Original scientific paper

UDC: 159.942.3(597)

Received: February, 13, 2026.
Revised: May, 15, 2026.
Accepted: August, 12, 2026.

10.23947/2334-8496-2026-14-2-291-303

Check for updates

Reaction-Time Patterns in an Emotional Face–Word Stroop Task among Lower Secondary School Students: A Pilot Study

Vinh-Long Tran-Chi[1,2], *Gia-Phuoc Tran-Thien*[3*]

[1] University of Health Sciences, Ho Chi Minh City, Vietnam, e-mail: tcvlong@uhsvnu.edu.vn
[2] Vietnam National University, Ho Chi Minh City, Vietnam
[3] Faculty of Psychology, Ho Chi Minh City University of Education, Ho Chi Minh City, Vietnam,
e-mail: tranthiengiaphuoc@gmail.com

Abstract: This pilot study examined whether incongruent facial expressions and Vietnamese emotion labels were associated with longer reaction times than congruent pairings in lower secondary school students. Six Vietnamese students in Grades 6 to 8 completed a 130-trial Emotional Face–Word Stroop Task comprising 70 congruent and 60 incongruent trials. Reaction time was the primary measure, while accuracy was summarized separately. Mean reaction time was longer for incongruent trials than for congruent trials, with an average difference of 175.87 milliseconds and a 95 percent confidence interval from 104.28 to 247.46 milliseconds. Overall accuracy was 81.3 percent and was descriptively lower for incongruent trials than for congruent trials. An exploratory analysis across three broader stimulus groups produced a condition-by-group pattern, but this result was interpreted as task-specific because trial frequencies were unequal and the corresponding cells differed in emotional composition. All participants completed the task, with completion times ranging from 7.76 to 11.12 minutes. The findings showed an observed within-sample congruency pattern in this Vietnamese-language task, but they do not support population-level or stable emotion-specific conclusions. Future studies should use larger and more diverse samples, balanced stimulus categories, prespecified reaction-time preprocessing rules, controlled testing conditions, and sensitivity analyses restricted to correct responses. This pilot provides procedural and behavioral information for refining the task and designing future developmental, educational, and cross-cultural research.

Keywords: *Emotional Stroop Task, attentional interference, reaction time, emotion processing, students.*

Introduction

The Stroop task is a well-established and widely used experimental paradigm for examining interference and cognitive control. First introduced by Stroop (1935), the task requires participants to respond to a task-relevant stimulus dimension while managing interference from a competing, task-irrelevant dimension. In the classic color–word Stroop task, participants name the ink color of words that themselves denote color names. Responses are generally faster and more accurate when the ink color and word meaning are congruent, whereas incongruent trials tend to produce longer reaction times and more errors, a pattern known as the Stroop interference effect (MacLeod, 1991; Stroop, 1935). Contemporary reviews continue to characterize the Stroop effect as robust, while indicating that it may involve contributions from multiple levels of processing and more than one locus of attentional selection (Parris et al., 2022).

Building on the classic paradigm, emotional Stroop tasks have been developed to examine how emotional meaning influences attentional performance. In a common version of the task, emotional and neutral words are presented in different ink colors, and participants are instructed to identify the ink color while ignoring the word meaning (Joyal et al., 2019; Straub et al., 2022; Wielgopolan et al., 2025). Emo-

*Corresponding author: tranthiengiaphuoc@gmail.com

tional words often produce longer reaction times than neutral words, although the magnitude and direction of this difference may vary across stimulus sets, task designs, and participant groups (Joyal et al., 2019; McKenna, 1986; Straub et al., 2022; Wielgopolan et al., 2025; Williams et al., 1996). This reaction-time difference is commonly used as an index of emotional interference or attentional bias, but it does not uniquely identify the cognitive mechanism producing the effect (Joyal et al., 2019; Straub et al., 2022).

Although classic and emotional Stroop tasks both involve interference during task performance, the sources of that interference can differ substantially across paradigms (Müller et al., 2025). In the classic task, interference arises when task-relevant color information competes with task-irrelevant word information and may involve both informational and task conflict (Parris et al., 2022). In emotional Stroop tasks, performance may be influenced by the emotional salience and semantic meaning of the stimuli, as well as by task demands, prior experience, and participant characteristics (Hsieh and Sharma, 2019; Imbir, Pastwa, et al., 2021; McKenna and Sharma, 2004; Williams et al., 1996). Emotional interference may therefore reflect multiple processes, including attentional capture and semantic or conceptual processing, with additional contributions from task-control and memory processes depending on the paradigm and participant characteristics (Hsieh and Sharma, 2019; Imbir, Pastwa, et al., 2021). This distinction is important because reaction-time differences alone cannot determine which processing component is responsible for the observed interference (Parris et al., 2022).

Emotional attention and interference can vary with the modality and configuration of the emotional stimuli used in a task (Lunn and Chen, 2022; Smolker et al., 2022). Face–word Stroop variants combine facial expressions and emotional language, creating conflict when the two sources of emotional information are inconsistent (Jamieson et al., 2023; Kuehne et al., 2019). Recent research indicates that emotional face–word Stroop paradigms can provide behavioral measures related to processing speed, inhibitory control, and emotional information processing in adolescent samples (Banich et al., 2019; Smolker et al., 2022). Studies of inconsistent emotional face–word signals further indicate that performance may involve cognitive control processes that regulate competition between task-relevant and task-irrelevant affective information (Banich et al., 2019; Jamieson et al., 2023; Kuehne et al., 2019). Behavioral measures such as reaction time and accuracy can characterize the interference pattern, but they do not by themselves identify the neural processes underlying that pattern (Jamieson et al., 2023; Kuehne et al., 2019).

Research has also examined whether emotional Stroop interference varies according to stimulus valence. Earlier studies reported greater interference for negative than for neutral or positive stimuli, which was interpreted as preferential attention to potentially threatening information (McKenna and Sharma, 1995; Pérez-Edgar and Fox, 2003). For example, Pérez-Edgar and Fox (2003) found that secondary school students responded more slowly to negative words than to neutral words, whereas responses to positive and neutral words did not differ. More recent evidence indicates that emotion perception can vary with developmental stage and stimulus presentation (Tsenkova et al., 2025). Within the emotional Stroop paradigm, performance may also depend on multiple affective dimensions, including interactions between valence and arousal and effects associated with subjective significance (Imbir, Duda-Goławska, et al., 2021). Emotional stimuli therefore do not produce a uniform interference pattern across task conditions, stimulus characteristics, or participant populations (Imbir, Duda-Goławska, et al., 2021; Joyal et al., 2019).

Several theoretical accounts have been proposed to explain emotional Stroop interference. Attentional-capture accounts propose that emotionally salient stimuli temporarily direct processing resources away from task goals, whereas threat-related accounts emphasize disruption by information with potential motivational relevance (Wentura et al., 2000; Williams et al., 1996). Response- and task-conflict accounts instead propose that interference can arise from competition at the response level and from competition between task sets, with their contributions depending partly on the stimulus–response structure and response requirements of the task (Hsieh and Sharma, 2019; Parris et al., 2022). These perspectives are not mutually exclusive. Rather, they suggest that observed interference may reflect multiple processes whose relative contributions can vary with task design, stimulus characteristics, and participant characteristics (Hsieh and Sharma, 2019; Joyal et al., 2019; Parris et al., 2022; Smolker et al., 2022).

Beyond stimulus characteristics, individual and contextual factors may also shape emotional Stroop performance (Ishikawa et al., 2021; Scrimin et al., 2019). Previous research has linked variation in Stroop-related emotional and attentional interference to anxiety, clinical status, and the personal or diagnostic relevance of the stimuli (Ishikawa et al., 2021; Joyal et al., 2019). In educational settings, students have shown attentional bias toward school-related stressors, with performance also varying according to individual emotional characteristics and classroom context (Scrimin et al., 2016; Scrimin et al., 2019). More broadly, contextual factors such as stress, support, cognitive stimulation, and school or classroom environments have been identified as relevant pathways or protective factors in youth cognitive and academic outcomes (Rakesh et al., 2025). Together, these findings support further examination of emotional interference in school-aged populations in relation to attentional self-regulation, academic stressors, and classroom context (Scrimin et al., 2016; Scrimin et al., 2019). Longitudinal evidence also indicates that emotion regulation is closely linked to school-related well-being during adolescence (Beaumont et al., 2023).

Emotional face–word Stroop tasks combine two socially meaningful sources of emotional information and create conflict when facial expressions and emotion labels are inconsistent (Smolker et al., 2022). Recent research has applied emotional Stroop paradigms to adolescent samples across different cultural settings. For example, Lunn and Chen (2022) examined separate facial and lexical emotional Stroop tasks in Taiwanese adolescents, whereas Smolker et al. (2022) used a large-scale emotional face–word Stroop paradigm to examine task performance and its associations with cognition and psychopathology in a large adolescent cohort. Despite these developments, three gaps remain relevant. First, although word-based emotional Stroop paradigms remain prominent in the literature, comparatively less work has focused on facial–verbal emotional conflict in school-aged samples (Smolker et al., 2022). Second, research specifically focused on children and adolescents remains comparatively limited, particularly relative to the extensive adult and clinical literature (Lunn and Chen, 2022; Smolker et al., 2022). Third, emotional Stroop performance in Vietnamese adolescents remains an underexplored area. More broadly, cognitive neuroscience applications in Vietnamese educational research have been described as still developing (Shih-Tseng and Tran-Chi, 2019). Addressing these gaps is important because developmental stage and cultural context can shape emotion recognition and emotional processing (Möller et al., 2022), while task structure determines how competing facial and verbal emotional information generates response conflict (Smolker et al., 2022). Linguistic adaptation may also require careful consideration when emotional Stroop tasks are used across languages (Lunn and Chen, 2022). A Vietnamese-language pilot can therefore provide useful procedural and behavioral information for future developmental, educational, and cross-cultural studies.

The present small pilot study addressed these gaps using an Emotional Face–Word Stroop Task adapted for Vietnamese lower secondary school students. Unlike conventional emotional Stroop paradigms based on color naming, the present task required participants to judge whether a Vietnamese emotion label matched a facial expression. The study was designed to describe task implementation and within-subject performance patterns rather than to estimate population-level effects. The primary aim was to examine whether incongruent trials were associated with longer reaction times than congruent trials within the present sample. A secondary exploratory aim was to describe whether the condition-related reaction-time pattern varied across broader neutral, positive, and negative stimulus groups, given the mixed patterns reported in previous research (McKenna and Sharma, 1995; Pérez-Edgar and Fox, 2003). This secondary analysis was intended to characterize the present task configuration rather than to provide confirmatory evidence of stable emotion-specific effects. The analyses were expected to inform task refinement and the design of larger hypothesis-driven studies.

Materials and methods

Participants

The study included six lower secondary school students recruited through convenience sampling from the researchers' personal networks and referrals. All participants were enrolled in Grades 6 to 8 at the time of data collection. The sample consisted of two female students (33.3%) and four male students (66.7%). One participant was in Grade 6 (16.7%), three were in Grade 7 (50.0%), and two were in Grade 8 (33.3%). Exact ages were not collected; therefore, participants' grade levels were used to indicate their educational stage.

Eligibility required current enrollment as a lower secondary school student, sufficient Vietnamese reading comprehension to understand the task instructions and emotion labels, written consent from a parent or legal guardian, verbal assent from the student, and willingness to complete the computerized task. No additional clinical or health-related exclusion criteria were applied, and no formal clinical, developmental, or visual screening was conducted. Participants who ordinarily used corrective glasses completed the task using their usual eyewear. Before the task began, the researcher explained the instructions and demonstrated the required keyboard responses. No participant was excluded or withdrew after enrollment.

Recruitment was conducted independently of any secondary school, and no school served as a recruitment or data-collection site. The task was administered individually in home-based settings; therefore, school administrative permission was not applicable.

Study Design

The study employed a within-subject pilot design using an Emotional Face–Word Stroop Task. All participants completed congruent and incongruent trials involving seven emotion categories. Reaction time (RT) was operationalized as the response latency from stimulus onset to keypress for trials in which a response was recorded, irrespective of response accuracy, and served as the primary performance measure. Response accuracy was analyzed separately as a secondary measure. The study was designed to describe task implementation and within-subject performance patterns in a pilot sample, rather than to estimate population-level effects.

Emotional Face–Word Stroop Task

Stimuli

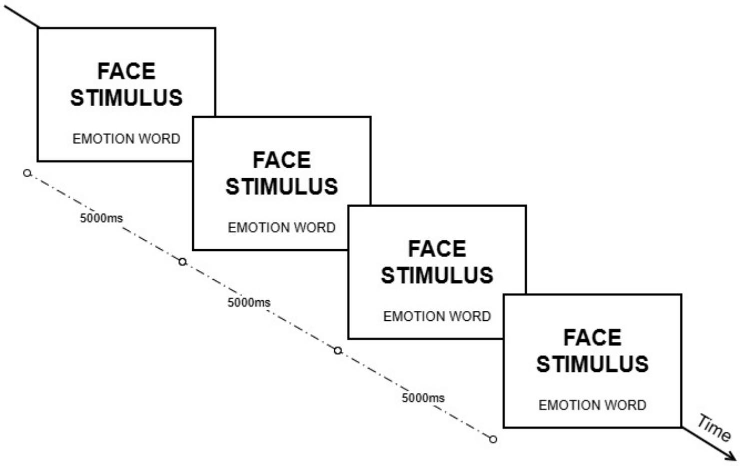
Facial stimuli were obtained from the Averaged Karolinska Directed Emotional Faces (AKDEF) database developed by Lundqvist and Litton (1998). The complete AKDEF set consists of 70 grayscale images depicting an averaged female face and an averaged male face, each displaying seven emotional expressions from five viewing angles. The expressions are afraid, angry, disgusted, happy, neutral, sad, and surprised, and the source images have a resolution of 562 × 762 pixels. The present task used all 70 images from the AKDEF set. These images were paired with Vietnamese emotion labels to construct 130 trials. The number of trials therefore refers to the number of face–label pairings, rather than the number of unique source images, and some images appeared in more than one pairing. Sad, angry, neutral, disgusted, surprised, and afraid expressions each appeared in 20 trials (15.4%), whereas happy expressions appeared in 10 trials (7.7%).

For the secondary exploratory analysis, trials were grouped as neutral (20 trials), positive (30 trials), and negative (80 trials). The positive group included happy and surprised trials, whereas the negative group included sad, angry, disgusted, and afraid trials. Surprise was provisionally classified as positive for this exploratory grouping. Because no happy-incongruent trials were presented, the positive cells were not compositionally equivalent across conditions. The positive congruent condition included both happy and surprised trials, whereas the positive incongruent condition included surprised trials only. This imbalance was explicitly considered when interpreting the exploratory stimulus-group analysis.

Although the task did not involve color–word naming, it followed a Stroop-like congruency framework in which participants responded to agreement or conflict between facial and verbal emotional information. Because the task required an explicit match–mismatch judgment, both the facial expression and the emotion label were relevant to the response. The manipulation therefore represented facial–verbal emotional congruency conflict rather than interference from a task-irrelevant dimension (Smolker et al., 2022).

Task Procedure

The task was programmed and administered using E-Prime version 1.1 (Psychology Software Tools, 2020). The English task instructions and emotion labels were translated into Vietnamese by bilingual members of the research team. Because the materials consisted of seven basic emotion terms and brief instructions, the translations were reviewed collaboratively for semantic clarity and age-appropriate comprehension. Given the limited and concrete nature of the materials, no formal back-translation or external linguistic validation was undertaken. Before the experimental trials began, the researcher explained the instructions orally, demonstrated the two response keys, allowed participants to ask questions, and confirmed that each participant understood the response procedure. No separate practice block was administered.



Each trial presented a facial expression together with a Vietnamese emotion label. Participants indicated whether the label matched the expression. The stimulus remained visible until a response was made or until the 5,000-ms response deadline was reached. The faces shown in the figure are schematic; the experimental stimuli were obtained from the AKDEF database (Lundqvist and Litton, 1998).

Figure 1. Example Trial Sequence in the Emotional Face—Word Stroop Task

No fixation cross was presented. On each trial, a facial image and a Vietnamese emotion label appeared simultaneously on the screen. Participants pressed the "1" key when the label matched the facial expression and the "2" key when the label did not match the expression. The facial stimulus and emotion label remained visible until the participant responded or the 5,000-ms response deadline was reached, and a keypress immediately terminated the stimulus display.

After each trial, feedback indicating a correct or incorrect response was displayed for approximately 1,000 ms. The next trial began immediately after the feedback display, with no separate intertrial interval. Because the task was brief, no scheduled breaks were provided. Across participants, total completion time ranged from 7.76 to 11.12 minutes, with a mean duration of 8.94 minutes. All participants received the same set of 130 trials, but trial order was randomized separately for each participant by E-Prime within a single mixed sequence. No additional counterbalancing or manually imposed restrictions on consecutive trial types were applied. Participants were instructed to respond as quickly and accurately as possible.

Apparatus and Software

The task was programmed and administered using E-Prime version 1.1 on the same Dell Precision 7550 laptop for all participants. The laptop had a 15.6-inch display, and responses were entered using its built-in keyboard. Using the same computer, display, keyboard, and software configuration standardized the administration conditions across participants. E-Prime recorded RT in milliseconds and automatically stored response accuracy.

Testing was conducted individually in a quiet and comfortable home-based setting. The researcher remained present throughout each session and made reasonable efforts to minimize noise, interruptions, and other external distractions. Timing precision was not independently benchmarked; however, the same hardware and software configuration was maintained across all testing sessions.

Data Collection and Measures

Data collection took place between May 2024 and October 2024. RT was operationalized as the interval between stimulus onset and the participant's keypress for trials with a recorded response, irrespective of response accuracy. Accuracy was coded separately as 1 for a correct response and 0 for an incorrect response when a response was recorded. The response window was limited to 5,000 ms. When no response was made before the deadline, E-Prime recorded a system-coded RT value of 0 ms, and no response accuracy was recorded. Two timeout trials occurred across the sample. During data preparation, the corresponding analysis values for these timeout trials were recoded as 99 and designated as user-missing in SPSS.

The task yielded 780 trial-level records across six participants. Each participant was presented with 130 trials, including 70 congruent trials (53.8%) and 60 incongruent trials (46.2%). All trial records were retained in the dataset; however, the two timeout trials were treated as missing for analyses requiring RT or accuracy. Participant-level condition means were therefore calculated from the available nonmissing observations within each condition. RT and accuracy were analyzed as separate performance indicators.

Data Analysis

Given the small pilot sample, the analyses were used primarily to characterize within-subject performance patterns, and the inferential findings were interpreted exploratorily rather than as population-level evidence. RT and accuracy were treated as separate performance indicators. RT analyses retained all recorded RTs irrespective of response accuracy; therefore, RTs from both correct and incorrect responses were included. The two timeout trials, for which no behavioral response was recorded, were designated as user-missing in SPSS and did not contribute to RT or accuracy calculations. No missing values were imputed. No RT trimming, winsorization, transformation, or distribution-based outlier exclusion was applied to the recorded RT values. Participant-level condition means were calculated from the available nonmissing RT observations within each condition.

For the primary within-subject comparison, participant-level mean RTs for congruent and incongruent trials were compared using a paired-samples t test. Given the sample size, distributional assumptions and inferential estimates were interpreted cautiously. The mean difference, its 95% confidence interval, and Cohen's d_z were reported to characterize the magnitude of the within-subject condition difference.

For the secondary exploratory analysis, a 2 (condition: congruent, incongruent) × 3 (broader stimulus group: neutral, positive, negative) repeated-measures analysis of variance was conducted using participant-level mean RTs. Because of the small sample size and unequal stimulus composition, this analysis was used to characterize patterns in the present task configuration rather than to provide confirmatory evidence of emotion-specific effects. Partial eta squared (η_p^2) was reported as an effect-size estimate. Sphericity was assessed using Mauchly's test, and Greenhouse–Geisser corrections were planned if the assumption was violated. Statistical analyses were conducted using IBM SPSS Statistics Version 26.

Results

Data Completeness and Trial Retention

All six participants completed the full task protocol, yielding 780 trial-level records. Two trials (0.3%) reached the 5,000-ms response deadline without a behavioral response. These timeout trials remained in the dataset but were treated as missing for analyses requiring RT or accuracy. Accordingly, 778 trials contained recorded behavioral responses. RTs from both correct and incorrect responses were retained in the RT analyses, whereas accuracy was analyzed separately.

Descriptive Statistics of Emotional Stroop Performance

Descriptive statistics for RTs across conditions and emotion categories are presented in Table 1. Mean RTs across individual emotion categories ranged from approximately 1,201 to 1,948 ms. Although the overall condition-level mean was higher for incongruent trials, patterns across individual emotion categories varied by condition. Therefore, the individual emotion results are presented descriptively, without inferential comparisons at that level.

Table 1. *Descriptive Statistics of Reaction Times (ms) Across Conditions and Emotion Groups*

Condition	Emotion Group	Emotion Name	Minimum	Maximum	Mean	Standard Deviation
Congruent	Positive	Happy	636	3381	1201	453
		Surprised	784	2578	1352	368
	Negative	Afraid	843	3997	1838	696
		Angry	607	3233	1555	594
		Disgusted	723	4739	1407	646
		Sad	626	4029	1768	639
	Neutral	Neutral	737	4187	1689	677
Incongruent	Positive	Happy	–	–	–	–
		Surprised	901	4250	1948	713
	Negative	Afraid	658	4715	1829	718
		Angry	735	3113	1620	545
		Disgusted	796	4992	1494	713
		Sad	862	3193	1772	604
	Neutral	Neutral	814	4970	1653	799

Note: Descriptive statistics are based on trial-level RT observations and are presented to characterize the observed RT distributions within the task. Timeout trials were treated as missing and were not included in these calculations.

Overall accuracy was 81.3%. Descriptive accuracy by condition is presented in Table 2. Mean accuracy was 86.19% (*SD* = 4.11%) for congruent trials and 75.56% (*SD* = 12.46%) for incongruent trials.

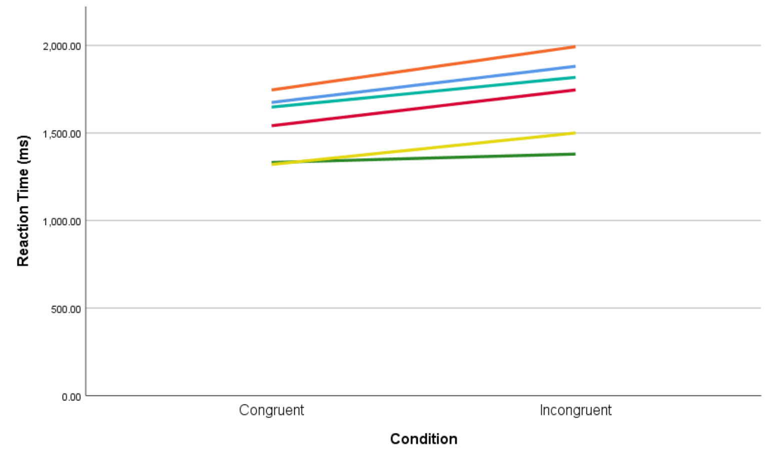
Table 2. *Overall Accuracy*

Condition	Mean Accuracy (%)	Standard Deviation (%)
Congruent	86.19	4.11
Incongruent	75.56	12.46

Note: Accuracy was calculated from trials with a recorded response; timeout trials were treated as missing

Condition Effect (Congruent vs. Incongruent)

The primary within-subject comparison examined whether mean RTs differed between congruent and incongruent trials. Participant-level mean RTs were longer for incongruent trials (*M* = 1,719.50 ms, *SD* = 234.40) than for congruent trials (*M* = 1,543.63 ms, *SD* = 180.87). An exploratory paired-samples *t* test indicated a condition difference, $t_{(5)} = 6.32$, $p = .001$, Cohen's $d_z = 2.58$. The mean difference was 175.87 ms, 95% CI [104.28, 247.46]. Within this six-participant pilot sample, the observed mean RT difference was in the expected direction; however, the inferential result is reported for exploratory purposes and should not be interpreted as providing a stable population-level effect estimate. Participant-level mean RTs for the two conditions are illustrated in Figure 2.



Each line represents an individual participant. Reaction times are shown as participant-level means

Figure 2. Reaction Times for Congruent and Incongruent Conditions

Exploratory Condition-by-Stimulus-Group Patterns

As a secondary exploratory analysis, a 2 (condition: congruent, incongruent) × 3 (broader stimulus group: neutral, positive, negative) repeated-measures analysis of variance was conducted using participant-level mean RTs. Because the task contained unequal numbers of trials across emotion categories and the positive cells differed in emotional composition across conditions, the analysis was used to characterize the observed pattern in the present task configuration rather than to test confirmatory valence-specific effects.

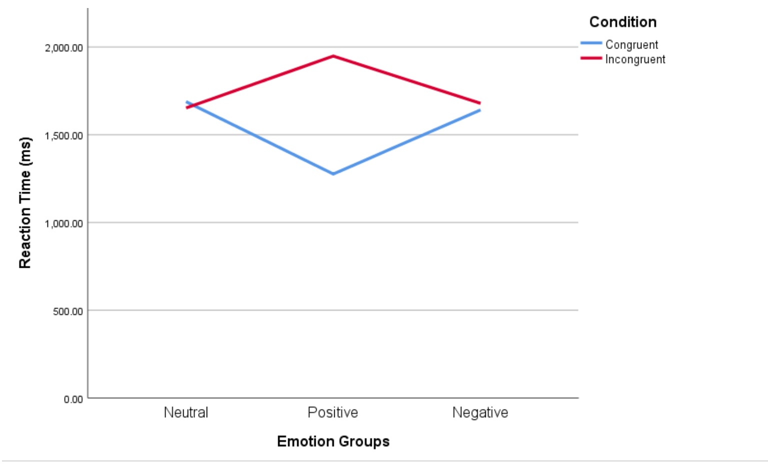
Table 3. *Mean Reaction Times (ms) Across Conditions and Broader Stimulus Groups*

Emotion Group	Congruent (Mean ± Standard Deviation)	Incongruent (Mean ± Standard Deviation)
Neutral	1688.68 ± 237.57	1652.57 ± 377.27
Positive	1276.33 ± 171.91	1948.30 ± 171.48
Negative	1641.64 ± 219.27	1678.92 ± 275.11

Note: Values represent participant-level means. The positive congruent condition included happy and surprised trials, whereas the positive incongruent condition included surprised trials only because no happy-incongruent trials were presented. The cells therefore differed in emotional composition, and comparisons involving the positive group should be interpreted descriptively.

Mauchly's tests were not statistically significant for either the broader stimulus-group effect or the Condition × Stimulus Group interaction (p = .457 and p = .696, respectively). Given the small pilot sample, however, these tests provided limited evidence regarding the sphericity assumption. For consistency with the original exploratory analysis, the uncorrected degrees of freedom are retained and reported descriptively.

The analysis yielded a main effect of condition, $F_{(1, 5)} = 42.99$, $p = .001$, $\eta_p^2 = .90$. The main effect of broader stimulus group was not statistically significant, $F_{(2, 10)} = 0.20$, $p = .83$, $\eta_p^2 = .04$. A Condition × Stimulus Group interaction was also observed, $F_{(2, 10)} = 22.74$, $p < .001$, $\eta_p^2 = .82$. Because the positive congruent and incongruent cells did not contain the same emotion categories, the interaction is reported as a task-specific exploratory pattern rather than evidence of valence-specific interference. Given the small pilot sample and the non-equivalent emotional composition of the positive cells, these ANOVA results are reported solely as exploratory summaries of the observed task configuration. The associated p values and effect-size estimates should not be interpreted as stable estimates of population-level effects or as confirmatory evidence of valence-specific interference.



Note: The positive congruent and incongruent cells differed in emotional composition because happy trials were present only in the congruent condition. The figure is therefore intended to illustrate the observed task pattern rather than a confirmed valence-related effect.

Figure 3. Descriptive Reaction-Time Patterns Across Broader Stimulus Groups and Conditions.

Accuracy and Error Patterns

Overall accuracy was adequate rather than near either extreme of the response scale. Mean accuracy was descriptively lower for incongruent trials than for congruent trials. Descriptively, the combination of longer RTs and lower accuracy in the incongruent condition was not consistent with the conventional pattern expected under a speed–accuracy trade-off. Because accuracy was not subjected to inferential testing, this interpretation remains descriptive.

Discussion

The present pilot study examined attentional processing in Vietnamese lower secondary school students using an Emotional Face–Word Stroop Task. Accordingly, the present paradigm is best understood as a Stroop-like facial–verbal congruency task rather than a conventional emotional color–word Stroop task in which emotional word meaning is task-irrelevant to the color-naming response. The primary focus was the within-subject difference in reaction time between congruent and incongruent trials, whereas patterns across broader stimulus groups were examined exploratorily. Overall, the results showed longer reaction times for incongruent trials, together with adequate descriptive accuracy across the sample. The exploratory analysis also indicated that condition-related reaction-time patterns differed across broader stimulus groups, although these differences require cautious interpretation because of the unequal stimulus composition.

The longer reaction times observed for incongruent trials are consistent with previous emotional face–word Stroop findings in adolescent samples (Smolker et al., 2022). In the present task, the incongruent condition was defined by a mismatch between facial and verbal emotional information rather than by conventional color–word conflict. The observed pattern suggests that evaluating mismatched emotional cues imposed additional processing demands, potentially engaging conflict-resolution and cognitive-control processes (Jamieson et al., 2023; Smolker et al., 2022). Given the pilot design, the result is best interpreted as a consistent within-sample condition pattern rather than as a stable estimate of a population-level effect. Nevertheless, it provides a useful basis for refining and evaluating this Vietnamese-language task in a larger sample.

Accuracy findings provided additional context for interpreting the reaction-time results. Mean accuracy was descriptively lower for incongruent trials than for congruent trials, while overall performance remained adequate. Descriptively, the combination of longer reaction times and lower accuracy in the incongruent condition was not consistent with the conventional pattern expected under a speed–accuracy trade-off. Because accuracy was not subjected to inferential testing, this interpretation remains descriptive.

The exploratory analysis further indicated that condition-related reaction-time differences varied across the broader stimulus groups. However, this pattern cannot be attributed confidently to emotional valence because the positive cells were not compositionally equivalent across conditions. The positive congruent condition included happy and surprised trials, whereas the positive incongruent condition included surprised trials only, and the numbers of neutral, positive, and negative trials were also unequal. The observed interaction may therefore reflect differences in stimulus composition, trial frequency, or their combination rather than a stable valence-specific effect. This cautious interpretation is consistent with evidence that emotional Stroop performance can vary as a function of valence and arousal (Imbir, Duda-Goławska, et al., 2021), while emotional face–word Stroop effects can also vary with distractor valence and the relative frequency of incongruent trials (Smolker et al., 2022). Accordingly, the present interaction is most appropriately viewed as a task-specific exploratory pattern that can inform the development of a more balanced stimulus design.

The study also has relevance for developmental and educational research. In educational settings, attentional bias and emotional interference have been associated with academic stressors, self-regulatory processes, and classroom context (Scrimin et al., 2016; Scrimin et al., 2019). Emotion regulation is also longitudinally related to school-related well-being during adolescence (Beaumont et al., 2023). Although these educational outcomes were not directly assessed, the present task provides a behavioral paradigm that could be incorporated into future studies examining how school-aged participants manage competing emotional information. The Vietnamese context also adds evidence from a comparatively underrepresented population in the emotional Stroop literature. Because emotion recognition and labeling may vary across developmental and cultural contexts, and because task translation can affect cross-language implementation, culturally and linguistically adapted paradigms are important for future comparative research (Lunn and Chen, 2022; Möller et al., 2022). The current findings therefore contribute preliminary procedural and behavioral information rather than direct evidence of cultural differences.

Several limitations should be considered. The pilot sample was small and recruited through convenience sampling, which limits the precision of effect-size estimates and the extent to which the observed patterns can be generalized. Exact ages were not collected, which also restricted developmental interpretation beyond grade level. In addition, the unequal stimulus distribution, particularly the absence of happy-incongruent trials, limits conclusions about valence-specific or emotion-specific effects. Testing was conducted in home-based settings rather than a laboratory; however, the same laptop, software configuration, researcher-administered procedure, and efforts to minimize distractions were used across all sessions. No separate practice block was administered, so early-trial performance may have reflected task-familiarization effects, particularly given the participants' developmental stage. The reaction-time analysis retained both correct and incorrect responses when an RT was recorded because the outcome was defined as overall response timing. Timeout trials without a recorded behavioral response were treated as missing. Future studies could prespecify sensitivity analyses restricted to correct-response trials to evaluate the robustness of this analytic choice. Finally, the translated materials were reviewed collaboratively but did not undergo formal linguistic validation, and the hardware-software timing configuration was not independently benchmarked. These limitations should be considered when interpreting the magnitude and robustness of the observed within-sample condition pattern and identify important priorities for task refinement and future confirmatory investigation.

Future research should examine emotional face–word Stroop performance in larger and more diverse samples of school-aged participants. A balanced design should include equivalent emotion categories and trial frequencies across congruent and incongruent conditions. Larger studies should also use prespecified reaction-time preprocessing rules, standardized task instructions, and controlled testing procedures, together with preregistered primary and exploratory analyses. Such studies could examine whether emotional face–word Stroop performance varies by developmental stage or is associated with academic stress, anxiety, emotion regulation, and student well-being. Overall, the present pilot provides a preliminary behavioral and procedural basis for larger hypothesis-driven research on emotional conflict and attentional processing in developmental, educational, and cross-cultural contexts.

Conclusion

The present pilot study examined reaction-time patterns in six Vietnamese lower secondary school students using an Emotional Face–Word Stroop Task that combined facial expressions with Vietnamese emotion labels. Mean reaction times were longer for incongruent than congruent trials, showing an observed within-sample condition pattern under the present task configuration. Accuracy was descriptively lower for incongruent trials but remained adequate overall. Descriptively, the combination of longer reaction times and lower accuracy in the incongruent condition was not consistent with the conventional pattern expected under a speed–accuracy trade-off.

The exploratory variation across broader stimulus groups should be interpreted as task-specific because the emotion categories and trial frequencies were not equivalent across conditions. The findings therefore provide preliminary behavioral and procedural information for refining a Vietnamese-language emotional face–word Stroop paradigm rather than population-level evidence of stable valence-specific effects. Future studies with larger samples, balanced stimuli, prespecified reaction-time preprocessing procedures, and more controlled testing conditions are needed to evaluate the stability of the observed pattern and extend its use in developmental, educational, and cross-cultural research.

Acknowledgements

The authors would like to thank the participating students and their parents or legal guardians for their time and cooperation. The authors also gratefully acknowledge Ho Chi Minh City University of Education for supporting this research. The first author was affiliated with Ho Chi Minh City University of Education during the data collection period.

Funding

This study was supported by Ho Chi Minh City University of Education as part of a university-level research project (Project code: CS.2022.19.24TĐ). No external funding was received.

Conflict of Interests

The authors declare that there is no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon request. Restrictions apply due to privacy considerations related to the small sample of minor participants.

Institutional Review Board Statement

At the time the protocol was reviewed, Ho Chi Minh City University of Education did not operate a separately designated institutional review board or research ethics board. Before data collection, the study protocol underwent prospective institutional review through the university-level research project approval process. A review panel convened by the university and the relevant academic unit evaluated the scientific protocol and the ethical safeguards for research involving minor participants. Following this review, the project was approved on May 10, 2022, under university-level research project code CS.2022.19.24TĐ.

The study procedures were guided by the ethical principles of the Declaration of Helsinki (World Medical Association, 2013) and the APA Ethics Code (American Psychological Association, 2017). Written informed consent was obtained from a parent or legal guardian of each participant, and verbal assent was obtained from each participating student before data collection. Participants and their parents or legal guardians were informed about the study procedures, the voluntary nature of participation, confidentiality protections, and the right to withdraw at any time without negative consequences. Recruitment and data collection were conducted independently of any secondary school, and no school records, facilities, or personnel were involved; therefore, permission from a secondary school administration was not applicable.

Author Contributions

Conceptualization, V.L.T.C.; methodology, V.L.T.C. and G.P.T.T.; software, V.L.T.C.; formal analysis, V.L.T.C. and G.P.T.T.; writing—original draft preparation, G.P.T.T.; writing—review and editing, V.L.T.C. All authors have read and agreed to the published version of the manuscript.

References

American Psychological Association. (2017). *Ethical Principles of Psychologists and Code of Conduct*. American Psychological Association. https://www.apa.org/ethics/code

Banich, M. T., Smolker, H. R., Snyder, H. R., Lewis-Peacock, J. A., Godinez, D. A., Wager, T. D., & Hankin, B. L. (2019). Turning Down the Heat: Neural Mechanisms of Cognitive Control for Inhibiting Task-Irrelevant Emotional Information During Adolescence. *Neuropsychologia*, *125*, 93–108. https://doi.org/10.1016/j.neuropsychologia.2018.12.006

Beaumont, J., Putwain, D. W., Gallard, D., Malone, E., Marsh, H. W., & Pekrun, R. (2023). Students' Emotion Regulation and School-Related Well-Being: Longitudinal Models Juxtaposing Between- And Within-Person Perspectives. *Journal of Educational Psychology*, *115*(7), 932–950. https://doi.org/10.1037/edu0000800

Hsieh, C. W., & Sharma, D. (2019). Priming Emotional Salience Reveals the Role of Episodic Memory and Task Conflict in the Non-color Word Stroop Task. *Frontiers in Psychology*, *Volume 10*, 1826. https://doi.org/10.3389/fpsyg.2019.01826

Imbir, K. K., Duda-Goławska, J., Pastwa, M., Jankowska, M., & Żygierewicz, J. (2021). Event-Related Potential Correlates of Valence, Arousal, and Subjective Significance in Processing of an Emotional Stroop Task. *Frontiers in Human Neuroscience*, *Volume 15*, 617861. https://doi.org/10.3389/fnhum.2021.617861

Imbir, K. K., Pastwa, M., Duda-Goławska, J., Sobieszek, A., Jankowska, M., Modzelewska, A., Wielgopolan, A., & Żygierewicz, J. (2021). Electrophysiological Correlates of Interference Control in the Modified Emotional Stroop Task With Emotional Stimuli Differing in Valence, Arousal, and Subjective Significance. *PLOS ONE*, *16*(10), e0258177. https://doi.org/10.1371/journal.pone.0258177

Ishikawa, K., Oyama, T., & Okubo, M. (2021). The Malfunction of Domain-Specific Attentional Process in Social Anxiety: Attentional Process of Social and Non-Social Stimuli. *Cognition and Emotion*, *35*(6), 1163–1174. https://doi.org/10.1080/02699931.2021.1935217

Jamieson, G. A., Page, J., Evans, I. D., & Hamlin, A. (2023). Conflict and Control in Cortical Responses to Inconsistent Emotional Signals in a Face-Word Stroop. *Frontiers in Human Neuroscience*, *Volume 17*, 955171. https://doi.org/10.3389/fnhum.2023.955171

Joyal, M., Wensing, T., Levasseur-Moreau, J., Leblond, J., Sack, A. T., & Fecteau, S. (2019). Characterizing Emotional Stroop Interference in Posttraumatic Stress Disorder, Major Depression and Anxiety Disorders: A Systematic Review and Meta-Analysis. *PLOS ONE*, *14*(4), e0214998. https://doi.org/10.1371/journal.pone.0214998

Kuehne, M., Schmidt, K., Heinze, H.-J., & Zaehle, T. (2019). Modulation of Emotional Conflict Processing by High-Definition Transcranial Direct Current Stimulation (HD-TDCS). *Frontiers in Behavioral Neuroscience*, *Volume 13*, 224. https://doi.org/10.3389/fnbeh.2019.00224

Lundqvist, D., & Litton, J. E. (1998). *The Averaged Karolinska Directed Emotional Faces - AKDEF [CD-ROM]*. Department of Clinical Neuroscience, Psychology Section, Karolinska Institutet.

Lunn, J., & Chen, J.-Y. (2022). Chronotype and Time of Day Effects on Verbal and Facial Emotional Stroop Task Performance in Adolescents. *Chronobiology International*, *39*(3), 323–332. https://doi.org/10.1080/07420528.2021.1998102

MacLeod, C. M. (1991). Half a Century of Research on the Stroop Effect: An Integrative Review. *Psychological Bulletin*, *109*(2), 163–203. https://doi.org/10.1037/0033-2909.109.2.163

McKenna, F. P. (1986). Effects of Unattended Emotional Stimuli on Color-Naming Performance. *Current Psychological Research & Reviews*, *5*(1), 3–9. https://doi.org/10.1007/BF02686591

McKenna, F. P., & Sharma, D. (1995). Intrusive Cognitions: An Investigation of the Emotional Stroop Task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(6), 1595–1607. https://doi.org/10.1037/0278-7393.21.6.1595

McKenna, F. P., & Sharma, D. (2004). Reversing the Emotional Stroop Effect Reveals That It Is Not What It Seems: The Role of Fast and Slow Components. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(2), 382–392. https://doi.org/10.1037/0278-7393.30.2.382

Möller, C., Bull, R., & Aschersleben, G. (2022). Culture Shapes Preschoolers' Emotion Recognition but Not Emotion Comprehension: A Cross-Cultural Study in Germany and Singapore. *Journal of Cultural Cognitive Science*, *6*(1), 9–25. https://doi.org/10.1007/s41809-021-00093-6

Müller, V. I., Cieslik, E. C., Ficco, L., Tyralla, S., Sepehry, A. A., Aziz-Safaie, T., Feng, C., Eickhoff, S. B., & Langner, R. (2025). Not All Stroop-Type Tasks Are Alike: Assessing the Impact of Stimulus Material, Task Design, and Cognitive Demand via Meta-Analyses Across Neuroimaging Studies. *Neuropsychology Review*, *35*(3), 449–482. https://doi.org/10.1007/s11065-024-09647-1

Parris, B. A., Hasshim, N., Wadsley, M., Augustinova, M., & Ferrand, L. (2022). The Loci of Stroop Effects: A Critical Review of Methods and Evidence for Levels of Processing Contributing to Color-Word Stroop Effects and the Implications for the Loci of Attentional Selection. *Psychological Research*, *86*(4), 1029–1053. https://doi.org/10.1007/s00426-021-01554-x

Pérez-Edgar, K., & Fox, N. A. (2003). Individual Differences in Children's Performance During an Emotional Stroop Task: A Behavioral and Electrophysiological Study. *Brain and Cognition*, *52*(1), 33–51. https://doi.org/10.1016/S0278-2626(03)00007-1

Psychology Software Tools. (2020). *E-Prime (Version 1.1)*. [Computer Software]. https://support.pstnet.com/

Rakesh, D., Lee, P. A., Gaikwad, A., & McLaughlin, K. A. (2025). Annual Research Review: Associations of Socioeconomic Status With Cognitive Function, Language Ability, and Academic Achievement in Youth: A Systematic Review of Mechanisms and Protective Factors. *Journal of Child Psychology and Psychiatry*, *66*(4), 417–439. https://doi.org/10.1111/jcpp.14082

Scrimin, S., Moscardino, U., Altoè, G., & Mason, L. (2016). Effects of Perceived School Well-Being and Negative Emotionality on Students' Attentional Bias for Academic Stressors. *British Journal of Educational Psychology*, *86*(2), 278–295. https://doi.org/10.1111/bjep.12104

Scrimin, S., Moscardino, U., & Mason, L. (2019). First-Graders' Allocation of Attentional Resources in an Emotional Stroop Task: The Role of Heart Period Variability and Classroom Climate. *British Journal of Educational Psychology*, *89*(1), 146–164. https://doi.org/10.1111/bjep.12228

Shih-Tseng, T. H., & Tran-Chi, V.-L. (2019). The Prospect of Applying Cognitive Neuroscience in Educational Research in Vietnam: A Lesson From Taiwan. *Ho Chi Minh City University of Education Journal of Science*, *16*(1), 30–38. https://doi.org/10.54607/hcmue.js.16.1.2332(2019)

Smolker, H. R., Wang, K., Luciana, M., Bjork, J. M., Gonzalez, R., Barch, D. M., McGlade, E. C., Kaiser, R. H., Friedman, N. P., Hewitt, J. K., & Banich, M. T. (2022). The Emotional Word-Emotional Face Stroop Task in the ABCD Study: Psychometric Validation and Associations With Measures of Cognition and Psychopathology. *Developmental Cognitive Neuroscience*, *53*, 101054. https://doi.org/10.1016/j.dcn.2021.101054

Straub, E. R., Schmidts, C., Kunde, W., Zhang, J., Kiesel, A., & Dignath, D. (2022). Limitations of Cognitive Control on Emotional Distraction – Congruency in the Color Stroop Task Does Not Modulate the Emotional Stroop Effect. *Cognitive, Affective, & Behavioral Neuroscience*, *22*(1), 21–41. https://doi.org/10.3758/s13415-021-00935-4

Stroop, J. R. (1935). Studies of Interference in Serial Verbal Reactions. *Journal of Experimental Psychology*, *18*(6), 643–662. https://doi.org/10.1037/h0054651

Tsenkova, N., Bahn, D., Kauschke, C., & Schwarzer, G. (2025). Developmental Differences in Perceiving Arousal and Valence From Dynamically Unfolding Emotional Expressions. *PLOS ONE*, *20*(8), e0329554. https://doi.org/10.1371/journal.pone.0329554

Wentura, D., Rothermund, K., & Bak, P. (2000). Automatic Vigilance: The Attention-Grabbing Power of Approach and Avoidance-Related Social Information. *Journal of Personality and Social Psychology*, *78*(6), 1024–1037. https://doi.org/10.1037/0022-3514.78.6.1024

Wielgopolan, A., Pastwa, M., Żuk, D., & Imbir, K. K. (2025). Emotional Stroop Test in Online and Offline Experimental Procedures: The Comparison of Two Different Modes of Conducting Studies. *Language and Cognition*, *17*, e31, Article e31. https://doi.org/10.1017/langcog.2024.77

Williams, J. M. G., Mathews, A., & MacLeod, C. (1996). The Emotional Stroop Task and Psychopathology. *Psychological Bulletin*, *120*(1), 3–24. https://doi.org/10.1037/0033-2909.120.1.3

World Medical Association. (2013). World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects. *JAMA*, *310*(20), 2191–2194. https://doi.org/10.1001/jama.2013.281053