1.2b (http://sccn.ucsd.edu/eeglab/) in Matlab 2014b (MathWorks Inc). It comprised of the following steps: (a) resampling to 250Hz; (b) low-pass filtering of 30Hz by FIR filter; (c) segmentation from 200ms before to 800ms after stimuli onset; (d) baseline correction from −200 to 0 ms; (e) manually rejecting salient muscle epochs and bad channels if any; (f) Independent Component Analysis (ICA); (g) visually inspecting and rejecting artifact components (horizontal and vertical eye movements and muscle components); (h) interpolating bad channels if any; (i) re-referencing offline to the average of all electrodes; (j) rejecting trials in which EEG voltages were out of range (−80, 80) μV.

## 2.6 | Behavioral statistics

We used SPSS 17.0 to perform statistical analyses, with the significance level at $p = .05$. Trials with response times (RTs) out of range (mean-$3 \times SD$, mean $+ 3 \times SD$) were removed for each participant. The impact of alexithymia on congruency was tested in 2 (Con vs. Incon) by 2 (HA vs. LA) repeated-measures ANOVA with the corrected performance (ACC/RT; speed-accuracy tradeoff) as the dependent variable. To change corrected performance between HA and LA into the common scales (because growing evidence suggested deficits of auditory emotional processing in alexithymia; Goerlich et al., 2011; Goerlich-Dobre et al., 2014), we normalized the corrected performance for each participant as the index of integration effect ([Con − Incon]/[Con + Incon]). This index signals the individual point of the best balance between ACC and RT and thus indicates each participant's "sweet spot" at which sensory integration is most efficient. Importantly, this procedure eliminates the impact of the variable of noninterest (main effect of alexithymia in auditory emotion processing) on the variable of interest (integration efficiency in the HA and LA groups). Repeated-measures ANOVA of normalized performance was conducted with alexithymia groups as the independent variable. To assess the difference between HA and LA in intensity rating of each stimulus, repeated-measures ANOVA was conducted.

**PSYCHOPHYSIOLOGY** SPR

## 2.7 | ERP statistics

Trials of ERP data were included for statistical analysis only if the behavioral response was correct. The current study focused on three ERP components, the frontal-central N1, the central P2, and the frontal-central N2. We also focused on the P2 to test neural correlates of alexithymia because alexithymia has been demonstrated to impact emotion recognition from both auditory and visual emotional stimuli, which can be indexed by P2 (Franz et al., 2004). Visual detection on the grand-averaged waveform in the specified channel and the topography across all conditions confirmed the peak latency of each component. The time window of each component was in line with previous studies (Crowley & Colrain, 2004; Folstein & Van Petten, 2008; Mercado et al., 2006; Miller et al., 2011; Mittag et al., 2013; Wang et al., 2012). Specifically, the N1 was identified in a window of 100–150 ms over the frontal-central electrode (FCz; Wang et al., 2012). The P2 was defined as the mean activity between 165 and 215 ms over the central electrode (Cz; Crowley & Colrain, 2004; Mercado et al., 2006; Miller et al., 2011). The N2 was measured in a time-window of 250–350 ms over the frontal-central electrode (FCz; Folstein & Van Petten, 2008; Mittag et al., 2013). The 2 (Con vs. Incon) by 2 (HA vs. LA) repeated-measures ANOVA was conducted for each component of interest. Please note that the minimum numbers of trials in each condition is higher than 55. The number of bad channels is $0.48 \pm 1.13$ (mean $\pm$ $SD$), with ranges from 0 to 5.

## 2.8 | Correlations among subscales of TAS-20, P2 and corrected performance

To examine the correlations of sub-dimensions of alexithymia with behavioral and electrophysiological differences between HA and LA (if difference is significant), we performed correlations among subscales of the TAS-20 (DIF, DDF, and EOT), the significant ERPs, and the significant behavioral performance.

## 3 | RESULTS

Note that anxiety, depression, and autism have been demonstrated to co-occur with alexithymia (Hendryx et al., 1991; Li et al., 2015). In our sample, we indeed found significant correlations of TAS-20 with BAI ($r = .635$, $p < .001$), BDI ($r = .635$, $p < .001$), and AQ ($r = .644$, $p < .001$). Thus, we performed additional ANCOVAs with BAI, BDI, and AQ scores as covariates to control for potential influences of these covariates on the results. Upon discovering null findings in our analyses, Bayes Factors (BFs) were computed for the respective comparisons to test whether the negative result supported the alternative or the null hypothesis. These were done using JASP (JASP Team, 2020) on both behavioral and ERP data. Descriptive data (mean $\pm$ $SD$) are given in Table 3.

### 3.1 | Behavioral results

For corrected performance, the 2 (Con vs. Incon) by 2 (HA vs. LA) repeated-measures ANOVA revealed main effects of Congruency ($F_{(1,46)} = 90.849$, $p < .001$, $\eta_p^2 = 0.664$; Con > Incon; 95% Confidence interval [CI] of difference between Con and Incon [0.075, 0.115]) and Group ($F_{(1,46)} = 12.197$, $p = .001$, $\eta_p^2 = 0.210$; HA < LA; 95% CI of difference between HA and LA [−0.293, −0.079]). No significant interaction effect between Congruency and Group was found ($F_{(1,46)} = 0.726$, $\eta_p^2 = 0.016$). After normalization, the significant main effect of Group disappeared ($F_{(1,46)} = 0.115$, $\eta_p^2 = 0.002$; 95% CI of difference between HA and LA [−0.015, 0.011]). The independent-sample BF $t$-test provided substantial evidence for the null hypothesis of no difference between HA and LA regarding the integration effect ($BF_{10} = 0.301$; Dienes, 2014). Furthermore, ANCOVA revealed same patterns with ANOVA. The ANCOVA in corrected performance revealed main effects of Congruency ($F_{(1,44)} = 92.453$, $p < .001$, $\eta_p^2 = 0.678$; Con > Incon; 95% CI

**TABLE 3** Behavioral and electrophysiological responses in each experimental condition of each group

| | HA | | LA | |
|---|---|---|---|---|
| | Con | Incon | Con | Incon |
| Correct performance | $1.46 \pm 0.20$ | $1.37 \pm 0.19$ | $1.65 \pm 0.18$ | $1.55 \pm 0.17$ |
| Integrative effect | $0.03 \pm 0.02$ | | $0.03 \pm 0.02$ | |
| Intensity rating | $5.85 \pm 0.77$ | | $5.72 \pm 1.42$ | |
| N1 | $-4.81 \pm 1.99$ | $-4.59 \pm 2.27$ | $-4.83 \pm 2.12$ | $-4.56 \pm 1.94$ |
| N2 | $-3.73 \pm 2.57$ | $-4.02 \pm 2.66$ | $-3.65 \pm 2.32$ | $-4.00 \pm 2.64$ |
| P2 | $2.02 \pm 2.07$ | $2.04 \pm 1.90$ | $1.20 \pm 1.72$ | $1.13 \pm 1.84$ |
| P2_indi | $2.33 \pm 2.20$ | $2.36 \pm 2.02$ | $1.58 \pm 1.71$ | $1.51 \pm 1.84$ |

*Note::* Descriptive data are presented as mean (standard deviation).

Abbreviations: Con, congruent condition; HA, individuals with alexithymia; Incon, incongruent condition; LA, individuals without alexithymia; P2_indi, P2 amplitudes based on individual peak.
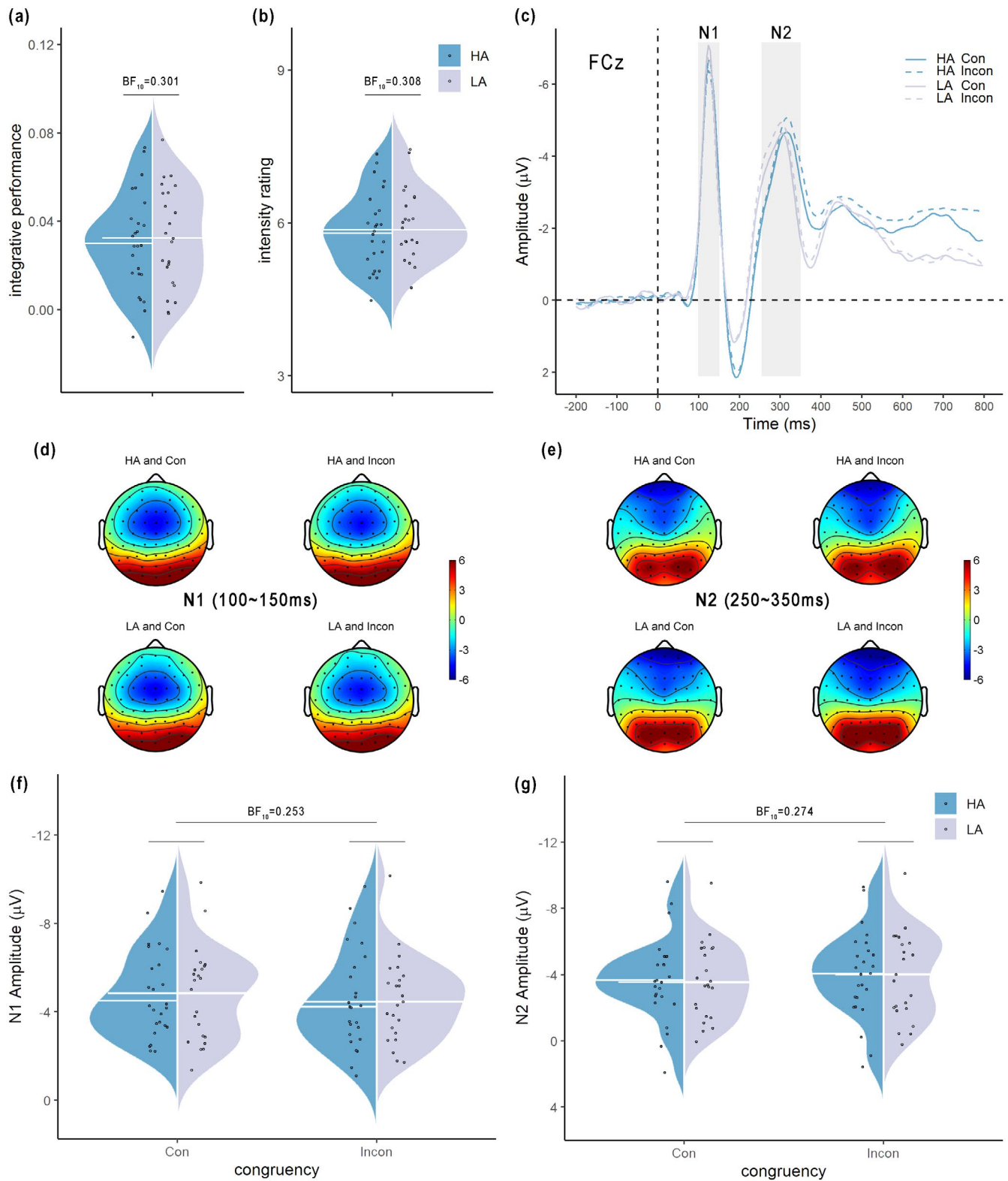
**FIGURE 2** Group differences in eMSI effect. (a) Integrative performance between HA and LA. (b) Intensity rating between HA and LA. (c) Time course at FCz electrode. (d & e) Topographic maps of each condition in N1 and N2. (f & g) Mean amplitude values of N1 and N2 in Con versus Incon between HA and LA. Con, congruent condition; HA, individuals with alexithymia; Incon, incongruent condition; LA, individuals without alexithymia

of difference between Con and Incon [0.075, 0.115]) and Group ($F_{(1,44)} = 4.325$, $p = .043$, $\eta_p^2 = 0.089$; HA < LA; 95% CI of difference between HA and LA [−0.329, −0.005];

Figure 3a). No significant interaction effect between Congruency and Group was found ($F_{(1,44)} = 0.002$, $\eta_p^2 < 0.001$). After normalization, the significant main effect of Group

disappeared ($F_{(1,44)} = 0.206$, $\eta_p^2 = 0.005$; 95% CI of difference between HA and LA [−0.015, 0.023]). The independent-sample BF $t$-test provided substantial evidence for the null hypothesis of no difference between HA and LA regarding the integration effect ($BF_{10} = 0.301$; Figure 2a; Dienes, 2014).

For intensity ratings, no significant group main effect was found from ANOVA ($F_{(1,46)} = 0.164$, $\eta_p^2 = 0.004$; 95% CI of difference between HA and LA [−0.524, 0.788]). Similarly, for this nonsignificant effect, the BF was calculated to test whether the negative result supported the alternative or the null hypothesis. Again, substantial evidence for the null hypothesis of an absence of group differences for intensity ratings was provided by the independent-sample BF $t$-test ($BF_{10} = 0.308$; Dienes, 2014). The same pattern was replicated when adding covariates. ANCOVA in intensity ratings found no significant group main effect ($F_{(1,44)} = 0.939$, $\eta_p^2 = 0.021$; 95% CI of difference between HA and LA [−0.510, 1.454]). Similarly, for this nonsignificant effect, the BF was calculated to test whether the negative result supported the alternative or the null hypothesis. Again, substantial evidence for the null hypothesis of an absence of group differences for

intensity ratings was provided by the independent-sample BF $t$-test ($BF_{10} = 0.308$; Figure 2b; Dienes, 2014).

## 3.2 | ERP analyses

### 3.2.1 | N1

For the N1, ANOVA showed a significant main effect of Congruency ($F_{(1,46)} = 5.244$, $p = .027$, $\eta_p^2 = 0.102$; 95% CI of difference between Con and Incon [−0.459, −0.030]). Specifically, Con elicited a more negative N1 than the Incon condition. No other significant effect was observed (main effect of Group: $F_{(1,46)} < 0.001$, $\eta_p^2 < 0.001$; 95% CI of difference between HA and LA [−1.206, 1.182]; interaction effect between Group and Congruency: $F_{(1,46)} = 0.053$, $\eta_p^2 = 0.001$). We then calculated the BF for this nonsignificant effect. The BF ANOVA test confirmed that there was no interaction effect between Group and Congruency for the N1 ($BF_{10} = 0.306$; Dienes, 2014). Importantly, we found same patterns in ANCOVA. Regarding the N1, there was a significant main effect of Congruency ($F_{(1,44)} = 4.970$, $p = .031$, $\eta_p^2 = 0.101$;
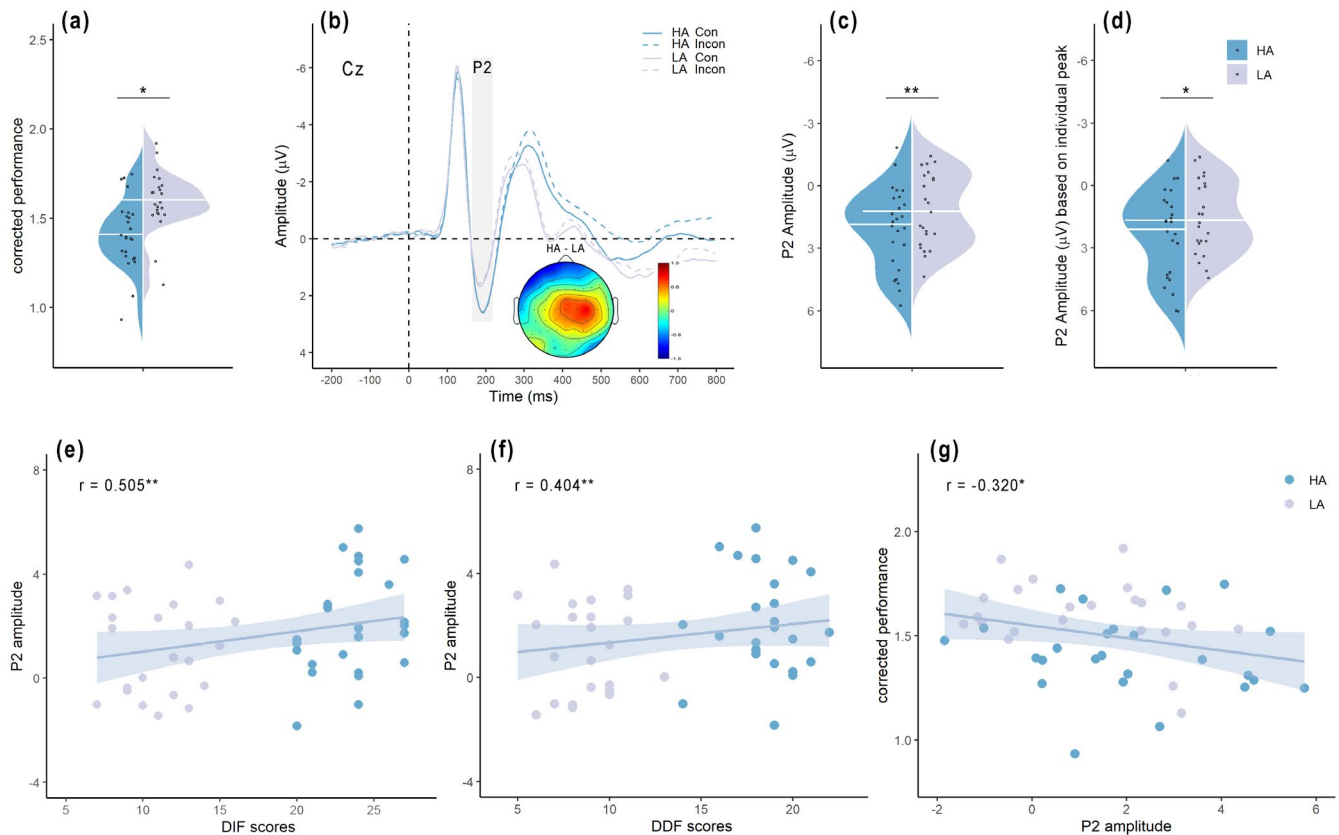


**FIGURE 3** Group differences (independent of congruency effect). (a) Corrected performance between HA and LA. (b) Time course at Cz electrode and the topographic map of the difference between HA and LA. (c) Mean amplitude values of P2 between HA and LA. (d) Mean amplitude values of P2 based on individual peak between HA and LA. (e-g) Correlations between DIF and P2, between DDF and P2, and between P2 and corrected performance. Con, congruent condition; Incon, incongruent condition; DDF, difficulty describing feelings; DIF, difficulty identifying feelings; HA, individuals with alexithymia; LA, individuals without alexithymia. **$p < .01$; *$p < .05$

95% CI of difference between Con and Incon [−0.462, −0.023]). Specifically, Con elicited a more negative N1 than the Incon condition. No other significant effect was observed (main effect of Group: $F_{(1,44)} = 0.333$, $\eta_p^2 = 0.008$; 95% CI of difference between HA and LA [−1.281, 2.309]; interaction effect between Group and Congruency: $F_{(1,44)} = 0.014$, $\eta_p^2 < 0.001$). We then calculated the BF for this nonsignificant effect. The BF ANCOVA test confirmed that there was no interaction effect between Group and Congruency for the N1 ($BF_{10} = 0.253$; Figure 2f; Dienes, 2014).

## 3.2.2 | P2

With regard to the P2, we did not observe a significant main Group effect from ANOVA ($F_{(1,46)} = 2.553$, $p = .117$, $\eta_p^2 = 0.053$; 95% CI of difference between HA and LA [−0.224, 1.947]). No other significant effect was found (main effect of Congruency: $F_{(1,46)} = 0.066$, $\eta_p^2 = 0.001$; 95% CI of difference between Con and Incon [−0.156, 0.202]; interaction effect between Group and Congruency: $F_{(1,46)} = 0.265$, $\eta_p^2 = 0.006$). We observed a significant main Group effect from ANCOVA ($F_{(1,44)} = 7.974$, $p = .007$, $\eta_p^2 = 0.153$; HA > LA; 95% CI of difference between HA and LA [0.615, 3.683]; Figure 3c). No other significant effect was found [main effect of Congruency: $F_{(1,44)} = 0.036$, $\eta_p^2 = 0.001$; 95% CI of difference between Con and Incon (−0.160, 0.194); interaction effect between Group and Congruency: $F_{(1,44)} = 0.617$, $\eta_p^2 = 0.014$]. To eliminate the impact of the previous negative-going (N1) peak on P2, we also computed the average waveform (50 ms) around the individual peak to validate our result. A similar pattern was observed: a significant main effect of Group ($F_{(1,44)} = 5.775$, $p = .021$, $\eta_p^2 = 0.116$; HA > LA; 95% CI of difference between HA and LA [0.311, 3.540]; Figure 3d) and no other significant effect (main effect of Congruency: $F_{(1,44)} = 0.023$, $\eta_p^2 = 0.001$; 95% CI of difference between Con and Incon [−0.167, 0.193]; interaction effect between Group and Congruency: $F_{(1,44)} = 0.835$, $\eta_p^2 = 0.019$). The significant effect of alexithymia on P2 disappeared when not controlling for anxiety, depression, and autism. It indicates that the absence of the significant group effect on P2 might be contaminated by these potential variables.

## 3.2.3 | N2

For the N2, ANOVA revealed a significant main effect of Congruency ($F_{(1,46)} = 6.715$, $p = .013$, $\eta_p^2 = 0.127$; 95% CI of difference between Con and Incon [0.071, 0.569]), whereas no significant main effect of Group or interaction effect between Congruency and Group was found (main effect of

Group: $F_{(1,46)} = 0.005$, $\eta_p^2 < 0.001$; 95% CI of difference between HA and LA [−1.514, 1.415]; interaction effect between Congruency and Group: $F_{(1,46)} = 0.072$, $\eta_p^2 = 0.002$). Specifically, Con evoked a less negative N2 than Incon. Once again, we conducted Bayesian statistics for the interaction effect between Congruency and Group to test whether this negative result supported the alternative or null hypothesis. The BF ANOVA test showed that there was no interaction effect between Group and Congruency for the N2 ($BF_{10} = 0.287$; Dienes, 2014). Importantly, ANCOVA replicated the results from ANOVA. A significant main effect of Congruency was found from ANCOVA ($F_{(1,44)} = 6.613$, $p = .014$, $\eta_p^2 = 0.131$; 95% CI of difference between Con and Incon [0.070, 0.576]), whereas no significant main effect of Group or interaction effect between Congruency and Group were found (main effect of Group: $F_{(1,44)} = 0.006$, $\eta_p^2 < 0.001$; 95% CI of difference between HA and LA [−2.128, 2.301]; interaction effect between Congruency and Group: $F_{(1,44)} = 0.296$, $\eta_p^2 = 0.007$). Specifically, Con evoked a less negative N2 than Incon. Once again, we conducted Bayesian statistics for the interaction effect between Congruency and Group to test whether this negative result supported the alternative or null hypothesis. The BF ANCOVA test showed that there was no interaction effect between Group and Congruency for the N2 ($BF_{10} = 0.274$; Figure 2g; Dienes, 2014).

Notably, the significant integrative effects (Con vs. Incon) at both behavioral and neural levels indicate a successful manipulation in the current study, consistent with previous studies comparing Con and Incon audiovisual conditions (Campanella & Belin, 2007; Klasen et al., 2012).

## 3.3 | Correlations among subscales of TAS-20, P2 and corrected performance

Given that the significant group difference in P2 was observed only when adding covariates, we conducted partial correlations among subscales of the TAS-20 (difficulty identifying feelings [DIF], difficulty describing feelings [DDF], and externally oriented thinking [EOT]), the P2, and corrected behavioral performance (with BAI, BDI and AQ as covariates). These post hoc analyses revealed a positive correlation between DIF and P2 ($r_{(43)} = .505$, $p = .004$; Figure 3e), between DDF and P2 ($r_{(43)} = .404$, $p = .006$; Figure 3f), and a negative correlation between P2 and corrected performance ($r_{(43)} = −.320$, $p = .032$; Figure 3g). Please note that analyses without covariates were as planned initially (e.g., ANOVA), whereas analyses with covariates were conducted post hoc, including ANCOVAs, partial correlations, and Bayesian statistics.

# 4 | DISCUSSION

The current study examined whether alexithymia is related to deficits in the integration of multisensory emotional information. Bayesian statistics showed no group differences in classical integration-related ERP components (N1 and N2). These results do not support the assumption of abnormal emotional multisensory integration (eMSI; hypo- and hyper integration) in relation to alexithymia. Instead, HA appear to show normal behaviors and neuropsychological responses during eMSI. However, the findings of worse performance and higher P2 amplitudes in HA versus LA, independent of emotional congruency, do support the notion that HA have deficits processing auditory emotional information.

In light of two stages of eMSI (Klasen et al., 2012; Wang et al., 2012), we observed that Con elicited an increased N1 and a decreased N2 compared with Incon. The larger N1 may reflect amplification of attention: that is attention triggered by stimulation of one sensory modality can be amplified by stimulation of another sensory modality (Ho et al., 2013; Klasen et al., 2012; Luck et al., 2000). The N1 effect has also been demonstrated in blind-sight patients who cannot consciously perceive a visual stimulus (but are able to react to it, implying that stimulus processing took place outside of conscious awareness), reflecting autonomous eMSI in the early stage (de Gelder et al., 2002). Therefore, the observed larger N1 in Con versus Incon conditions may indicate that attention to auditory stimuli, unconsciously amplified by implicit visual (i.e., facial) processing, facilitates early processing of eMSI in the congruent context, resulting in better performance (i.e., multisensory facilitation). The increased N2 has been repeatedly reported for processing incongruent compared with congruent stimuli in the Stroop task (Folstein & Van Petten, 2008), reflecting conflict processing (Mittag et al., 2013; Zinchenko et al., 2017). The larger N2 evoked in the current Incon versus Con conditions may indicate inhibition of conflicting stimuli (multisensory inhibition).

However, the current results failed to support the hypothesis of either hypo- or hyper-integration in relation to alexithymia. Instead, our evidence supports normal behavioral and neuropsychological responses of eMSI in HA, resulting from Bayesian analysis. Bayes Factor has long been considered as a central tool to enable conclusions that could support a null hypothesis over a specific hypothesis (Gallistel, 2009). To counteract the problem of publication bias, reporting a lack of differences is as important as are claims of "significant effects" (Gallistel, 2009; Tendeiro & Kiers, 2019). Indeed, being able to provide substantial evidence for the absence of differences plays a crucial role in the development of theories in neuroscience and cognitive science (Dienes, 2014). Based on two stages of eMSI (Klasen et al., 2012; Wang et al., 2012), observed absence of N1 difference in Con versus Incon between HA and LA suggests that individuals with high alexithymia levels are able to successfully integrate an attended emotional voice with an unattended facial expression, thus implying intact emotional multisensory facilitation. It goes without saying, however, that strong conclusions regarding intact eMSI in alexithymia can only be made with more studies (using different tasks) supporting our finding.

No N2 difference in Con versus Incon between HA and LA were observed, and integrative behavioral performance was also unaffected. This may reflect successful inhibition of the incongruent emotional information in individuals with high alexithymia levels. However, previous studies did report deficits in inhibiting conflict processing (Zhang et al., 2011), especially in emotional conflicting inhibition (Zhang et al., 2012). One possible explanation is the task difference. Although previous studies adopted ANT and Go/Nogo tasks with unisensory stimuli (Zhang et al., 2011, 2012), we, instead, used a Stroop-like task with multisensory (visual-auditory) stimuli. Consequently, the absence of alexithymia-related differences in conflict inhibition may be task-specific or modality-specific. Another possible reason could be that the TAS-20 may not be sensitive to alexithymia-related differences in conflict inhibition. Koven and Thomas (2010) used principal components analysis on the TAS-20 and the Mood Awareness Scale to decompose alexithymia into two facets: emotional clarity and emotional monitoring. They found that only the emotional clarity component correlated with behavioral inhibition (Koven & Thomas, 2010). Thus, it is possible that the use of the TAS-20 as the only alexithymia measure may have prevented us from detecting conflict inhibition differences in relation to alexithymia. Future studies should employ multiple alexithymia assessment tools to identify the nature of differences in behavioral inhibition and executive control in relation to alexithymia.

It should be noted that a previous study (with western participants) found that alexithymia was associated with deficits in emotional integration during an audiovisual affective priming task with a stimulus onset asynchrony (SOA) of 200 ms (the presentation of audio stimuli occurred 200 ms after visual stimuli; Goerlich et al., 2011). One potential explanation for the different results of our current and the previous study might be the asynchrony between auditory and visual stimuli in that study. Indeed, time delay has been reported to be essential for congruency effects (Föcker & Röder, 2019). It has been demonstrated that the maximum eMSI appeared when the temporal difference between auditory and visual stimuli was less than 100 ms, and integrative performance decreased as the SOA increased (Ren et al., 2017). It is conceivable that the 200 ms SOA used in the study by Goerlich and colleagues (2011) might be a limiting factor for the integration of emotions from multisensory stimulus modalities, revealing alexithymia-related differences. It would be interesting to test the effect of asynchrony of stimulus presentation in future studies on eMSI in relation to alexithymia. Another reason

might be attributed to cultural differences between western participants in the previous study (Goerlich et al., 2011) and eastern participants in our present study. It has been suggested that culture plays an important role in emotion processing in relation to alexithymia (Zhu et al., 2007). Indeed, experience and expression of emotions, defining features of alexithymia, are significantly different between eastern and western cultures (Le et al., 2002). Therefore, it is possible that cultural differences contributed to the absence of eMSI differences in relation to alexithymia in the present study. One may argue that the absence of the behavioral performance difference between Con and Incon in HA versus LA may be due to the similarity of the emotions. Indeed, the angry and sad stimuli we used are similar in that both convey negative information. However, despite sharing some characteristics, different emotion expressions are thought to serve different functions (Horstmann, 2003). For example, expressions of anger (an approach emotion) are thought to cue the perceiver to modulate his or her behavior, whereas expressions of sadness (a withdrawal emotion) are thought to serve an emotion-sharing function (Garcia & Tully, 2020). We also found significant congruency effects (Con vs. Incon) both at behavioral and neural (N1 and N2) levels, further indicating that angry and sad emotions were perceived as different stimuli. Thus, the absence of the behavioral performance difference between Con and Incon in HA versus LA was likely not due to the similarity of the emotions. Instead, the absence of effects, in combination with Bayesian analysis, support normal behavioral and neuropsychological responses of eMSI in HA.

We also investigated abnormal emotion recognition from the voice and its corresponding neuropsychological mechanism. It is well-known that alexithymia impacts emotion recognition from both audio and visual emotional stimuli (Franz et al., 2004; Goerlich-Dobre et al., 2014; Kano et al., 2003; Lee et al., 2011; Van der Velde et al., 2013). For example, HA exhibited a larger P2 during the processing of negative pictures than LA (Franz et al., 2004). The P2 component is linked to the early processing of categorizing different stimuli, especially negative emotions (Crowley & Colrain, 2004; Zhu et al., 2015). The P2 component has also been implicated in the involvement of attentional resources (Mercado et al., 2006; Miller et al., 2011). Indeed, it has been reported that HA need more attentional resources to process emotional information (Franz et al., 2004; Van der Velde et al., 2013). Therefore, worse performance and increased P2 in HA versus LA in the current study may reflect that individuals with high alexithymia levels called upon attentional resources to a stronger extent to categorize vocal emotions. This deficit was further corroborated by correlations between the DIF and DDF subscales of alexithymia and P2, and between P2 and corrected behavioral performance. Recent reviews on alexithymia have identified DIF and—to a somewhat lesser extent—DDF as the most important TAS-20 contributors

to deficits in perceiving and recognizing negative emotions during the early, automatic (unconscious) processing level (Donges & Suslow, 2017; Goerlich, 2018). The present finding that P2 was correlated with DIF and DDF provides further evidence for altered emotion processing at early, automatic stages in relation to alexithymia and suggests that individuals with difficulty identifying and describing their feelings need more attentional resources to recognize negative emotions from the voice. However, these conclusions should be made cautiously given poor Cronbach's alpha coefficient of DDF.

Several limitations of the present study should be mentioned. First, we adopted the Stroop-like visual-auditory task with the judgement of auditory emotion only (rather than also including the visual judgement task). Although stronger eMSI effects were found in the voice task than in the visual task in Chinese participants (Wang et al., 2012), whether our findings can be extended to the visual task remains to be tested. Second, static facial pictures were used, whereas in real life, we recognize emotion from dynamic visual information, for example facial movements (Collignon et al., 2008; Jessen & Kotz, 2011; Kokinous et al., 2015). Future studies should aim to explore eMSI in a dynamic context for higher ecological validity. Third, we only adopted negative stimuli in the current study. It is thus not clear whether the current conclusions would extrapolate to neutral or positive emotions during eMSI. Although differences in processing negative emotions seem to be at the core of alexithymia ((Kano et al., 2003; Lee et al., 2011), both positive and negative emotions are associated with processing differences in alexithymia (Van der Velde et al., 2013). Future studies would thus benefit from investigating positive, neutral, and negative emotion integration in relation to alexithymia. Next, we did not measure the sound intensity participants received, despite the approximate equal intensity of all vocal stimuli. Participants also reported that they can hear the voice clearly without any uncomfortable state. However, it has been demonstrated that intensity affects eMSI (Pan et al., ). Future studies might thus benefit from assessing perceived sound intensities. Finally, although the voice stimuli selected from the dataset were matched in emotional properties between angry and sad stimuli, we did observe stronger intensity and more negative valence being perceived from angry than sad utterances. Importantly, however, the absence of significant group differences for these parameters suggests that the current result of intact eMSI in alexithymia was not confounded by basic stimulus characteristics.

To conclude, both behavioral and electrophysiological data indicated intact emotional multisensory integration in relation to alexithymia. These results do not support alexithymia-related differences in the integration of emotional information. However, we provide evidence (replicating previous research) for an emotion recognition deficit from the voice and show that this deficit is driven by the DIF

and DDF factors of the TAS-20. Our work provides valuable insights into the effects of alexithymia on the neuropsychological mechanisms of eMSI and the processing of emotions expressed by the voice.

## CONFLICT OF INTEREST

The authors have indicated they have no potential conflicts of interest to disclose.

## AUTHOR CONTRIBUTION

**Zhihao Wang:** Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Project administration; Visualization; Writing-original draft; Writing-review & editing. **Mai Chen:** Conceptualization; Investigation; Methodology; Project administration; Visualization; Writing-original draft. **Katharina S. Goerlich:** Conceptualization; Supervision; Writing-review & editing. **André Aleman:** Conceptualization; Supervision; Writing-review & editing. **Pengfei Xu:** Conceptualization; Funding acquisition; Supervision; Writing-review & editing. **Yuejia Luo:** Conceptualization; Funding acquisition; Supervision; Writing-review & editing.

## DATA AVAILABILITY STATEMENT

The data are available from the corresponding author P. Xu, upon reasonable request.

## ORCID

*Zhihao Wang* https://orcid.org/0000-0001-6292-9307

## REFERENCES

Bagby, R. M., Parker, J. D. A., & Taylor, G. J. (1994). The twenty-item Toronto Alexithymia scale—I. Item selection and cross-validation of the factor structure. *Journal of Psychosomatic Research*, *38*(1), 23–32. https://doi.org/10.1016/0022-3999(94)90005-1

Bagby, R. M., Taylor, G. J., & Parker, J. D. A. (1994). The twenty-item Toronto Alexithymia scale—II. Convergent, discriminant, and concurrent validity. *Journal of Psychosomatic Research*, *38*(1), 33–40. https://doi.org/10.1016/0022-3999(94)90006-X

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The Autism-Spectrum Quotient (AQ): Evidence from …. *Journal of Autism and Developmental Disorders*, *31*(1), 5–17.

Beck, A. T. (1967). *Depression: Clinical, experimental, and theoretical aspects*. University of Pennsylvania Press.

Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology*, *56*(6), 893–897. https://doi.org/10.1037/0022-006X.56.6.893

Bird, G., & Cook, R. (2013). Mixed emotions: The contribution of alexithymia to the emotional symptoms of autism. *Translational Psychiatry*, *3*(7), e285. https://doi.org/10.1038/tp.2013.61

Boersma, P., & Weenink, D. (2021). *Praat: Doing phonetics by computer* (Version 6.1.16). http://www.praat.org

Campanella, S., & Belin, P. (2007). Integrating face and voice in person perception. *Trends in Cognitive Sciences*, *11*(12), 535–543. https://doi.org/10.1016/j.tics.2007.10.001

Campanella, S., Bruyer, R., Froidbise, S., Rossignol, M., Joassin, F., Kornreich, C., Noël, X., & Verbanck, P. (2010). Is two better than one? A cross-modal oddball paradigm reveals greater sensitivity of the P300 to emotional face-voice associations. *Clinical Neurophysiology*, *121*(11), 1855–1862. https://doi.org/10.1016/j.clinph.2010.04.004

Chen, X., Pan, Z., Wang, P., Yang, X., Liu, P., You, X., & Yuan, J. (2016). The integration of facial and vocal cues during emotional change perception: EEG markers. *Social Cognitive and Affective Neuroscience*, *11*(7), 1152–1161. https://doi.org/10.1093/scan/nsv083

Collignon, O., Girard, S., Gosselin, F., Roy, S., Saint-Amour, D., Lassonde, M., & Lepore, F. (2008). Audio-visual integration of emotion expression. *Brain Research*, *1242*, 126–135. https://doi.org/10.1016/j.brainres.2008.04.023

Crowley, K. E., & Colrain, I. M. (2004). A review of the evidence for P2 being an independent component process: Age, sleep and modality. *Clinical Neurophysiology*, *115*(4), 732–744. https://doi.org/10.1016/j.clinph.2003.11.021

De Gelder, B., & Bertelson, P. (2003). Multisensory integration, perception and ecological validity. *Trends in Cognitive Sciences*, *7*(10), 460–467. https://doi.org/10.1016/j.tics.2003.08.014

de Gelder, B., Pourtois, G., & Weiskrantz, L. (2002). Fear recognition in the voice is modulated by unconsciously recognized facial expressions but not by unconsciously recognized affective pictures. *Proceedings of the National Academy of Sciences*, *99*(6), 4121–4126. https://doi.org/10.1073/pnas.062018499

Delle-Vigne, D., Kornreich, C., Verbanck, P., & Campanella, S. (2014). Subclinical alexithymia modulates early audio-visual perceptive and attentional event-related potentials. *Frontiers in Human Neuroscience*, *8*, 1–14. https://doi.org/10.3389/fnhum.2014.00106

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, *5*, 1–17. https://doi.org/10.3389/fpsyg.2014.00781

Donges, U. S., & Suslow, T. (2017). Alexithymia and automatic processing of emotional stimuli: A systematic review. *Reviews in the Neurosciences*, *28*(3), 247–264. https://doi.org/10.1515/revneuro-2016-0049

Donohue, S. E., Todisco, A. E., & Woldorff, M. G. (2013). The rapid distraction of attentional resources toward the source of

incongruent stimulus input during multisensory conflict. *Journal of Cognitive Neuroscience*, *25*(4), 623–635. https://doi.org/10.1162/jocn_a_00336

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Föcker, J., & Röder, B. (2019). Event-related potentials reveal evidence for late integration of emotional prosody and facial expression in dynamic stimuli: An ERP study. *Multisensory Research*, *32*(6), 473–497. https://doi.org/10.1163/22134808-20191332

Folstein, J. R., & Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: A review. *Psychophysiology*, *45*(1), 152–170. https://doi.org/10.1111/j.1469-8986.2007.00602.x

Franz, M., Schaefer, R., Schneider, C., Sitte, W., & Bachor, J. (2004). Visual event-related potentials in subjects with alexithymia: Modified processing of emotional aversive information? *American Journal of Psychiatry*, *161*(4), 728–735. https://doi.org/10.1176/appi.ajp.161.4.728

Gallistel, C. R. (2009). The importance of proving the null. *Psychological Review*, *116*(2), 439–453. https://doi.org/10.1037/a0015251

Garcia, S. E., & Tully, E. C. (2020). Children's recognition of happy, sad, and angry facial expressions across emotive intensities. *Journal of Experimental Child Psychology*, *197*, 104881. https://doi.org/10.1016/j.jecp.2020.104881

Goerlich, K. S. (2018). The multifaceted nature of alexithymia - A neuroscientific perspective. *Frontiers in Psychology*, *9*, 1–7. https://doi.org/10.3389/fpsyg.2018.01614

Goerlich, K. S., Witteman, J., Aleman, A., & Martens, S. (2011). Hearing feelings: Affective categorization of music and speech in alexithymia, an ERP study. *PLoS One*, *6*(5), e19501. https://doi.org/10.1371/journal.pone.0019501

Goerlich-Dobre, K. S., Witteman, J., Schiller, N. O., van Heuven, V. J. P., Aleman, A., & Martens, S. (2014). Blunted feelings: Alexithymia is associated with a diminished neural response to speech prosody. *Social Cognitive and Affective Neuroscience*, *9*(8), 1108–1117. https://doi.org/10.1093/scan/nst075

Gong, X., Huang, Y. X., Wang, Y., & Luo, Y. J. (2011). Revision of the Chinese facial affective picture system. *Chinese Mental Health Journal*, *25*(1), 40–46.

Hendryx, M. S., Haviland, M. G., & Shaw, D. G. (1991). Dimensions of alexithymia and their relationships to anxiety and depression. *Journal of Personality Assessment*, *56*(2), 227–237. https://doi.org/10.1207/s15327752jpa5602_4

Ho, H. T., Schröger, E., & Kotz, S. A. (2013). Selective attention modulates early human evoked potentials during emotional face–voice processing. *Journal of Cognitive Neuroscience*, *2015*, 1–10. https://doi.org/10.1162/jocn

Horstmann, G. (2003). What do facial expressions convey: Feeling states, behavioral intentions, or actions requests? *Emotion*, *3*(2), 150–166. https://doi.org/10.1037/1528-3542.3.2.150

JASP Team. (2020). *JASP (Version 0.14.1)*. https://jasp-stats.org/

Jessen, S., & Kotz, S. A. (2011). The temporal dynamics of processing emotions from vocal, facial, and bodily expressions. *NeuroImage*, *58*(2), 665–674. https://doi.org/10.1016/j.neuroimage.2011.06.035

Kano, M., Fukudo, S., Gyoba, J., Kamachi, M., Tagawa, M., Mochizuki, H., Itoh, M., Hongo, M., & Yanai, K. (2003). Specific brain processing of facial expressions in people with alexithymia: An H215O-PET study. *Brain*, *126*(6), 1474–1484. https://doi.org/10.1093/brain/awg131

Klasen, M., Chen, Y. H., & Mathiak, K. (2012). Multisensory emotions: Perception, combination and underlying neural processes. *Reviews in the Neurosciences*, *23*(4), 381–392. https://doi.org/10.1515/revneuro-2012-0040

Kokinous, J., Kotz, S. A., Tavano, A., & Schröger, E. (2015). The role of emotion in dynamic audiovisual integration of faces and voices. *Social Cognitive and Affective Neuroscience*, *10*(5), 713–720. https://doi.org/10.1093/scan/nsu105

Koven, N. S., & Thomas, W. (2010). Mapping facets of alexithymia to executive dysfunction in daily life. *Personality and Individual Differences*, *49*(1), 24–28. https://doi.org/10.1016/j.paid.2010.02.034

Le, H.-N., Berenbaum, H., & Raghavan, C. (2002). Culture and alexithymia: Mean levels, correlates and the role of parental socialization of emotions. *Emotion*, *2*(4), 341–360. https://doi.org/10.1037/1528-3542.2.4.341

Lee, B. T., Lee, H. Y., Park, S. A., Lim, J. Y., Tae, W. S., Lee, M. S., Joe, S. H., Jung, I. K., & Ham, B. J. (2011). Neural substrates of affective face recognition in alexithymia: A functional magnetic resonance imaging study. *Neuropsychobiology*, *63*(2), 119–124. https://doi.org/10.1159/000318086

Li, S., Zhang, B., Guo, Y., & Zhang, J. (2015). The association between alexithymia as assessed by the 20-item Toronto Alexithymia Scale and depression: A meta-analysis. *Psychiatry Research*, *227*(1), 1–9. https://doi.org/10.1016/j.psychres.2015.02.006

Liu, P., & Pell, M. D. (2012). Recognizing vocal emotions in Mandarin Chinese: A validated database of Chinese vocal emotional stimuli. *Behavior Research Methods*, *44*(4), 1042–1051. https://doi.org/10.3758/s13428-012-0203-3

Liu, P., Rigoulot, S., & Pell, M. D. (2015). Culture modulates the brain response to human expressions of emotion: Electrophysiological evidence. *Neuropsychologia*, *67*, 1–13. https://doi.org/10.1016/j.neuropsychologia.2014.11.034

Luck, S. J. (2014). *An Introduction to the event-related potential technique*. MIT press.

Luck, S. J., Woodman, G. F., & Vogel, E. K. (2000). Event-related potential studies of attention. *Trends in Cognitive Sciences*, *4*(11), 432–440. https://doi.org/10.1016/S1364-6613(00)01545-X

Luminet, O., Bagby, R. M., & Taylor, G. J. (Eds.). (2018). *Alexithymia*. Cambridge University Press. https://doi.org/10.1017/9781108241595

Mercado, F., Carretié, L., Tapia, M., & Gómez-Jarabo, G. (2006). The influence of emotional context on attention in anxious subjects: Neurophysiological correlates. *Journal of Anxiety Disorders*, *20*(1), 72–84. https://doi.org/10.1016/j.janxdis.2004.10.003

Miller, M. W., Rietschel, J. C., McDonald, C. G., & Hatfield, B. D. (2011). A novel approach to the physiological measurement of mental workload. *International Journal of Psychophysiology*, *80*(1), 75–78. https://doi.org/10.1016/j.ijpsycho.2011.02.003

Mittag, M., Alho, K., Takegata, R., Makkonen, T., & Kujala, T. (2013). Audiovisual attention boosts letter-speech sound integration. *Psychophysiology*, *50*(10), 1034–1044. https://doi.org/10.1111/psyp.12085

Müller, V. I., Cieslik, E. C., Kellermann, T. S., & Eickhoff, S. B. (2013). Crossmodal emotional integration in major depression. *Social Cognitive and Affective Neuroscience*, *9*(6), 839–848. https://doi.org/10.1093/scan/nst057

Pan, Z., Liu, X., Luo, Y., & Chen, X. (2017). Emotional intensity modulates the integration of bimodal angry expressions: ERP evidence. *Frontiers in Neuroscience*, *11*(349), 1–9. https://doi.org/10.3389/fnins.2017.00349

Ren, Y., Yang, W., Nakahashi, K., Takahashi, S., & Wu, J. (2017). Audiovisual integration delayed by stimulus onset asynchrony between auditory and visual stimuli in older adults. *Perception*, *46*(2), 205–218. https://doi.org/10.1177/0301006616673850

Stein, B. E., & Stanford, T. R. (2008). Multisensory integration: Current issues from the perspective of the single neuron. *Nature Reviews Neuroscience*, *9*(4), 255–266. https://doi.org/10.1038/nrn2331

Tang, X., Wu, J., & Shen, Y. (2016). The interactions of multisensory integration with endogenous and exogenous attention. *Neuroscience and Biobehavioral Reviews*, *61*, 208–224. https://doi.org/10.1016/j.neubiorev.2015.11.002

Taylor, G. J., Bagby, R. M., & Parker, J. D. (2003). The 20-item Toronto alexithymia scale. *Journal of Psychosomatic Research*, *55*(3), 277–283. https://doi.org/10.1016/S0022-3999(02)00601-3

Taylor, G. J., Bagby, R. M., Ryan, D. P., Parker, J. D., Doody, K. F., & Keefe, P. (1988). Criterion validity of the Toronto alexithymia scale. *Psychosomatic Medicine*, *50*(5), 500–509. https://doi.org/10.1097/00006842-198809000-00006

Tendeiro, J. N., & Kiers, H. A. L. (2019). A review of issues about null hypothesis Bayesian testing. *Psychological Methods*, *24*(6), 774–795. https://doi.org/10.1037/met0000221

Van der Velde, J., Servaas, M. N., Goerlich, K. S., Bruggeman, R., Horton, P., Costafreda, S. G., & Aleman, A. (2013). Neural correlates of alexithymia: A meta-analysis of emotion processing studies. *Neuroscience and Biobehavioral Reviews*, *37*(8), 1774–1785. https://doi.org/10.1016/j.neubiorev.2013.07.008

Wang, W. Y., Hu, L., Valentini, E., Xie, X. B., Cui, H. Y., & Hu, Y. (2012). Dynamic characteristics of multisensory facilitation and inhibition. *Cognitive Neurodynamics*, *6*(5), 409–419. https://doi.org/10.1007/s11571-012-9197-x

Zhang, L., Ye, R., Yu, F., Cao, Z., Zhu, C., Cai, Z., Hu, P., Pu, H., & Wang, K. (2012). How does emotional context modulate response inhibition in alexithymia: Electrophysiological evidence from an ERP study. *PLoS One*, *7*(12), e51110. https://doi.org/10.1371/journal.pone.0051110

Zhang, L., Zhu, C., Ye, R., Cao, Z., Tian, Y., Yang, P., Hu, P., & Wang, K. (2011). Impairment of conflict processing in alexithymic individuals. *Neuroscience Letters*, *504*(3), 261–264. https://doi.org/10.1016/j.neulet.2011.09.043

Zhu, C., He, W., Qi, Z., Wang, L., Song, D., Zhan, L., Yi, S., Luo, Y., & Luo, W. (2015). The time course of emotional picture processing: An event-related potential study using a rapid serial visual presentation paradigm. *Frontiers in Psychology*, *6*, 1–10. https://doi.org/10.3389/fpsyg.2015.00954

Zhu, X., Yi, J., Yao, S., Ryder, A. G., Taylor, G. J., & Bagby, R. M. (2007). Cross-cultural validation of a Chinese translation of the 20-item Toronto Alexithymia Scale. *Comprehensive Psychiatry*, *48*(5), 489–496. https://doi.org/10.1016/j.comppsych.2007.04.007

Zinchenko, A., Obermeier, C., Kanske, P., Schröger, E., Villringer, A., & Kotz, S. A. (2017). The influence of negative emotion on cognitive and emotional control remains intact in aging. *Frontiers in Aging Neuroscience*, *9*(349), 1–16. https://doi.org/10.3389/fnagi.2017.00349

## SUPPORTING INFORMATION

Additional Supporting Information may be found online in the Supporting Information section.