Psychiatry Research 229 (2015) 960–967

Contents lists available at ScienceDirect

Psychiatry Research

journal homepage: www.elsevier.com/locate/psychres

Valence-based Word-Face Stroop task reveals differential emotional interference in patients with major depression

Zeynep Başgöze [a], Ali Saffet Gönül [b], Bora Baskak [c], Didem Gökçay [a,*]

[a] Middle East Technical University, Informatics Institute, Ankara, Turkey
[b] Ege University Medical School, Psychiatry, Izmir, Turkey
[c] Ankara University Medical School, Psychiatry, Ankara, Turkey

ARTICLE INFO

Article history:
Received 21 May 2014
Received in revised form
20 May 2015
Accepted 23 May 2015
Available online 29 June 2015

Keywords:
Major depressive disorder
Emotional interference
Word-Face Stroop
Negative bias
Analytical rumination

ABSTRACT

Word-Face Stroop task creates emotional conflict between affective words and affective faces. In this task, healthy participants consistently slow down while responding to incongruent cases. Such interference related slowdown is associated with recruitment of inhibitory processes to eliminate task-irrelevant information. We created a valence-based Word-Face Stroop task, in which participants were asked to indicate whether the words in the foreground are positive, negative or neutral. Healthy participants were faster and more accurate than un-medicated patients with major depression disorder (MDD). In addition, a significant congruence by group interaction is observed: healthy participants slowed down for incongruent cases, but MDD patients did not. Furthermore, for the negative words, healthy individuals made more errors while responding to incongruent cases but MDD patients made the lowest number of errors for this category. The emotional percepts of the patients were intact, because correct response rates in word valence judgments for positive/negative words, and reaction times for happy/sad faces had similar patterns with those of controls. These findings are supported by the analytical rumination interpretation of depression: patients lose speed/accuracy in laboratory tasks due to processing load spent during continuous rumination. However, for tasks in line with their preoccupation, continual practice makes the patients more vigilant and adept.

1. Introduction

Among several deficits in major depression disorder (MDD), problems in conflict resolution are reported frequently (Paelecke-Habermann et al., 2005; Holmes and Pizzagalli, 2007). Mood-congruent processing biases in MDD patients perpetuate depressive symptoms by influencing the salience of social stimuli (Clasen et al., 2013). An enhanced sensitivity to negative situations makes patients more inclined to misinterpret emotionally conflicting situations (Zetsche and Joormann, 2011), resulting in a dysfunction in conscious perceptions and social interactions (Victor et al., 2010), which, in turn, increases the intensity of their depressive symptoms (Gotlib et al., 2004).

The classical Stroop task measures cognitive aspects of conflict resolution by focusing on the dominant effect of word meaning over word color. While major depression patients are consistently reported to be slower than healthy controls in the Stroop task, they still exhibit the expected inhibitory slowdown in reaction times for incongruent cases (Markela-Lerenc et al., 2006; Gohier et al., 2009). On the other hand, an emotional load that is task-irrelevant may also create conflicts. Cognition must be protected from interference by irrelevant emotional stimuli for effective mental functioning (Etkin et al., 2006). There exists an array of emotional Stroop paradigms, for which a general response characteristic emerges in the depressed patient populations. A meta-analysis of emotional Stroop tasks reveal that depressed populations exhibit longer response latencies compared to healthy controls (Epp et al., 2012). Furthermore, a qualitative difference between clinically depressed, dysphoric and mood induced groups is observed: 'the general emotional bias is stronger among clinically depressed individuals and there is some content-specificity favoring negative stimuli' (Epp et al., 2012). The emotional bias in the depression patients toward negative stimuli is usually associated with sad mood, while task design parameters such as stimulus duration are reported to be effective for observing differential behavior between patients and controls (Gotlib et al., 2004).

In order to investigate emotional aspects of conflict resolution, it is imperative to consider not only congruence, but also emotional conditions (i.e. happy/sad or positive/negative). Word-Face Stroop tasks that generate interference between affective words and affective faces allow for investigation of both of these factors simultaneously. Compared to the classical Stroop task, the Word-Face Stroop task is less investigated. Nevertheless, in healthy and depressed populations, interference effects are observed for conflicting emotional

* Correspondence to: METU Üniversiteler Mah., Dumlupınar Blv. No. 1, 06800 Çankaya Ankara, Turkey. Fax: +90 312 210 3745.
E-mail address: dgokcay@metu.edu.tr (D. Gökçay).

http://dx.doi.org/10.1016/j.psychres.2015.05.099
0165-1781/

content in words and pictures consistently [1] (Stenberg et al., 1998; Haas et al., 2006; Etkin et al., 2006; Egner et al., 2008, Başgöze, 2008; Zhu et al., 2010; Hu et al., 2012; Strand et al., 2013, Chechko et al., 2013). In a Word-Face Stroop task, conflict occurs between emotional words overlaid on emotional facial expressions. In the congruent condition, the word and the face bear the same emotional category but in the incongruent condition, the emotional categories of the word and the face differ. In some variants, neutral words are used along with neutral faces in a third condition.

Current applications of the Word-Face Stroop task can be summarized as follows. When the words are chosen as targets, the participants are asked to evaluate the emotion of the word on the foreground (Stenberg et al., 1998; Haas et al., 2006; Başgöze, 2008; Zhu et al., 2010; Strand et al., 2013). When the faces are chosen as targets, the participants are asked to evaluate the emotion of the face on the background (Etkin et al., 2006; Egner et al., 2008; Zhu et al., 2010; Hu et al., 2012; Strand et al., 2013; Chechko et al., 2013). The evaluation may be based on categorical (e.g. happy/sad; Etkin et al., 2006; Egner et al., 2008; Zhu et al., 2010; Strand et al., 2013; Chechko et al., 2013) or valence related (e.g. positive/negative; Stenberg et al., 1998; Haas et al., 2006; Basgoze, 2008; Hu et al., 2012; Strand et al., 2013) judgments. So far, the Word-Face Stroop task is administered in several languages such as Swedish (Stenberg et al., 1998), English (Haas et al., 2006; Etkin et al., 2006; Egner et al., 2008), Turkish (Başgöze, 2008), Chinese (Zhu et al., 2010; Hu et al., 2012), Norwegian (Strand et al., 2013) and German (Chechko et al., 2013). Word stimuli may consist of emotion words (Etkin et al., 2006; Egner et al., 2008; Zhu et al., 2010; Chechko et al., 2013) or positive/negative nouns that are previously rated (Stenberg et al., 1998; Haas et al., 2006; Başgöze, 2008; Hu et al., 2012; Strand et al., 2013). Face stimuli are chosen from databases (Ekman and Friesen, 1976; Ekman et al., 2002; Gur et al., 2002; Minear and Park, 2004; Bai et al., 2005; Tottenham et al., 2009) with expressions of 6 basic emotions.

Findings from these Word-Face Stroop experiments indicate that healthy controls' response time for the incongruent condition is longer than the congruent condition. This is expected because interference effects brought by the competing emotions of the word versus the face increase processing load (Stenberg et al., 1998; Etkin et al., 2006; Haas et al., 2006; Başgöze, 2008; Egner et al., 2008; Zhu et al., 2010; Hu et al., 2012; Chechko et al., 2013; Strand et al., 2013). Findings regarding the incongruent versus congruent conditions differ with respect to the diagnostic category of the MDD groups. For instance, Strand et al. (2013) and Chechko et al. (2013) observed the existence of interference related slowdown in remitted MDD patients and a group of medicated MDD patients respectively. However, Başgöze (2008) and Hu et al. (2012) did not observe this interference related slowdown in un-medicated MDD patients. Unfortunately, none of these studies reported significant interaction in reaction times for the congruence and group effect.[2] Most studies reported that the accuracy of the subjects for the congruent condition is higher (Etkin et al., 2006; Zhu et al., 2010; Chechko et al., 2013; Strand et al., 2013). Strand et al. (2013) and Chechko et al. (2013) indicated that although less accurate than healthy subjects, MDD patients also held the same response pattern: being less accurate when emotional interference is encountered. Differences between emotional conditions such as positive versus negative or happy versus sad are less addressed. Etkin et al. (2008) did not find any differences in this regard in healthy subjects, but Stenberg et al. (1998) and Hu et al. (2012) observed positive word advantage (i.e. reaction times for positive words were shorter) in the healthy group. Hu et al. (2012) indicated that for the un-medicated patient group, this advantage was absent. The fact that these tasks are performed on several different populations such as remitted patients (Strand et al., 2013), patients with minor depression (Başgöze, 2008; Chechko et al., 2013), medicated patients with major depression (Chechko et al., 2013) or un-medicated patients (Başgöze, 2008; Hu et al. 2012), with different task parameters is a prohibiting factor for generalizing the results across clinical groups.

In the literature, attention biases in the MDD patients for negative stimuli are reported frequently (Joormann and Gotlib, 2007). Depressed people have difficulty in disengaging their attention from negative information and respond slowly when the distracter has negative content (Siegle et al., 2002). However, depressed people might also exhibit enhanced performance toward negative stimuli, in terms of either reaction times or accuracy when conflict monitoring (Erickson et al., 2005; Karparova et al., 2007) or memory (Fritzsche et al., 2010) is a prominent part of the task. These behavioral differences have a physiological underlay such that individuals with major depressive disorder show increased neural activity in response to sad, but diminished neural activity in response to happy stimuli in emotion-related brain circuits (Leppänen, 2006). This pattern is reverse for healthy controls (Suslow et al., 2010). Furthermore, this pathophysiology is present even for subliminal stimuli (i.e. briefly presented and masked happy faces and sad faces) (Suslow et al., 2010; Victor et al., 2010). Interestingly, after medication, the physiological profile of the MDD patients is reversed, so that there is more brain activity for happy compared to sad faces in limbic areas such as amygdala and ventral striatum (Victor et al., 2010).

From these findings we may draw two main hypotheses for Word-Face Stroop tasks: (1) The emotional interference effect (i.e. slowdown for emotionally competing stimuli) in the healthy population will be diminished in an un-medicated MDD population. (2) The positive word advantage seen in the healthy population will be observed in the MDD population as well, albeit due to a totally different cause. The rationale for these hypotheses is rooted in the same cause: the mood-congruent bias in the patients. If the patients cannot disengage from negative stimuli, then the distracting stimuli in the incongruent situations will not work, because in the incongruent condition, at least one of the simultaneously presented stimuli is negative. On another front, the negative bias slows down the patients' response for negative targets. Hence positive word advantage will be observed in the patients due to the slowness induced by the negative words.

In this study, our main motivation was to design and administer a Word-Face Stroop task such that the differential behavior of the MDD patients in emotional conflict resolution is pronounced. Earlier studies did not report statistically significant interaction between the subject groups. However, if the task design promotes the differences in the patient group, the interaction in congruency and subject groups might become significant. For this purpose, task parameters for which MDD patients are sensitive must be manipulated carefully. Three parameters seem to be important for observing differences in depression: (1) Stimulus duration. (2) Emotional category of the evaluated feature. (3) Modality of the stimulus that creates interference. As reported by Joorman and Gotlib (2007) and Gotlib et al. (2004), investigators have found attention biases in depression consistently when tasks had relatively long stimulus durations (i.e. 1 s or more) but for stimuli with shorter durations, the mood-congruent attention biases were variable and inconsistent. On the other hand, due to survival-related priorities, people keep tracking different perceptual channels for significant events and make crude evaluative emotional categorization automatically (De Houwer and Hermans, 1994). Such categorization rather reflects good/bad, pleasant/unpleasant, positive/negative judgments at

[1] Two studies are excluded from this list. These employ a variant of the Word-Face Stroop task with right/left visual hemi-field presentations (Anes and Kruer, 2004) and faces presented in a color overlay (Isaac et al. 2012).

[2] Başgöze (2008) and Hu et al. (2012) emphasized that there was a trend towards significant interaction in the reaction times for incongruent versus congruent conditions and subject groups.

an early stage of information processing (De Houwer and Hermans, 1994). This mechanism develops early (Posner et al., 2005) and allows for representation of more sophisticated emotions through the semantic palette later (Gökçay, 2011). Since MDD patients ruminate continually about negative evaluations (Andrews and Thomson, 2009), emotional categorization related to positive/negative judgments is more relevant to observe their behavioral differences. Finally, detecting behavioral differences in interference is of utmost importance in this task. Interference will only be observed when the irrelevant stimulus modality has privileged access compared to the target stimulus modality (De Houwer and Hermans, 1994). Verbal stimuli require longer processing times while their emotional salience is detected. Given the biological salience of faces, face stimuli can be processed rapidly, and even subliminally (Victor et al., 2010). In a Word-picture Stroop study administered by De Houwer and Hermans (1994), response to pictures with positive/negative valence were faster and more automatic, and interference was observed only when responding to words, but not pictures. Due to the profound difficulties experienced by clinically depressed subjects in social interactions, human faces expressing emotions are particularly powerful stimuli for these individuals (Gotlib et al., 2004). In a Word-Face Stroop study conducted by Strand et al. (2013) on remitted depression patients and healthy controls, interference effect is studied in two separate experiments: either words as targets or faces as targets. Results indicated interference effect only when words were used as targets and faces were used as distracters. In contrast, no interference effect was found when faces were targets and words were distracters (Strand et al., 2013). This indicates that facial information is perceived prior to lexical emotional information when both are presented at the same time. We took into account these factors while designing our Word-Face Stroop task and decided to set stimulus duration on the order of 2 s, ask emotional evaluations to be made on positive/negative judgments and let words to be evaluated as targets rather than faces.

2. Method

The study was funded through joint research projects[3] between Middle East Technical University (METU), and psychiatry departments of Ege University and Ankara University medical colleges. Regional ethical review board approvals were obtained from METU and Ankara University to conduct the procedures in accordance with the Helsinki declaration of the world medical association assembly.

2.1. Participants

Sixteen right handed medication naïve native Turkish female patients (age: 30.2 ± 7.7, education: 9.5 ± 3.8) were recruited from the outpatient service of Ege University and Ankara University Psychiatry Departments. The study was restricted with female participants because presentation rate of male patients was very low. Patients met major depressive episode criteria in DSM-IV-TR, and their diagnosis was done with SCID-I (First et al., 1996). Axis I co-morbidities such as substance/alcohol abuse and anxiety disorder were excluded. Hamilton Depression Rating Scale (HAM-D) was administered before the Word-Face Stroop task. Patients with HAM-D scores higher than 19 were admitted to the study after signing written informed consent (Total HAM-D score average of the patients was 27.1 ± 6.7). Sixteen healthy right-handed female Turkish participants (age 31 ± 9.6, education 10.8 ± 3.1), were recruited as control subjects from the university, among the networks of the co-workers upon written consent. Exclusion criteria were history of brain injury, neurological disorder, or use of mood-altering medication.

2.2. Stimuli

A total of 96 affective words[4] were selected from the in-house developed Turkish word database TUDADEN (Gokcay and Smith, 2012). Words were equally distributed across three categories: Neutral, Positive, Negative (32 words each). The distribution of the valence and arousal of the words are shown in Fig. 1. On the valence scale, the ranges and mean values of the words on a 9-item scale were as follows: Neutral range: 4.5–6 ($M=5.57 \pm 0.62$), Negative range: 1.4–4, ($M=2.89 \pm 0.79$), Positive range: 6.5–8.6 ($M=7.29 \pm 0.49$). Valence of the three word categories differed significantly from each other ($F(2,93)=369{,}690$, $MSe=157.528$, $p<0.001$). On the arousal scale, the range was 4–6.5 ($M=5.35 \pm 0.56$) and there was no significant difference across the three categories. Frequency ($M=97.4 \pm 80.3$) and length ($M=5.41 \pm 1.4$) of the words were also balanced across the categories.

As seen in Appendix A, total of 12 face pictures (4 happy, 4 sad, and 4 neutral) were chosen from The Productive Aging Lab Face Database (Minear and Park, 2004). It is important to emphasize that word and face stimuli conformed to medium intensity arousal while being evenly distributed to positive, negative and neutral categories across the valence dimension.

2.3. Procedure

The stimuli were shown on gray background at the center of a personal laptop monitor, from which the subjects were seated 60 cm away. The software program E-Prime was used to design the experiment. Participants were asked to judge the emotional valence of the words superimposed on affective faces, which necessitated inhibition of the emotion induced by the face. If participants failed to respond within 2000 ms, the experiment proceeded to a new trial. In between trials, a fixation cross appeared for 1500 ms. The participants were instructed to use their right hand's index finger to respond via 3 marked keys on the keyboard, corresponding to positive, negative and neutral answers. Trials were categorized as congruent, incongruent and neutral appearing in random order such that, each affective word is presented twice, both in congruent (valence of the word and the face are the same) or incongruent (valence of the word and the face are different) settings. Neutral words appeared only on neutral faces. The subjects responded to a total of 160 trials as illustrated in Fig. 1.

3. Results

3.1. Analyses for task validation

3.1.1. Word valence judgments

For the congruent condition with positive words, significant positive correlations were found between the number of correct responses and the valence values for both healthy ($r=704$, $p<0.01$) and depressed ($r=749$, $p<0.01$) populations. When the words had higher positive values in the valence scale (e.g. 7 or more), correct response rates increased. For the congruent condition with negative words, significant negative correlations were found between the number of correct responses and the valence values for both healthy ($r=-644$, $p<0.01$) and depressed ($r=-489$, $p<0.01$) populations. When the words had more negative values in the valence scale (e.g. 3 or less) correct response rates increased.

3.1.2. Facial expression perception

A 2×2 mixed design ANOVA on RTs with emotional face (happy, sad) as a within-subject factor and depression (healthy, depressed) as a between-subject factor, adding education and age as covariates, revealed no significant effect of any of the factors. This indicates that pictures were processed similarly across the emotion categories in both groups.

3.1.3. Gender effects

A 2×2 mixed design ANOVA on RTs with gender category (male, female) of the face stimuli as a within-subject factor and depression (healthy, depressed) as a between-subject factor, adding education and age as covariates, revealed no significant effect of any of the factors. This indicates that male and female pictures were processed similarly in both groups.

3.2. Correct response rate analysis

Table 1 summarizes the correct response rates (CRR) for all experimental conditions, where neutral stands for the condition containing neutral words appearing on neutral pictures, 'congpos'/

[3] The authors are grateful to METU BAP office and TUBITAK Turkish Scientific and Research Council for funding (no: 109E081).

[4] All the words used in the study are listed in Appendix B along with their English translations.

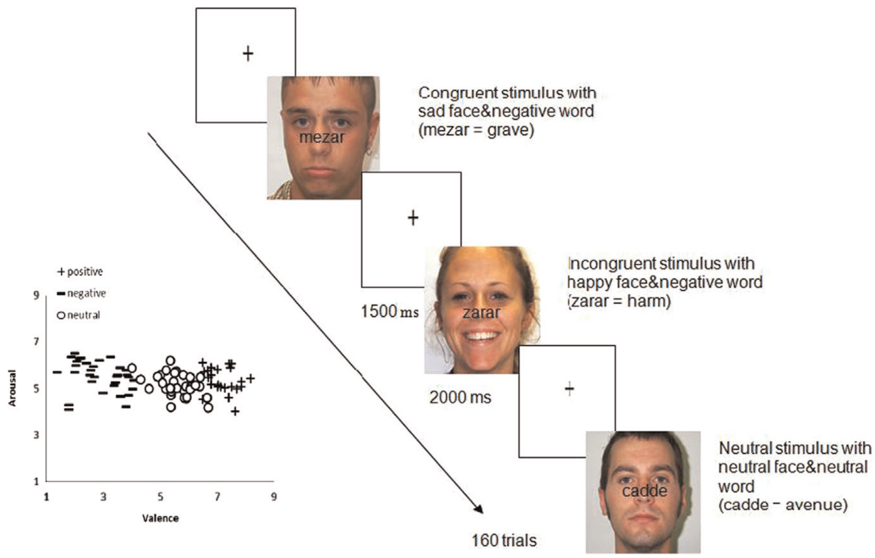

Fig. 1. The flow of an experiment session.
(Words' valence versus arousal values are depicted on the left).

Table 1
Correct response rates.

CRR (%)	Neutral	Congpos	Congneg	Incongpos	Incongneg
MDD	39.26 ± 29.20	45.71 ± 22.73	46.88 ± 22.13	38.09 ± 24.35	51.17 ± 21.77
Healthy	47.46 ± 22.74	71.49 ± 12.55	70.90 ± 15.59	69.14 ± 13.25	66.99 ± 18.68

cong: congruent, incong: incongruent, pos: positive, neg: negative.

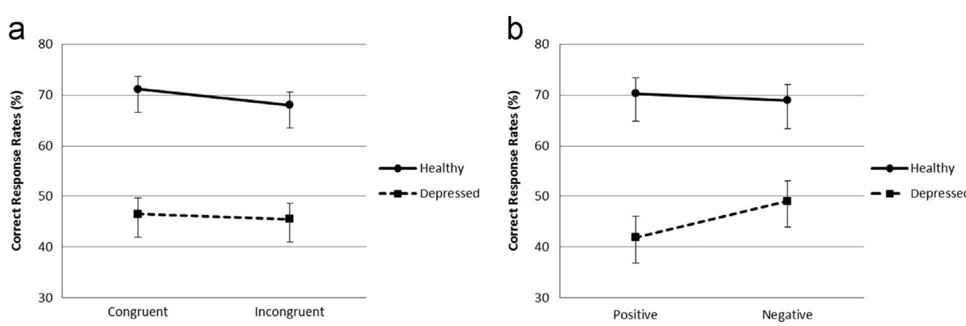

Fig. 2. Correct response rates among MDD patients and healthy controls, with SE depicted. (a) Correct response rates with respect to congruency (b) Correct response rates with respect to valence.

'congneg' are the conditions where positive/negative words are displayed under the congruent condition, and 'incongpos'/'incongneg' are the conditions where positive/negative words are displayed under the incongruent condition. As seen from Table 1, the neutral condition has the lowest CRR, both in healthy and patient groups. A $2 \times 2 \times 2$ mixed design ANOVA with congruency (congruent, incongruent) and valence (positive, negative) as within-subject factors and depression (healthy, depressed) as a between-subject factor, adding education and age as covariates, revealed a significant main effect of depression ($F(1,28)=12.491$, $\eta^2=0.308$, $p<0.01$): depressed individuals made more errors than healthy individuals regardless of the categories. Furthermore, a marginally significant 3 way interaction between congruency, valence and depression was found ($F(1,28)=4.213$, $\eta^2=0.131$, $p=0.05$). For the positive words, healthy individuals did not show any difference for the congruency factor whereas patients made more errors in incongruent cases. For the negative words, healthy individuals made more errors in responding to incongruent cases whereas patients made more errors in responding to congruent cases, as illustrated in Table 1 and Fig. 2a, b.

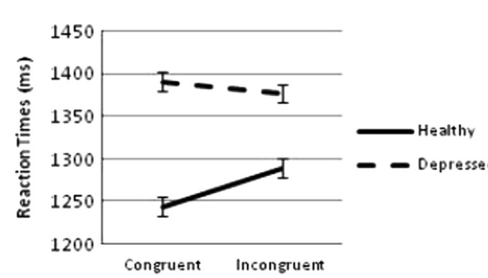

Fig. 3. Reaction times among MDD patients and healthy controls, with SE depicted.

3.3. Reaction time analysis

Table 2 summarizes the reaction times exclusively selected for correct responses of all experimental conditions, where neutral stands for the condition containing neutral words appearing on neutral pictures, 'congpos'/'congneg' are the conditions where positive/negative words are displayed under the congruent condition, and 'incongpos'/

Table 2
Reaction times for correct responses.

RT (ms)	Neutral	Congpos	Congneg	Incongpos	Incongneg
MDD	1456.71 ± 198.82	1355.66 ± 164.40	1353.07 ± 139.83	1371.50 ± 179.38	1378.84 ± 159.09
Healthy	1364.08 ± 156.62	1171.32 ± 154.94	1261.61 ± 180.68	1206.53 ± 145.45	1282.29 ± 169.20

cong: congruent, incong: incongruent, pos: positive, neg: negative.

'incongneg' are the conditions where positive/negative words are displayed under the incongruent condition. The reaction time (RT) data is converted into *z*-scores and outliers are filtered out by removing data with values above 3 and below −3. 12% of the data is removed at this step afterwards, RT data for only correct responses are extracted. For the patients and controls, approximately 45% and 70% of the data survived this step. As seen from Table 2, the neutral condition has the highest RT, both in healthy and patient groups. With lowest CRR and highest RTs, the neutral condition deserves separate in-depth consideration. Therefore, results in this category will be reported elsewhere. A 2 × 2 × 2 mixed design ANOVA with congruency (congruent, incongruent) and valence (positive, negative) as within-subject factors and depression (healthy, depressed) as a between-subject factor, adding education and age as covariates, revealed a significant difference only for depression $F(1,28)=6.298$, $\eta^2=0.184$, $p<0.05$. Patients were significantly slower than healthy controls in all categories. No significant interaction or congruence or valence differences are found between the patients and healthy controls.

3.4. Item-wise RT analysis

When only correct responses are analyzed, data loss is large, especially for the patient group. Therefore we repeated the RT analysis by entering correctness as a factor for each trial. This necessitated an item-wise analysis, based on individual words. Since the number of words is higher than the number of subjects in this study, we anticipated that item-wise analysis which takes into account each word's RT across subjects might return stronger statistical significance regarding effects between conditions. We conducted a 2 × 2 × 2 × 2 mixed design ANOVA with congruency (congruent, incongruent), correctness (correct, incorrect), group (healthy, depressed) and valence (positive, negative) as factors on reaction times. Significant differences for the RTs in congruency ($F(1,50)=20.505$, $\eta^2=0.291$, $p<0.001$), and group ($F(1,50)=100.878$, $\eta^2=0.669$, $p<0.001$) are found. The main effect on valence was marginally significant ($F(1,50)=3.942$, $\eta^2=0.998$, $p=0.053$) such that regardless of the categories, reaction times for negative words were larger compared to positive words. The interactions between congruency and group ($F(1, 50)=4.553$, $\eta^2=0.083$, $p<0.05$) as well as congruency and correctness ($F(1, 50)=5.582$, $\eta^2=0.100$, $p<0.05$) were significant. As seen in Fig. 3, the healthy population slows down significantly for the incongruent cases, whereas this effect is absent in the MDD subjects.

4. Discussion

'The human brain protects the processing of task-relevant stimuli from interference (i.e. conflict) by task-irrelevant stimuli via attentional biasing mechanisms' (Egner et al., 2008). Such mechanisms may require effortful processing to inhibit the interference, which in turn is observed as a slowdown for emotionally incongruent situations. The valence-based Word-Face-Stroop task used in this study measures interference effects related to the valence aspect of emotion exclusively, since arousal levels of the words lie in the neutral range. During the valence-based Word-Face Stroop task, healthy controls exhibited the two well-known effects: (1) Stroop interference effect. (2) Positive word advantage. Faster response rates and higher number of correct responses were observed for the congruent and positive conditions as seen from Tables 1, 2 and Figs. 2, 3. When reaction times are considered according to the item-wise analysis, the emotional interference effect is significant, and the positive word advantage is marginally significant. These results indicate that our valence-based Word-Face Stroop task replicated the findings in the literature in emotional conflict resolution (Epp et al., 2012) as well as those in other Word-Face Stroop tasks (Stenberg et al., 1998; Haas et al., 2006; Etkin et al., 2006; Başgöze, 2008; Egner, 2008; Zhu et al., 2010; Hu et al., 2012; Strand et al., 2013; Chechko et al., 2013).

On the other hand, when compared to healthy controls, MDD patients exhibited slower reaction times and lower correct response rates independent of experimental conditions. This was evident from standard ANOVA analyses on correct response rates and reaction times. This is also a replication of the earlier findings regarding slowness due to mood-congruent biases in MDD patients (Epp et al., 2012). Unfortunately, the reaction time differences between experimental conditions within subjects are subtle, and error rates are extremely high for MDD patients. These factors prohibit detailed analysis between experimental conditions using standard methods which exclude erroneous responses. On the other hand, item-wise analysis allows for the inclusion of correct responses as a factor. Furthermore, the number of items in experimental conditions is higher than the number of subjects, which facilitates reaching at more powerful statistical significance. Using item-wise analysis, a new finding is achieved: There exists a statistically significant interaction between congruency and group for reaction times – in addition to significant main effects in all three conditions, congruency, group and valence.[5] The healthy group slowed down while processing words overlaid on emotionally incongruent facial expressions. But MDD patients failed to demonstrate such an interference effect; they responded similarly for words that appeared on incongruent facial expressions compared to words that appeared along with a congruent facial expression. Previously, two other studies (Başgöze, 2008, Hu et al., 2012) have indicated a trend towards significant interaction in congruency in a similar group of subjects, namely un-medicated MDD patients and healthy controls. Although the observed interaction between patients and healthy controls was significant, the underlying interference effect was subtle. Perhaps the reason for not detecting a statistically significant behavioral difference between the patients and controls in those studies was due to adherence with standard statistical analysis instead of item-wise analysis.

Finally, several statistical tests indicated the validity of the experimental parameters used in our valence-based Word-Face Stroop task. When reaction times are considered, there were no statistically significant differences for the expressions or gender of the face pictures in the background between healthy controls and patients. When correct response rates of patients and healthy controls were studied for evaluation of the words' valence, both groups had higher correct response rates when the emotional intensity of the word increased. From these results we can conclude that in the patient group, there was no deficit in the processing of emotional attributes. The MDD

[5] The main effect in valence was marginally significant: $p<0.053$.

patients' emotional percepts and cognitive abilities were intact.

Regarding this task, the behavioral differences in the patient group seem to be restricted to the processing of emotionally incongruent information. The implicit emotional processing of faces in the background might be a weaker distracter for the patients, probably due to repetition effects brought by continual rumination related with conflicts. The negative attention bias reported in the MDD patients in earlier studies which used words as targets (Erickson et al., 2005; Karparova et al., 2007; Fritzsche et al., 2010) is insufficient to account for the differential behavior of the patients in our study. If it was true that the attention biases in the patients facilitated the processing of negative words, then effects of this should have been apparent for both congruent and incongruent negative words. However, according to Tables 1 and 2, in the congruent condition, the processing of negative words is not facilitated. On the other hand, according to the depressive rumination account, when given a laboratory task, depressed people ruminate about other things, and consume limited cognitive resources. Such rumination diminishes their ability to perform well on the assignment (Andrews and Thompson, 2009). It can be claimed that when the laboratory task on which performance is to be evaluated is related closely to the depressogenetic problem such as sustained analysis or problem solving on a complex task, patients might perform better (Andrews and Thompson, 2009). We have observed that MDD patients have higher correct response rates for the incongruent negative condition (Table 1). In addition, their responses for the incongruent condition do not demonstrate the interference related slowdown. This performance is consistent with their continual rumination about conflicts and negative concepts. It has been suggested that the patients' performance is restored when another task is initiated first, to distract them from their ruminations (Andrews and Thompson, 2009). This might be tried, to see whether upon such distraction, the patients start showing the natural interference related slowdown for the incongruent condition. Another test could be changing the design of the Word-Face Stroop task to include emotional primes before each trial (Hart et al., 2010). This way, it may be possible to identify facilitation effects the patients naturally employ.

Some of the limitations of this study are as follows. Due to time restrictions of patients in the outpatient clinic, IQ assessment was not done. Instead, years of education was used. The emotional stimuli were not evaluated across the valence and arousal dimensions by the participants. Instead, the norm values obtained from healthy populations in the TUDADEN (Gokcay and Smith, 2012) and Productive Aging Lab Face Database (Minear and Park, 2004) were used. Furthermore, the admission procedure did not include record-keeping on the menstrual cycle of the participants, hence we were unable to add this information as a covariate. Clinical assessment of the patients was done according to the HAM-D scores but for healthy controls, only short interviews were conducted, inquiring their sanitary background and mood within the past months. Finally, in order to be able to generalize the findings, a larger cohort must be studied – the low number of participants is an important limitation of this study.

In conclusion, the results we have obtained indicate that the valence-based Word-Face task presented here may have clinical relevance. This task may be administered to a larger pool of medicated/un-medicated MDD, remitted MDD and dysphoric patients to see whether it has premise to distinguish among levels of depressive experience.

Hopefully in the future, existing theoretical (Beck, 1979, 2008; Bower, 1991) and neuroscientific (Mayberg, 1999, 2003) models of depression may be augmented in the light of these findings.

Acknowledgment

The data collection in this study is partially funded by TUBITAK project 109E081, and by BAP institutional funds of METU.

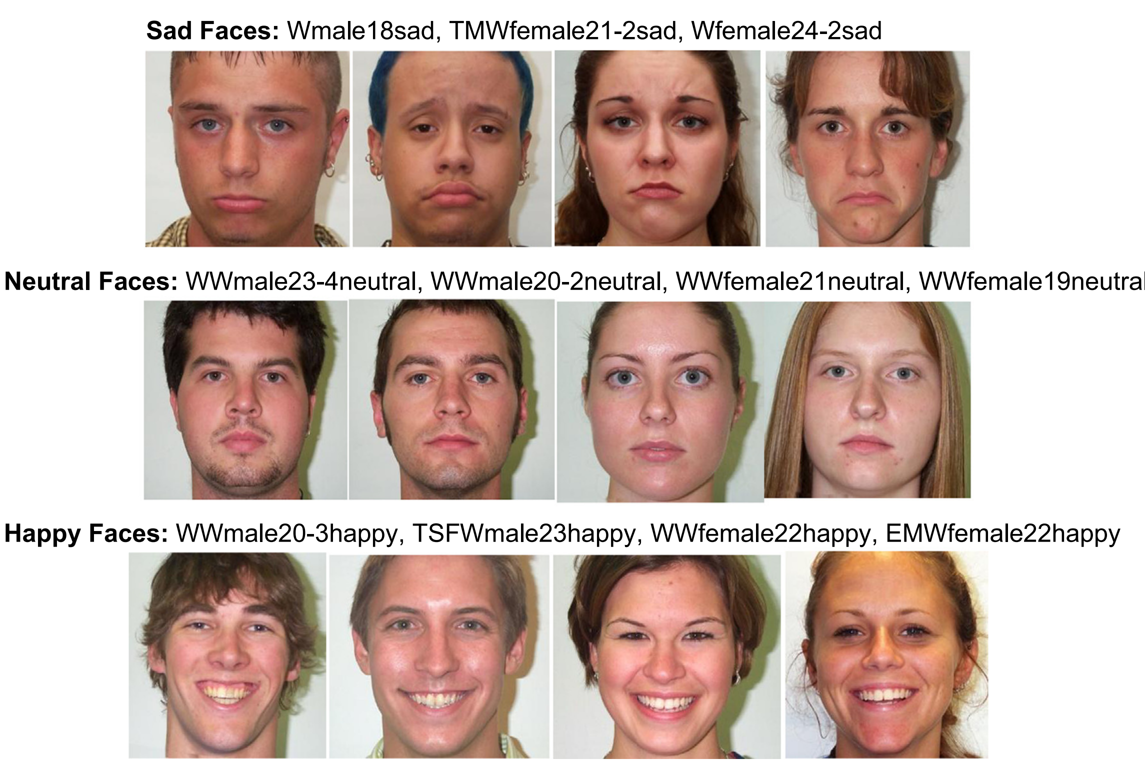

Fig. A1. Face stimuli.

Appendix A

See Fig. A1 here.

Appendix B

Word stimuli (*: no exact translation), (+: positive, 0: neutral, −: negative).

+ Words in Turkish	English translation	− Words in Turkish	English translation	0 Words in Turkish	English translation
Hayal	Imagination	Gurbet (*)	Absence from home	Ruh	Soul
Melek	Angel	Eksiklik	Lack	Kısmet	Destiny
Gönül (*)	Heart/mind	Cehennem	Hell	Gelenek	Tradition
Ilham	Inspiration	Kuşku	Suspicion	şaşkınlık	Maze
şans	Luck	Yalnızlık	Loneliness	Dürtü	Impulse
Fikir	Idea	Sıkıntı	Problem	Ciddiyet	Solemnity
Dilek	Wish	Nefret	Hatred	Korunma	Protection
Sevgi	Love	üzüntü	Sadness	Vicdan	Conscience
Cesaret	Courage	Dert	Distress	şüphe	Doubt
Sevda (*)	Passionate love	Kötülük	Malice	Prensip	Principle
şeref	Honor	özlem	Longing	Yorum	Comment
Ilgi	Interest	Hüzün	Gloom	Itiraz	Objection
Dostluk	Friendship	Bunalım	Melancholy	Ibadet	Worship
Itibar	Reputation	ısrar	Insistence	Deyim	Idiom
Uyum	Harmony	Zarar	Harm	Kural	Rule
Zevk	Pleasure	Uyarı	Warning	Etken	Factor
Kokteyl	Cocktail	çöplük	Trash	Fırın	Oven
Sinema	Cinema	Iltihap	Inflammation	Karayolu	Highway
Madalya	Medal	Leke	Stain	Perde	Curtain
Yazar	Writer	Sis	Fog	Arı	Bee
Koşu	Running	Yara	Wound	Davul	Drum
Posta	Mail	Gözyaşı	Tear	Demir	Iron
Konser	Concert	Sivrisinek	Mosquito	Ray	Rail
Kas	Muscle	Mezar	Grave	Cadde	Avenue
Bulut	Cloud	çamur	Mud	şapka	Hat
şarap	Wine	Tütün	Tobacco	Vadi	Valley
Ayna	Mirror	Gübre	Fertilizer	Asansör	Elevator
Salata	Salad	Idrar	Urine	Kaya	Rock
Elma	Apple	Toz	Dust	Pizza	Pizza
Altın	Gold	Cenaze	Funeral	Kedi	Cat
Bebek	Baby	Ilaç	Medicine	Kumaş	Fabric
Gemi	Ship	Hastane	Hospital	ördek	Duck

References

Andrews, P.W., Thomson, J.A., 2009. The bright side of being blue: depression as an adaptation for analyzing complex problems. Psychol. Rev. 116 (3), 620–654.

Anes, M.D., Kruer, J.L., 2004. Investigating hemispheric specialization in a novel face-word stroop task. Brain Lang. 89 (1), 136–141.

Bai, L., Ma, H., Huang, Y.X., Luo, Y.J., 2005. The development of native Chinese affective picture system—a pretest in 46 college students. Chin. Mental Health J. 19, 11.

Başgöze, Z., 2008. Emotional Conflict Resolution in Healthy and Depressed Populations (Master's thesis). Informatics Institute, Middle East Technical University, Ankara, Turkey.

Beck, A.T., Rush, A.J., Shaw, B.F., Emery, G., 1979. Cognitive Therapy of Depression: A Treatment Manual. Guilford Press, New York.

Beck, A.T., 2008. The evolution of the cognitive model of depression and its neurobiological correlates. Am. J. Psychiatry 165, 969–977.

Bower, G.H., 1991. Mood congruity of social judgments. In: Forgas, J.P. (Ed.), Emotion and Social Judgments. Pergamon Press, Oxford.

Chechko, N., Augustin, M., Zvyagintsev, M., Schneider, F., Habel, U., Kellermann, T., 2013. Brain circuitries involved in emotional interference task in major depression disorder. J. Affect. Disord. 149, 136–145.

Clasen, P.C., Wells, T.T., Ellis, A.J., Beevers, C.G., 2013. Attentional biases and the persistence of sad mood in major depressive disorder. J. Abnorm. Psychol. 122 (1), 1–25.

De Houwer, J., Hermans, D., 1994. Differences in the affective processing of words and pictures. Cognit. Emot. 8 (1), 1–20.

Egner, T., Etkin, A., Gale, S., Hirsch, J., 2008. Dissociable neural systems resolve conflict from emotional versus nonemotional distracters. Cereb. Cortex 18 (6), 1475–1484.

Ekman, P., Friesen, W.V., 1976. Pictures of Facial Affect, ,. Consulting Psychologists, Palo Alto, CA.

Ekman, P., Friesen, N., Hager, J.C., 2002. The Facial Action Coding System. A Human Face, Salt Lake City, Utah.

Epp, A.M., Dobson, K.S., Dozois, D.J.A., Frewen, P.A., 2012. A systematic meta-analysis of the stroop task in depression. Clin. Psychol. Rev. 32, 316–328.

Erickson, K., Drevets, W.C., Clark, L., Cannon, D.M., Bain, E.E., Zarate Jr., C.A., Charney, D.S., Sahakian, B.J., 2005. Mood-congruent bias in affective go/no-go performance of unmedicated patients with major depressive disorder. Am. J. Psychiatry 162 (11), 2171–2173.

Etkin, A., Egner, T., Peraza, D.M., Kandel, E.R., Hirsch, J., 2006. Resolving emotional conflict: a role for the rostral anterior cingulate cortex in modulating activity in the amygdala. Neuron 51 (6), 871–882.

First, M., Gibbon, M., Spitzer, R., Williams, J., 1996. User's Guide For the Structured Clinical Interview For DSM IV Axis I Disorders Biometrics Research Department. New York State Psychiatric Institute, New York.

Fritzsche, A., Dahme, B., Gotlib, I.H., Joormann, J., Magnussen, H., Watz, H., Nutzinger, D.O., Leupoldt, A., 2010. Specificity of cognitive biases in patients with current depression and remitted depression and in patients with asthma. Psychol. Med. 40 (5), 815–826.

Gohier, B., Ferracci, L., Surguladze, S.A., Lawrence, E., El Hage, W., Kefi, M.Z., Allain,

P., Garre, J.B., Le Gall, D., 2009. Cognitive inhibition and working memory in unipolar depression. J. Affect. Disord. 116 (1–2), 100–105.
Gökçay, D., Smith, M.A.,2012. TUDADEN: Türkçede Duygusal ve Anlamsal Değerlendirmeli Norm Veri Tabanı, In: Bingöl H. (Ed.), Bilgisayar ve Beyin, Pan Yayıncılık.
Gökçay, D., 2011. Emotional Axes: Psychology, Psychophysiology and Neuroanatomical Correlates, In: Gökçay, D., Yıldırım, G. (Eds.), Affective Computing And Interaction: Psychological, Cognitive and Neuroscientific Perspectives. IGI Global.
Gotlib, I.H., Krasnoperova, E., Yue, D.N., Joormann, J., 2004. Attentional biases for negative interpersonal stimuli in clinical depression. J. Abnorm. Psychol. 113, 127–135.
Gur, R.C., Sara, R., Hagendoorn, M., Marom, O., Hughett, P., Macy, L., Turner, T., Bajcsy, R., Posner, A., Gur, R.E., 2002. A method for obtaining 3-dimensional facial expressions and its standardization for use in neurocognitive studies. J. Neurosci. Methods 115, 137–143.
Haas, B.W., Omura, K., Constable, R.T., Canli, T., 2006. Interference produced by emotional conflict associated with anterior cingulate activation. Cogni. Affect. Behav. Neurosci. 6 (2), 152–156.
Hart, S.J., Green, S.R., Casp, M., Belger, A., 2010. Emotional priming effects during stroop task performance. NeuroImage 49, 2662–2670.
Holmes, A.J., Pizzagalli, D.A., 2007. Task feedback effects on conflict monitoring and executive control: relationship to subclinical measures of depression. Emotion 7 (1), 68–76.
Hu, Z., Liu, H., Weng, X., Northoff, G., 2012. Is there a valence-specific pattern in emotional conflict in major depressive disorder? An exploratory psychological study. Plos One 7 (2), e31983.
Isaac, L., Vrijsen, J.N., Eling, P., Oostrom, I., Speckens, A., Becker, E.S., 2012. Verbal and facial-emotional stroop tasks reveal specific attentional interferences in sad mood. Brain Behav., 74–83.
Joormann, J., Gotlib, I.H., 2007. Selective attention to emotional faces following recovery from depression. J. Abnorm. Psychol. 116 (1), 80–85.
Karparova, S.P., Kersting, A., Suslow, T., 2007. Deployment of attention in clinical depression during symptom remission. Scand. J. Psychol. 48, 1–5.
Leppänen, J.M., 2006. Emotional information processing in mood disorders: a review of behavioral and neuroimaging findings. Curr. Opin. Psychiatry 19 (1), 34–39.
Markela-Lerenc, J., Kaiser, S., Fiedler, P., Weisbrod, M., Mundt, C., 2006. Stroop performance in depressive patients: a preliminary report. J. Affect. Disord. 94, 261–267.
Mayberg, H.S., Liotti, M., Brannan, S.K., McGinnis, S., Mahurin, R.K., Jerabek, P.A., Silva, J.A., Tekell, J.L., Martin, C.C., Lancaster, J.L., Fox, P.T., 1999. Reciprocal limbic-cortical function and negative mood: converging PET findings in depression and normal sadness. Am. J. Psychiatry 156 (5), 675–682.
Mayberg, H.S., 2003. Modulating dysfunctional limbic-cortical circuits in depression: towards development of brain-based algorithms for diagnosis and optimised treatment. Br. Med. Bull. 65, 193–207.
Minear, M., Park, D.C., 2004. A lifespan database of adult facial stimuli. Behavior Research Methods. Behav. Res. Methods Instrum. Comput. 36 (4), 630–633.
Paelecke-Habermann, Y., Pohl, J., Leplow, B., 2005. Attention and executive functions in remitted major depression patients. J. Affect. Disord. 89, 125–135.
Posner, J., Russell, J.A., Peterson, B., 2005. The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology. Dev. Psychopathol. 17 (3), 715–734.
Siegle, G.J., Steinhauer, S.R., Thase, M.E., Stenger, V.A., Carter, C.S., 2002. Can't shake that feeling: event-related fMRI assessment of sustained amygdala activity in response to emotional information in depressed individuals. Biol. Psychiatry 51 (9), 693–707.
Stenberg, G., Wiking, S., Dahl, M., 1998. Judging words at face value: interference in a word processing task reveals automatic processing of affective facial expressions. Cognit. Emot. 12, 755–782.
Strand, M., Oram, M.W., Hammar, A., 2013. Emotional information processing in major depression remission and partial remission: faces come first. Appl. Neurophyschol.: Adult 20, 110–119.
Suslow, T., Konrad, C., Kugel, H., Rumstadt, D., Zwitserlood, P., Schöning, S., Ohrmann, P., Bauer, J., Pyka, M., Kersting, A., Arolt, V., Heindel, W., Dannlowski, U., 2010. Automatic mood-congruent amygdala responses to masked facial expressions in major depression. Biol. Psychiatry 67 (2), 155–160.
Tottenham, N., Tanaka, J.W., Leon, A.C., McCarry, T., Nurse, M., 2009. The NimStim set of facial expressions: judgments from untrained research participants. Psychiatry Res. 168, 242–249.
Victor, T.A., Furey, M.L., Fromm, S.J., Öhman, A., Drevets, W.C., 2010. Relationship between amygdala responses to masked faces and mood state and treatment in major depressive disorder. Arch. Gen. Psychiatry 67 (11), 1128–1138.
Zetsche, U., Joormann, J., 2011. Components of interference control predict depressive symptoms and rumination cross-sectionally and at six months follow-up. J. Behav. Ther. Exp. Psychiatry 42 (1), 65–73.
Zhu, X., Zhanga, H., Wua, T., Luob, W., Luo, Y., 2010. Emotional conflict occurs at an early stage: evidence from the emotional face-word Stroop task. Neurosci. Lett. 478, 1–4.