Original Papers

Reyhaneh Barani Toroghi*
Zahra Zohoorian*
Majid Ghoshuni*

Item Discrimination of IELTS Reading Comprehension Section: Evidence from Event Related Potentials

Abstract: *The development of international proficiency tests such as IELTS, which entail important decision making about people's academic lives, requires complex processes to ensure item discrimination. Previous research has indicated that IELTS has been ineffective in omitting distractor components, which may offer limitations in differentiating among the candidates. Among all the sections, particular attention has been paid to the reading comprehension component and it is considered as a criterion for determining whether a person is academically literate or not. While there seems to be a deep linkage between brain and reading comprehension, Event Related Potentials (ERP), as one of the methods used for measuring brain activity, allows researchers to observe reading- related brain processes and can document neural patterns at the millisecond level. This study aimed at examining item discrimination of the reading comprehension section of the IELTS exam through ERP. With a sample of 10 participants holding the band scores from 6 to 8, the reading comprehension items of a retired version of IELTS were given to the test takers as the task or the stimulus. It was found out that there is a mismatch between the proposed category of difficulty given for the reading comprehension item types and ERP evidence.*

Keywords: *IELTS, Reading Comprehension, Item Discrimination, Event Related Potentials*

1. INTRODUCTION

The International English Language Test System (IELTS) is one of the most prominent tests employed to assess the language skills of its candidates to make crucial decisions (Hashemi & Daneshfar, 2018). This test is used in highstakes decisions-making processes concerning the candidates' eligibility for job application or work in an English speaking country (O'sullivan, 2018). When tests are used extensively, their validity and item discrimination become even more important (Moore, Morton & Price, 2012).

Item discrimination is one of the criteria for de- termining the validity of a test (Henson, Dibello, & Stout, 2018; Khairani & Shamsuddin, 2016). While much research has been done on item discrimination, few researchers have focused on finding the discrimination of the items through employing neuroscience. Understanding the hidden processes in interpreting a linguistic structure in the brain requires using some of the neuroscience methods (Meyer, 2018; Nieuwland et al., 2018; Romeo et al., 2019; Steinhauer & Connolly, 2008). This study proposes the employment of neurological models to evaluate item discrimination of tests such as IELTS.

1.1 The Reading Section of IELTS

Developing a test is a complex process that may be subject to revision even if the items are developed by skillful item writers (Danuwijaya, 2018). IELTS exam in this sense is not an exception. Research shows that IELTS cannot be a careful predictor of candidate's proficiency level of reading comprehension (Hyatt & Brook, 2009; Isaacs& Trofimovich, 2012). It seems that prior research for revalidating IELTS exam has not been successful in omitting some distracting components including luck, guessing the choice, anxiety levels, the speed of decision making and test taking capacity (Kovaleno,2018). This leads to a limitation for predicting and making differentia- tion between IELTS candidates' levels of reading comprehension (Kane & Bejar, 2014, cited in Kovalenko, 2018).

The reading comprehension section of the IELTS exam is one of the most difficult parts (Kovalenko, 2018). It is considered as a key element for educationists,

* **Department of Languages, Mashhad Branch, Islamic Azad University, Mashhad Iran**

Corresponding author: Reyhaneh Barani Toroghi, baran2020reyhani3232@gmail.com

psychologists, and test designers (Lui, 2010 cited in Kovalenko, 2018), and can determine a person's academic literacy (Moore, Morton & Price, 2012). Accordingly, test designers in different widespread language tests such as IELTS should give more prominence to revalidating the reading section (Alderson, 2000; Urquhart et al., 1998; Weir, 1993, cited in Kovalenko, 2018).

Different scholars have developed several taxonomies for the reading comprehension section of IELTS (Moor& Morton & Price, 2012). For instance, Munby's taxonomy has an influential effect on test designers (Moore, Morton & Price, 2012), or the taxonomy developed by Weir and Urquhart, which was used for the IELTS Academic Reading test (Weir et al., 2009). Among various taxonomies presented by different researchers, the current study considers the taxonomy presented by the IELTS website itself. Texts in the reading comprehension section of IELTS involve different goals (O'Reilly & Lester,2017) including following instructions, finding the main ideas, finding the relationship between main ideas, identifying the underlying concepts, and drawing logical inferences (Alderson, 2000). Thus, all the IELTS reading comprehension questions fit into this framework (Moor, Price, Morton, 2012). The IELTS taxonomy categorizes the reading comprehension questions into five levels:

Level 1: Literal Comprehension/ Following Instruction

Literal comprehension is in accordance with following instruction category. Items can be answered directly from the text (Day & Park, 2005; Gocer, 2014). Test takers should follow the instructions of the question, recall, and recognize the ideas or information in the text (Gocer, 2014).It is known as the easiest level. In this research, true/ false and not given items of passages one and three were considered in this category.

Level 2: Reorganization/ Finding the Main Idea

In this level, identifying the main points, summarizing, and synthesizing them are very important (Brown, 2019; Gocer, 2014; Lord, 2015; Riddle, 2019). Its purpose is to discover the main idea, which is a straight forward step to discover the message of the text in general (Lord, 2015; Prebianca, 2019; Riddle, 2019). Reorganization questions' level of comprehension is higher than straight forward literal comprehension questions (Day & Park, 2005; Lord, 2015). In this research, the first section of reading comprehension passage two (choosing the correct heading for each paragraph from the list of headings) was categorized in this level.

Level 3: Inferential Comprehension/ Drawing Logical Inferences

This level is one of the most challenging parts of comprehension (Ackermann, 2019; Day & Park, 2005; Gocer, 2014; Iralde & Allain, 2019; Westby, 2019). In this level, the reader should make inferences from details in the text (Van den Broek, Beker, & Oudega, 2015; Day & Park, 2005) based on the implicit information in the text and prior knowledge (Ackermann, 2019; Day& Park, 2005).

In this research the second section of passage three (complete each sentence with the correct ending) is categorized in the inferential/drawing logical inferences level.

Level 4: Evaluation/Identifying Relationships between the Main Ideas

This level deals with judgments of the reader about different paragraphs and main ideas of a text and the relevance of main ideas in a certain text (Day & Park, 2005; Gocer, 2014). Here, the reader must identify and integrate various information from different parts of a text (Day & Park, 2005; Rouet &Potocki, 2018) which leads to integration (Beker, Jolles, Lorch, & Broek, 2016; O'Brien & Cook, 2016; Raney & Bovee, 2016). In this research the second section of the second passage (match person with the correct idea) was considered in this level.

Level 5: Appreciation/Identifying the Underlying Concept

As the last level it is very difficult and challenging (Chang-Jun, 2017; Epstein, Lazarus, Calvano, & Mattews, 2002) and the reader must recognize the author's aim (Day & Park, 2005; Gocer, 2014). It deals with emotions and feelings of the reader (Chang-Jun, 2017; Day & Park, 2005; Veeravagu, Muthusamy, & Marimuthu, 2010).Comprehension in this level is a combination of three components of explicit information, implicit information, and feelings and emotions of the reader and prior knowledge of the reader toward the text (Chang-Jun, 2017; Day & Park, 2005; Veeravagu, Muthusamy, & Marimuthu, 2010). In this research the first section of the third passage (multiple choice questions) was considered in this category.

1.2 Event Related Potentials and Reading Comprehension

As far as reading is concerned it is well-accepted that there is a deep linkage between brain and reading comprehension. There has been much research on reading comprehension and linguistics through integration of some fields such as neuroscience, cognition, and education (Aldahhan, Kirby, Brien, & Munoz, 2017). Neurolinguistics is one of the subcategories of neuroscience which deals with the relationship between language and functioning of the brain. Understanding the stable points about how brain functions when it deals with a text to read is an arising step in neuroscience(Mahaney, 2018). One of the fundamental features of ERP components, as a method of brain activity measuring (Nidal& Malik, 2014), is its function of observingand documentingreading- related brain processes at the millisecond level (Molfese,Molfese, & Pratt, 2006).

ERP is an extraction of Electroencephalogram (EEG) and it is a task-related tool for measuring the electrical brain activities associated with a stimulus (Ward, 2015). There are four relevant language- related components (Friederici, 2004) each of which shows a specific subprocess (Gouvea, Phillips, Kazanina, & Poeppel, 2010) whose understanding has opened a new vision for scholars (Small &Hickok,2015; Rastelli, 2018). These components enable scholars to identify some areas in the brain that

relate to language activities including comprehension and production (Faust, 2015; Meyer, 2018). The presence or lack of a specific component reflects a sign about the source of a component and its subcategories. One of these components relates to semantic processes, the other one relates to prosodic and phonetic processes, and two of them relate to syntax and structure processes (Small & Hickok, 2015; Luck & Kappenman, 2011). The four components are the N400, P600, ELAN, and CPS. In the current study, P600 and N400 were elicited and studied.

The N400 effect relates to semantics (Loeches, et al., 2017; Romeo et al., 2019; Bornkessel-Schlesewsky, & Schlesewsky, 2008; Steinhauer & Connolly, 2008). Its study has revealed basic insights about comprehension processes (Cheytte & Plaut, 2017; Loeches, et al., 2017) and reflects the difficulty or easiness of semantic processes for challenging or easy reading comprehension texts (Cheyette & Plaut, 2017). Recent research has shown the reflection of the N400 amplitudes about implicit semantic processes (Rabovsky, Hansen, & Clelland, 2016; Zhu, et al., 2019). N400 can unfold neural dynamics of comprehension while a reader is influenced by lexical or contextual violations (Draschkow, et al., 2019; Dudschig, et al., 2019; Payne & Federmeir, 2019; Zhu, et al., 2019). It is elicited according to brain response to literal andmeta-phorical sentences with violation (Aldunate, et al., 2019).

The P600 component deals with syntactic processing and structural analysis (Osterhout & Holcomb, 1992, cited in Steinhauer & Connolly, 2008). Although many scholars have demonstrated that violation or ambiguity may evoke the P600 component (Courteau, Martignetti Netti, Royle, & Steinhauers, 2019), Kaan,Harris, Gibson, and Holcomb (2000) believed that sentences with complex structures may cause the elicitation of the P600 components as well (Steinhauer & Connolly, 2008). This component can be a reliable way to distinguish between difficult sentences/ reading comprehension questions and easy ones. "The P600 reflects late syntactic reanalysis"(Sassenhagen & Fie-bach, 2019), which in this study has been considered. In this case, the reader most often should re-read a word that appeared earlier in the sentence, and when the reader follows one interpretation of the sentence he/she may realize later that this interpretation had been wrong (Sassenhagen & Fiebach, 2019; Spotorno, Chylus, Henst, & Noveck, 2013). It is considerably task dependent (Lecky& Federmeier, 2019; Gouvea, Phillips, Kazanina, & Poepple, 2010). That is, for eliciting the P600 some tasks should be given to the target participant (Sassenha-gen & Fiebach, 2019; Herten, Chwilla, &Kolk, 2005).

Due to the importance of reading comprehension section of IELTS exam, this study aims at examining item discrimination in reading comprehension sections, as stimuli, through ERP. ERP can probe the hidden processes involved in brain activity (Sassenhagen & Fiebach, 2019).

1.3 Item Discrimination

Item discrimination evaluation is an important phase in the development of a test (Ramonda & Sevigny, 2019). It is useful for test designers in determining which items to keep, modify, or omit in a test. It is done either for obtaining difficulty level or the discrimination power (Sim & Rasiah, 2006).It is used to reveal how well an item is able to discriminate between good and poor students (Buyukturan & Sireci, 2018). It also indicates the extent to which success on an item corresponds to success on the whole test (De Ayala, 2009; Lewis, Jacobes, & Fitchett, 2012; Buyukturan & Sireci, 2018). Item discrimination is viewed from two perspectives in this study namely; traditional item discrimination (De Ayala, 2013) and neurological item discrimination.

1.4 Traditional Item Discrimination

Item Discrimination looks for appropriate items in a test to differentiate examinees (Henson& Stout, 2018). There are many ways for calculating item discrimination, but the common formula for calculating item discrimina-tion is based on classical test theory as a psychometric theory to predict difficulty level (Brown& Knight, 2012). It ranges from +1 to -1. Negative discrimination index is the opposite form of the positive discrimination (Morrow, Mood,Disch, &Kang, 2015). The criterion for selection and rejection of an item is its positive discrimination. The closer the index is to +1, the more effective the item will be (Morrow, Mood, Disch, &Kang, 2015).

1.5 Neurological Item Discrimination

The nature of item discrimination is differentiating. In this research, the researchers' aim was to recognize item discrimination of reading comprehension items of the IELTS exam without any computing procedure. The researchers of this study propose the Neurological item discrimination based on ERP language-related compo-nents. It was proposed that while participants deal with stimuli, their brain waves can be recorded and different and particular patterns with a special latency can be elicited, which can relate to the easiness or difficulty of each item. Even no change in brain waves can trigger a special sign.

Thus, the researchers attempted to monitor changes in the brain waves of the participants while they were dealing with each item in the reading comprehension section of IELTS. Items were like manipulators or stimuli, which affected the brain reactions. ERP shows the manipulating power or discriminating power of each item or how much an item can make a difference in brain waves based on its challenging level. The superiority of using ERP rather than the traditional forms is that by using ERP components the item is measured directly without any influence of metacognitive strategies. In traditional item discrimination methods, measuring item discrimination cannot omit some distracters such as "luck, anxiety levels, guessing, or the speed of decision making (Kovalenko, 2018).

2. MATERIAL AND METHODS

The present study employed a mixed methods design, which is widely used by interdisciplinary researchers (Bergin, 2018; Kozleski, 2017; Murphy, 2018; Prichard

& Trowler, 2018).It is a process in which the researcher collects, analyses, and mixes the quantitative and qualitative data (Creswell,2002) and produces explanations and numbers (Patton, 2002; Walliman, 2017). As an explanatory sequential method, qualitative phase was the first phase followed by the quantitative phase. The researchers recorded the brain waves of each participant by using PSYTASK software for showing the stimuli and WINEEG software. Using the PSYTASK and WINEEG software was simultaneous. By using WINEEG, the researchers extracted ERP graphs from brain waves. Then, the researchers changed the qualitative data (ERP graphs) to quantitative data (numbers). For analysing and interpreting the data, MATLAB software was used.

2.1 Participants

10 individuals who had taken IELTS with an average age of 29 to 35 were selected (3 females and 7 males). Entry criterion for the participants was having an IELTS band score of 6-8. Also, the participants' agreement to participate in the study through a signed written consent form was ensured. Furthermore, because the participants' concentration and attention were really important specifically for drawing a conclusion, they filled out a written personal form in order to confirm their normal and healthy neurological status and deny any psychiatric illnesses, neurological disorders, and use of sleeping pills a night before brain wave recording. Participants were paid per session. The sampling procedure was purposive sampling. The setup of conducting this research was at the biomedical engineering laboratory. Authorization from the ethics committee at the laboratory to conduct this research was received. The total number of the participants was 10 initially. However, in each section of the reading comprehension test the number of participants differed. The number of participants decreased from 10 to an average of 9- 5 in each section due to the presence or absence of artifacts in their data. Artifacts are defined as physical actions resulting from the interactions between the brain and body which may lead to biased interpretations of the EEG (Gebodh, et al., 2019; Labate, Foresta, Mammone, & Morabito, 2015). Thus, omitting data with artifacts in each section of the IELTS reading comprehension test would make the rest of data more valid and reliable.

2.2 Instrumentation

2.2.1 Reading Comprehension Section of the IELTS Retired Version

A retired version of the IELTS published in 2018 was used. Three academic reading passages were extracted each of which included a different number of reading comprehension questions. Some of the items in each category were omitted because more items could lead to fatigue which could negatively affect the participants' brainwaves. Fatigue causes malfunction on ERPs (Boksem, Meijman, & Lorist, 2005). Also, some items were removed because at the end of the analysis their single trial

ERP average was calculated as zero and their artifact which shows the noise was ten. Also, the open-ended items of the IELTS were omitted due to the fact that filing the spaces would be related to creativity and the creativity component in still unknown in ERP (Reymond, Delahunty, & Seery, 2018). Thus, the final number of items was 19 (Appendix A). Eliminating the data with artifacts is used for filtering the ERPs (Calderson & Luck, 2014) and in order to have more reliable ERP results, data with high artifacts also had tobe deleted (Salo, Mutanen, Vaalto, & IImoniemi, 2020).

2.2.2 ERP Recording Related Software

ERP is an effective non-invasive brain stimuli (Shahsavar, Ghoshuni, & Talaei, 2018) which can effectively record brain activity during reading comprehension. To carry this out, the Psytask software was employed which enables the researchers to design various psychological tasks (Aleksandrov et.al, 2019) and present them. In this study reading comprehension questions were used as the stimuli. To design the stimuli the first step was adding the new task to the Psytask software with BMP pictures and then organizing the presentation order. Finally, the stimuli were randomized. Each stimulus was presented 10 times with a duration of 2500ms. Additionally, stimuli were randomized. Finally, the researchers analysed and interpreted the data by WINEEG and MATLAB software. WINEEG is a software that allows researchers to record, edit, and analyse the participants' brainwaves (www. Mitsar-medical.com).MATLAB is a numerical computing system and it integrates computation, data analysis, engineering graphics, and algorithm developments (Pavlov, 2017).

2.3 Research Procedure

The stages of the study were explained to each participant separately. Three academic reading passages extracted from a retired version of IELTS were given to the participants to be read. They had to answer the reading comprehension questions on a paper answer sheet. According to IELTS standards, 20 minutes was given for each passage. The first passage consisted of 6 true-false and not given items; the second one included 6 items 4 of which were omitted as they related to the choosing headline for each paragraph, also there were 4 questions of underlying concept in the format of matching names with headlines one of which was omitted, and the third passage consisted of 3 sections, the first section included 5 multiple choice items two of which were omitted. The next section of the third passage had 6 matching ideas items 4 of which were omitted, and the third part of the third passage had 3 true-false items. Totally, there were 6 sections with 19 items.

The answer sheets of the previous phase were taken from the participants, but reading passages were still available to them. The stimuli were presented in ODD-BALL paradigm which is an experimental design used within psychology research (Jared, Jouravlev, & Alexander, 2018). Presentations of sequences of repetitive stimuli

are frequently interrupted by a deviant stimulus. The reaction of the participants to this stimulus is recorded (Cohen, Sullivan, & Rico, 2010). Some stimuli are called rare and others are frequent (Cohen, Sullivan, & Rico, 2010). Stimuli that are rare are more likely to elicit the ERP component (Jared, Jouravlev, &Alexander, 2018). There are two types of stimuli in ODDBALL paradigm. These stimuli are presented sequentially and randomly to the participants (Kropotov, 2010). In ODDBALL paradigm, frequent and rare stimuli were rmixed with each other and the participants had to choose the rare ones (Kahane, Lhatoo, & Luders, 2019).

To record brain waves, the skin surface was cleaned and then the attachment of the electrodes was done. Electro-Gel was used for decreasing the impedance of skin surface into 10Ώ in order to increase the signals. The electrode impedances were also kept below 10 kΏ. Initially, the EEG signal was recorded from 3 main electrodes (Fz, Cz, Pz,) attached to the scalp of the participants and both earlobes were connected. These channels (Fz, Cz, and Pz) were used because they are the representative channels. For instance Fz shows the signals of frontal part, Cz indicates the signal of the central area of the brain, and Pz signifies the signals of posterior part of the brain (Moctezuma & Molinas ,2020). Thus, these channels were comprehensive for this study. And it seemed that there was no need to consider all the channels and the researchers decided to remove15 electrodes (O1, O2, F7, F3, F4, F8, T3, C3, C4, T5, P3, P4, T6, Fp1, Fp2) and just consider representative channels including Fz, Cz, and Pz positions according to the international 10-20 system as some scholars including Shahsavar, Ghoshuni, and Talaei (2018) conducted their experiment based on the Fz, Cz, and Pz positions. Furthermore, another reason for removing 15 channels in different studies is because of their role in making EM6 noises and artifacts. Also, Signals which are recorded from T4 channel are not reliable in this study because this channel shows Muscle contractions. Furthermore, Moctezuma and Molinas (2020) stated that "Interestingly, when using all the EEG channels available, lower accuracies were achieved compared to the case when EEG channels were selected by NSGA-II or NSGA-III; i.e in patient 19 we obtained an accuracy of 0.95 using all the channels and 0.975 using only two channels selected by NSGA-III. The results obtained are encouraging and it has been shown that it is possible to classify epileptic seizures using a few electro-

des, which provide evidence for the future development of portable EEG seizure detection devices". A 15-inch wide screen monitor was used for visual stimuli. To show the stimuli, the Psytask software was employed. Nineteen reading comprehension questions were shown separately and their answers were shown randomly, and the participants were allowed to choose just one answer for each reading comprehension question. The participants were requested to left click as soon as they saw the right answer. Each item was shown 10 times randomly according to the ODDBALL paradigm. While the participants were doing the tasks, simultaneously, their EEG signals were recorded. Finally, ERP signals were analysed by MATLAB and WINEEG software in order to identify item discrimination of reading comprehension questions.

2.4 Statistical Analysis

In order to extract ERPs from EEG signals, the parameters for ERP computation including artifact processing level, thresholds for channels, and synchronization type were determined in the WINEEG software. Moreover, the numerical labels were given to target and non-target items in order to extract ERPs. Figure 1 shows the process of adjusting the parameters to extracting ERPs from EEG signals. Also, figure 2 shows the raw ERPs recorded from one of the participants for one of the reading comprehension sections of the IELTS exam.

Then, the group information of ERPs was extracted which included total, averaged, error, omission, commission, artifact, RT1, RT2, Var(RT1), and Var (RT2) to illustrate clear information about every single trial of ERP. In this research, the required data were just averaged of a single trial of ERP and artifacts and other data related to the information group of ERPs were not considered. Table 1 presents a report for one of the participants with reliable data.

Table 2 presents data with artifacts. In order to have effective ERPs, the noisy data (data that contains artifacts) must be deleted (Throne & Hammand, 2015; Patel et al., 2017; Hekmatmanesh, et al., 2019). Thus, in the current research the researchers omitted ERPs with an average less than 5, or in other words, the single trial of ERPs with an amplitude that exceeded 100 microvolts were removed.

Finally, the number of participants with reliable data decreased from 10 to an average of 9-5 in different sections. The data format was changed to txt files and
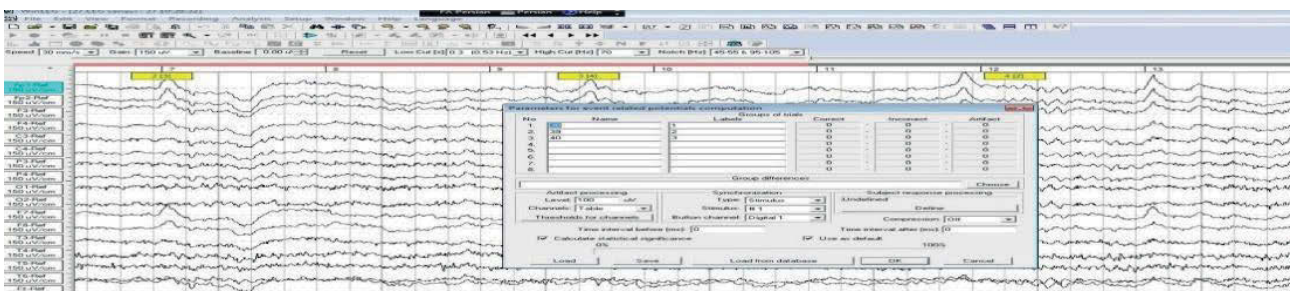

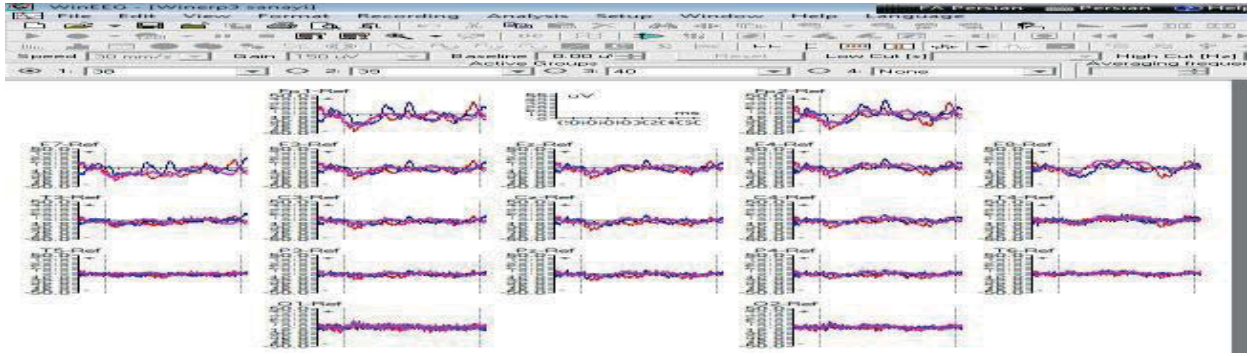**Figure 1 The Process of Extracting ERPs in WINEEG software**

Figure 2 Extracted Raw ERPs for a sample reading comprehension section

Table 1. ERPs Information group with less than 5 artifacts for a participant

Group name	Total	Averaged	Error	Omisson	Comission	Artefact	RT1	RT2	var(RT1)	var(RT2)
i	10	9	0	0	0	1	0	0	0.0	0.0
ii	10	10	0	0	0	0	0	0	0.0	0.0
iii	10	9	0	0	0	1	0	0	0.0	0.0
iv	10	10	0	0	0	0	0	0	0.0	0.0
v	10	10	0	0	0	0	0	0	0.0	0.0
vi	10	10	0	0	0	0	0	0	0.0	0.0
vii	10	9	0	0	0	1	0	0	0.0	0.0
viii	10	10	0	0	0	0	0	0	0.0	0.0

Table 2. ERPs Information group with more than 5 artifacts for a participant

Group name	Total	Averaged	Error	Omisson	Comission	Artefact	RT1	RT2	var(RT1)	var(RT2)
27A	10	2	0	0	0	8	0	0	0.0	0.0
27B	10	1	0	0	0	9	0	0	0.0	0.0
27C	10	2	0	0	0	8	0	0	0.0	0.0
27D	10	3	0	0	0	7	0	0	0.0	0.0

renamed in order to import to MATLAB software. Matrixes (size=250) were made for each participant. The size of each matrix had be the same in order to have homogenous analysis. In the next level, ERP graphs for each item in each section of reading comprehension part were extracted from MATLAB software. In these ERP graphs (example figure 3) there is a comparison between the internal items of a section. They illustrate the challenging level of items of one category. The ERP analyses focused on two time windows of N400 component (300-500 m.s) and the P600 component (500-800 m.s).

In order to change the qualitative data into quantitative data, MATLAB software was used. The integral for each item was calculated to compare the item discrimina-
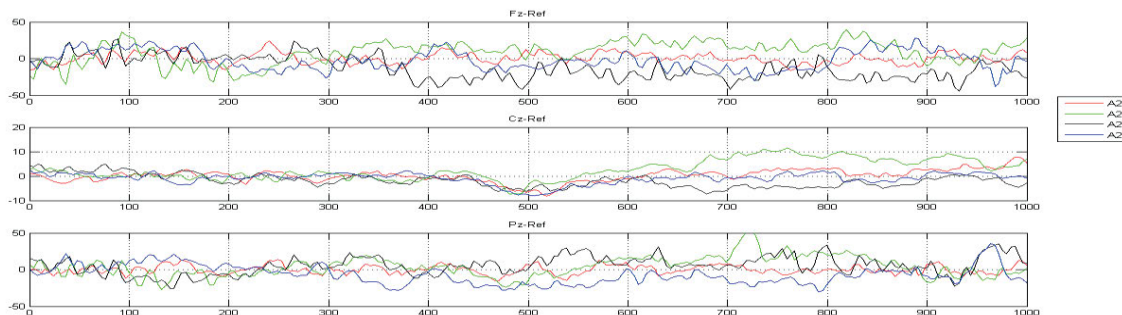

Figure 3 ERP graphs of items in the second section of reading passage two

tion of sections not the separate items. The sum and difference of target items from non-target items in each section was also calculated for N400 and P600 components. Moreover, an independent samples T-test was run to investigate the significant differences between reading comprehension sections concerning N400 and P600. Finally, bar graphs for each two sections were extracted. Figures 4 and 5 display the processes of calculating the

channel, the researchers did not consider the results of CZ and Fz channels. If one channel shows a contradictory result from other channels, the whole results will be omitted because they are not reliable (Scheffler, 2019). Table 3 displays the p-values.

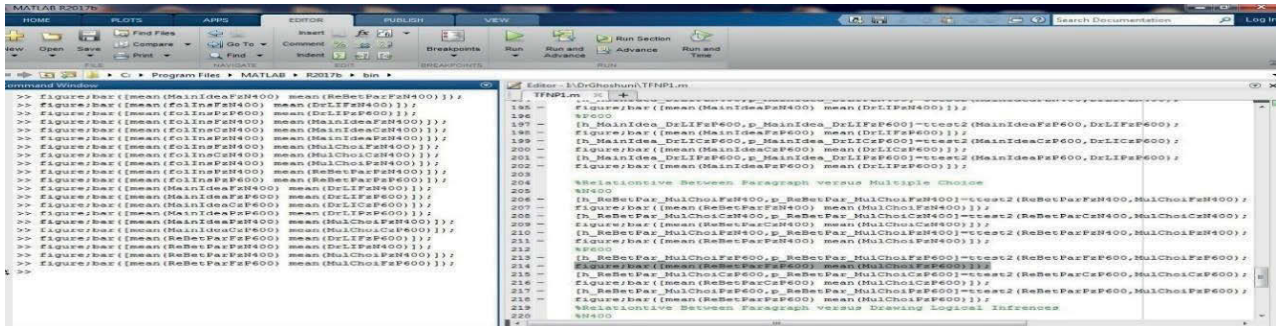Figures 5 and 6 display the ERPs of items of "following instruction" and "finding main idea" sections.

Figure 4 The processes of calculating the sum and difference of reading comprehension sections
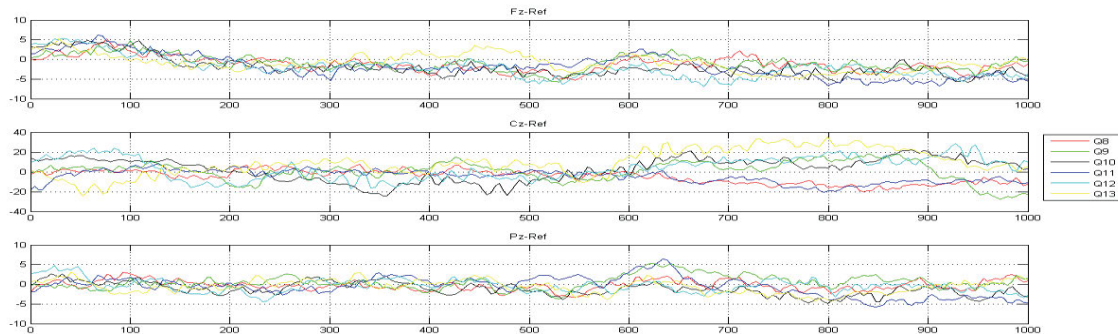

Figure 5 ERPs of "Following instructions" items in 3 channels

sum and the difference of reading comprehension sections in order to compare each two sections and extract the bar graphs using MATLAB software. Figure 4 presents the ERPs from Fz, Cz, and Pz channels elicited from each item of the second section of reading comprehension passage two. In Figure 5 four ERPs that relate to 4 items of the reading section 2 are illustrated in three channels and the amplitude of A20 is different from other items in different channels.

2.5 Statistical Analysis for Comparing the ERPs for Different Reading Sections

Results of the analysis for the first research question (Is there any significant difference in item discrimination of "following instruction" section and "finding main ideas" section based on ERP?) indicated that there was not a significant difference in item discrimination of "following instruction" and "finding main idea" sections. Thus, the first null-hypothesis was confirmed. Figures 5 and 6 display the related ERPs. Results of independent Samples T-test showed that although Cz and Fz channel reported P=0.01, P≤0.05 and P=0.03, P≤0.05 for N400 component, Pz channel showed a contradictory result from other channels. Due to the contradictory result of Pz

Table 3. P-values of "following instruction" and "finding main idea" sections

Independent sample T-test P-value		
	N400	**P600**
Cz	0.01	
Fz	0.03	
Pz	0.06	

*p≤0.05

Figure 7 displays the difference between "following instruction" (1) and "finding main idea" (2) item discrimination. The higher bar graph shows the easier section and it means that this section included less mental activity.

Results of the analysis for the second research question (Is there any significant difference in item discrimination of "following instruction" section and "identifying the underlying concept" section based on ERP?) indicated that there was a non-significant difference in item discrimination of "following instruction" and "identifying the underlying concept" sections. Thus, the second null-hypothesis was confirmed. Results of independent samples T-test showed that although Cz and Pz channels reported
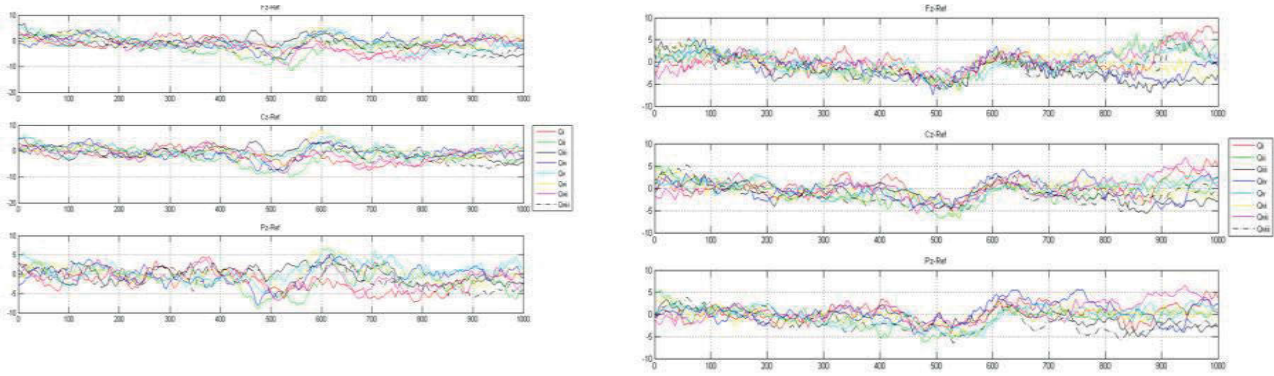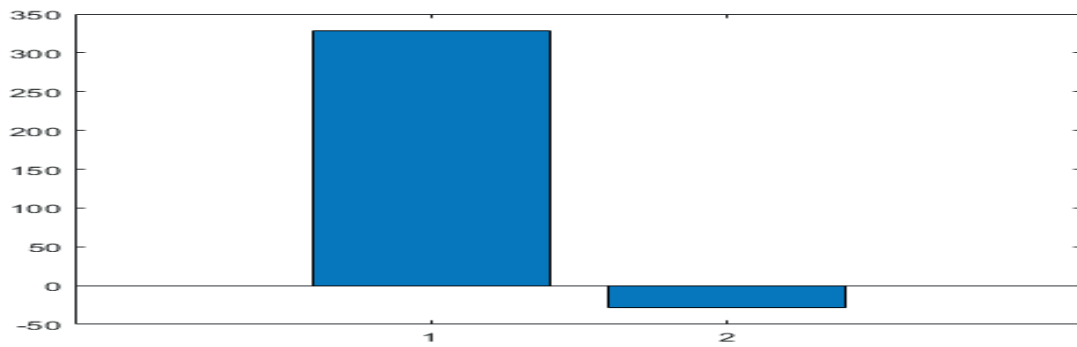
Figure 6 ERPs of "Main idea" section items

Figure 7 ERP bar graphs of item discrimination of "following instruction" and "main idea" sections in Cz channel of N400

P=0.01, P≤0.05 and P=0.03, P≤0.05 for N400 component, Fz channel showed a contradictory result from other channels because of which the researchers did not consider the results of CZ and Pz channels. Table 4 displays the p--values.

Figures 8 and 9 display the ERPs of items of "following instruction" and "identifying the underlying concept" sections.

Figure 10 displays the difference between "following instruction" (1) and "underlying concept" (2) item

Table 4. P-values of "following instruction" and "identifying underlying concept" sections

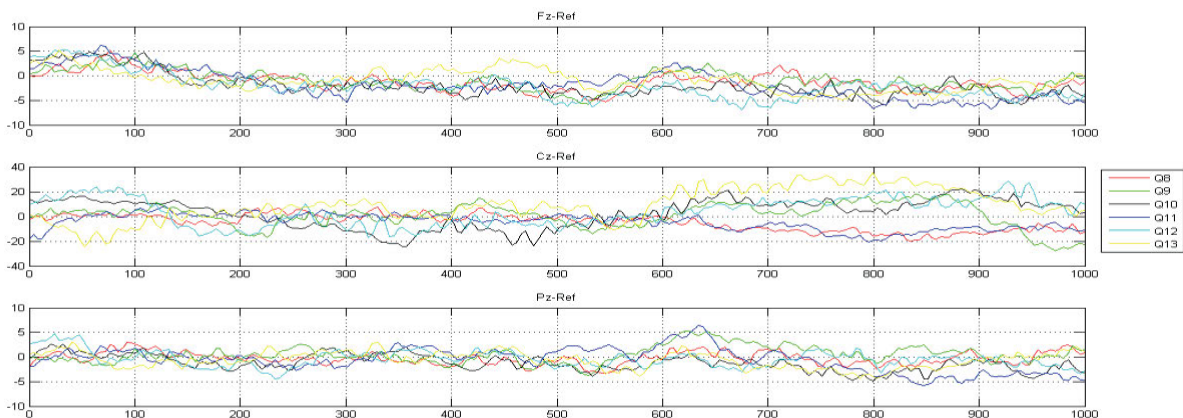|	Independent sample T-test P-value	
	N400	**P600**
Cz	0.02	
Fz	0.07	
Pz	0.01	

*p≤0.05

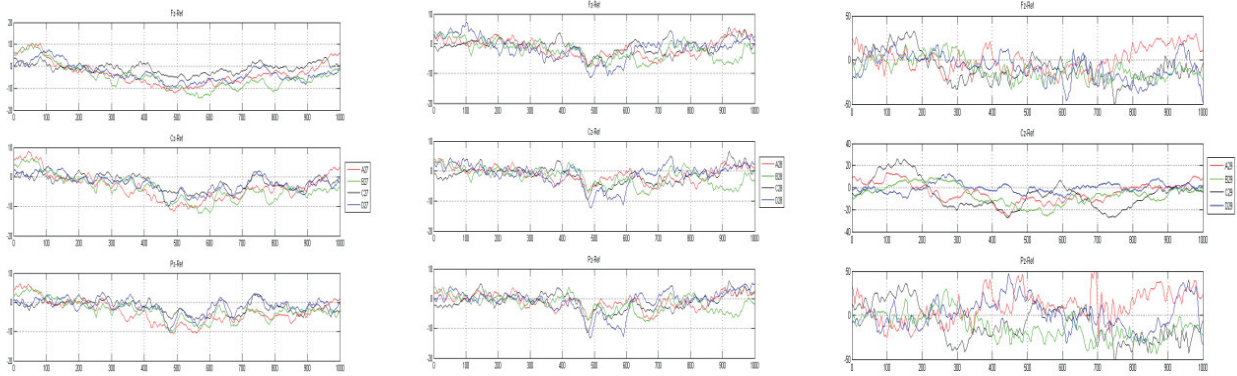**Figure 8 ERPs of "following instruction" items in 3 channels**

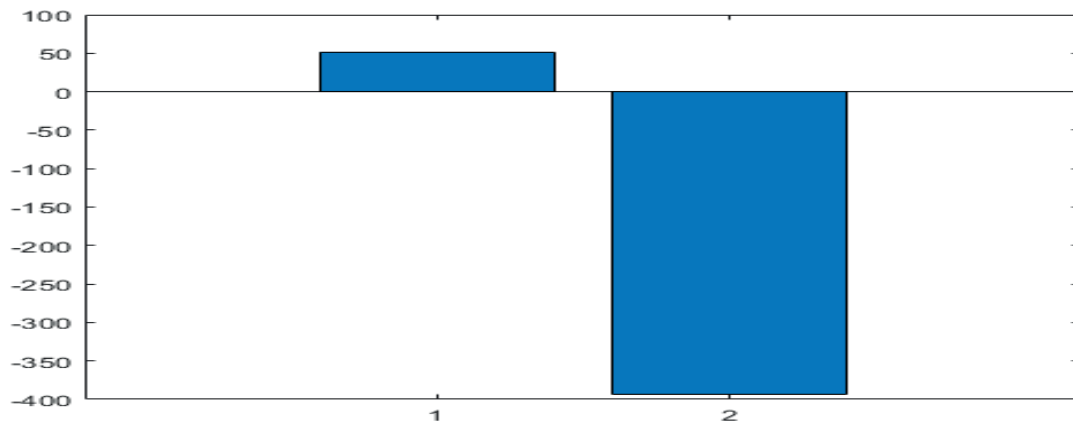**Figure 9 ERPs of "identifying underlying concept" items in 3 channels**

Figure 10 ERP bar graphs of item discrimination of "following instruction" and "underlying concept" sections in Cz channel of N400

discrimination. The higher bar graph shows the easier section and it means that this section included less mental activity.

Results of the analysis for the third research question (Is there any significant difference in item discrimination of "following instruction" section and "identifying relationships between the main ideas" section based on ERP?) indicated that the third null-hypothesis was rejected. Thus, there was a significant difference in item discrimination of "following instruction" and "identifying the relationships between main ideas" sections. The independent T-test reported p=0.01 p≤0.05 from Pz channel for the N400 component. Moreover p=0.0 , p≤0.05 was observed from Pz channel for P600 component. Furthermore, ERP results in Pz channels for P600 and N400 show that "following instruction" section is more challenging than "identifying relationships between the main ideas" section. Table 5 displays the p-values.

Figures 11 and 12 display the ERPs of items of "following instruction" and "identifying relationships between the main idea" sections.

Figures 13 and 14 display the ERP bar graphs that indicate the difference between item discrimination of "following instruction" (1) and "identifying relationships
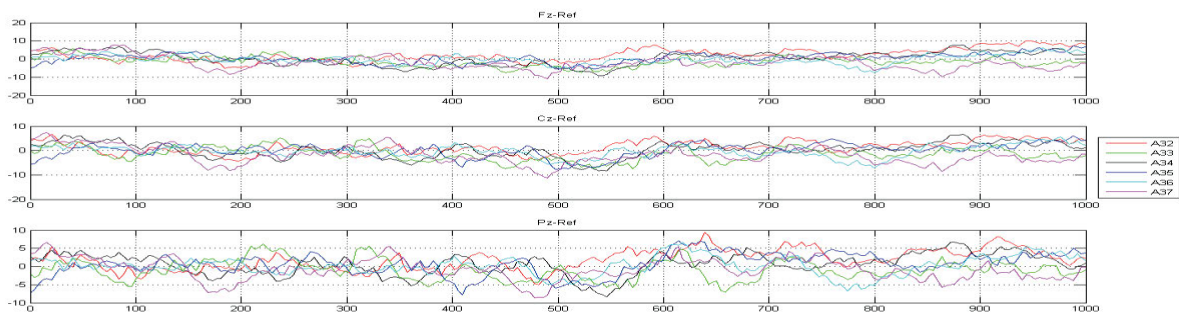

Figure 11 ERPs of following instruction items in 3 channels
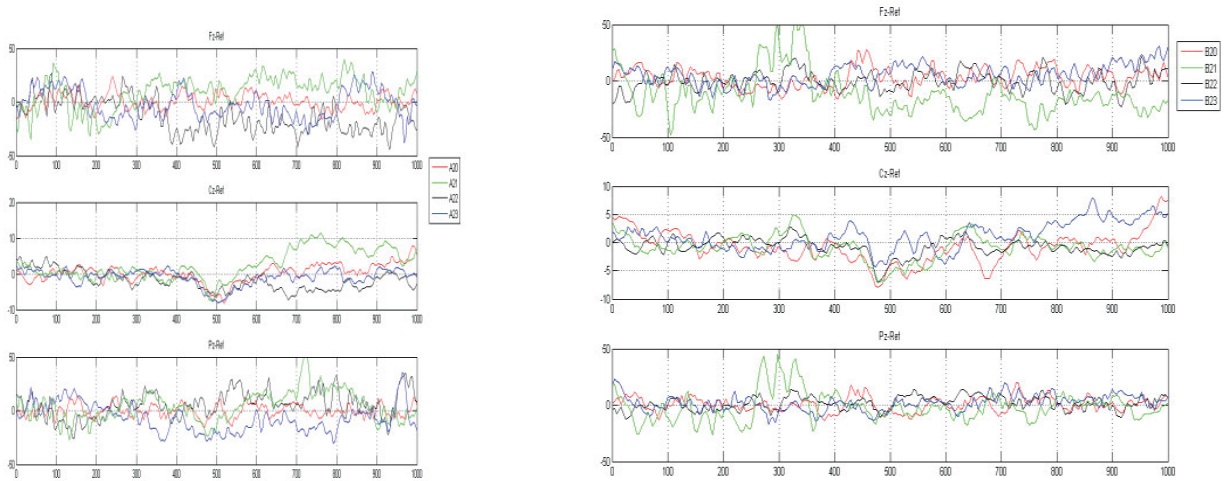
Figure 12 ERPs of "identifying relationship between main ideas" items in 3 channels

Table 5. P-values of "following instruction" and "identifying relationships between main ideas" sections

	Independent sample T-test P-value	
	N400	**P600**
Cz		
Fz		
Pz	0.01	0.00

*p≤0.05

between the main ideas" (2) sections, regarding N400and P600. This bar graph suggests a manipulation or development for the Barrett's taxonomy and IELTS taxonomy of reading comprehension categories because this ERP bar graph rejected the taxonomy presented by Barret and IELTS website.

Results of the analysis for the fourth research question (Is there any significant difference in item discrimination of "following instruction" section and "drawing logical inferences" section based on ERP?) indicated that the fourth null-hypothesis was rejected. The

Figure 13 ERP bar graphs of item discrimination of "following instruction" and "identifying relationships between the main ideas" in Fz channel of N400

Figure 14 ERP bar graphs of item discrimination of" following instruction" and "identifying relationships between the main ideas" in Pz channel of P600
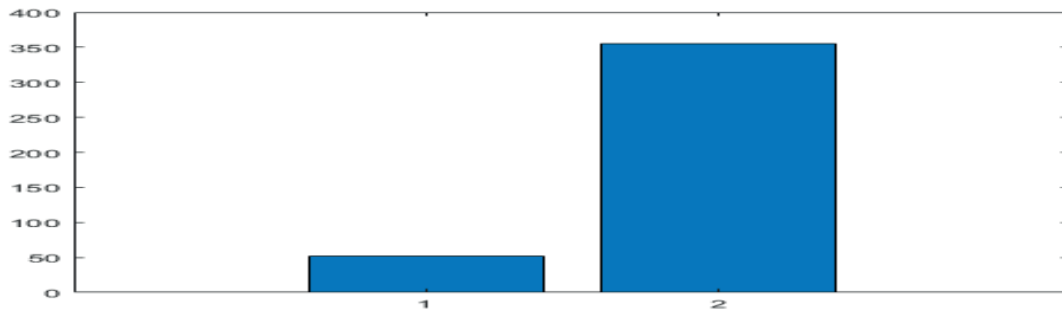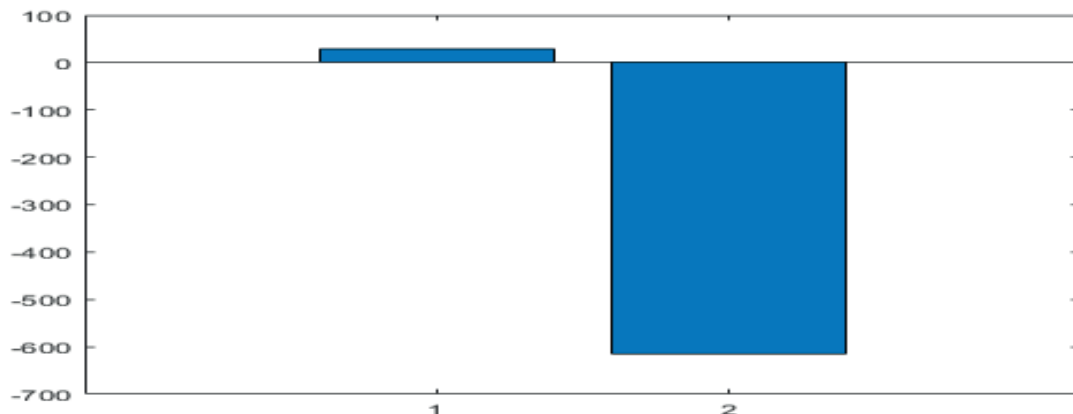
independent T-test reported *p=0.05, p≤0.05 from Pz channel for the P400 component*. Thus, there was a significant difference in item discrimination of "following instruction" and "drawing logical inferences" sections. Furthermore, ERP results in Pz channels for P600 shows that "following instruction" section is more challenging than "drawing logical inferences" section. Table 6 displays the reported p-value in Pz channel for N400 component.

Figures 15 and 16 display the ERPs of items of "following instruction" and "drawing logical inferences" sections.

Figure 17 displays the ERP bar graph that indicates the difference between item discrimination of "following instruction" (1) and "drawing logical inferences" (2) sections, regarding P600 component.

This ERP bar graph suggests a manipulation or development for the Barrett's taxonomy and IELTS taxonomy of reading comprehension categories because this ERP bar graph rejected the taxonomy presented by Barret and IELTS website.

Table 6. P-values of "following instruction" and "drawing logical inferences" sections

	Independent sample T-test P-value	
	N400	**P600**
Cz		
Fz		
Pz		0.05

*p≤0.05

Results of the analysis for the fifth research question (Is there any significant difference in item discrimination of "finding main ideas" section and "identifying the underlying concept" section based on ERP?) indicated that the fifth null-hypothesis was rejected. The independent T-test reported p=0.03, p≤0.05 from Pz channel for the N400 component and P=0.02, p≤0.05 was observed in Cz channel for P600 component. Table 7 displays the p-values.

Figure 15 ERPs of following instruction items in 3 channels
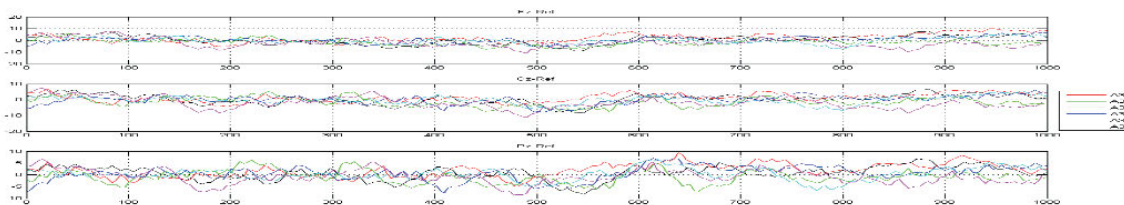
Figure 16 ERPs of "drawing logical inferences" items in 3 channels

Figure 17 ERP bar graphs of item discrimination of" following instruction" and "drawing logical inferences" sections in Pz channel of P600

Table 7. P-values of "finding main ideas" and "identifying the underlying concept" sections

	Independent sample T-test P-value	
	N400	**P600**
Cz		0.02
Fz		
Pz	0.03	

*p≤0.05

Although figures 22 and 23 show that there was a difference in item discrimination of "finding main ideas" items and "identifying relationships between main ideas" items, the item discrimination of the two sections was not significant and no significant p-value was reported.

Results of the analysis for the seventh research question (Is there any significant difference in item discrimination of "finding main ideas" section and "drawing logical inferences" section based on ERP?) indicated that the seventh null-hypothesis was rejected. Thus, there was a significant difference in item discrimination of "finding main idea" and "drawing logical inferences" sections. The independent T-test reported
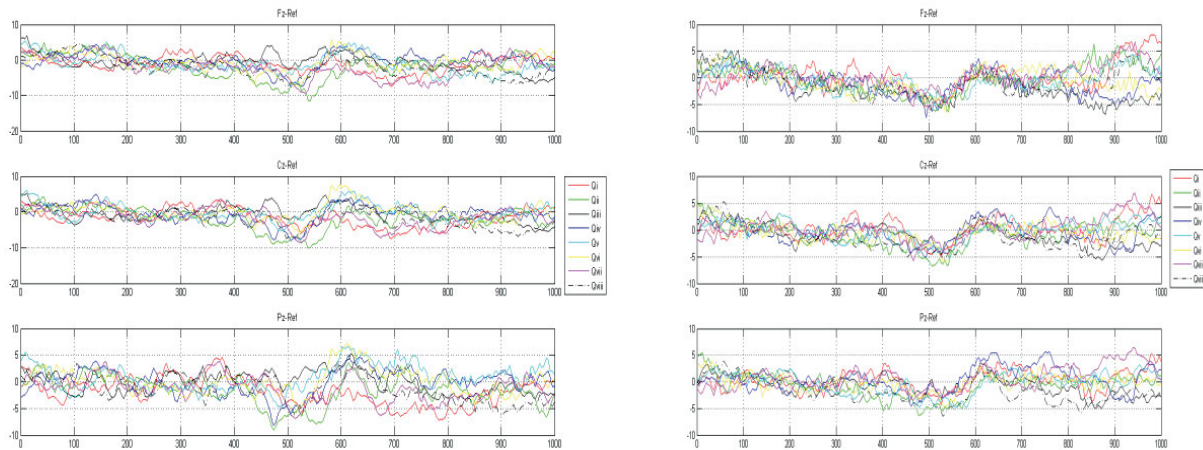
Figure 18 ERPs of "finding main idea" items in 3 channels
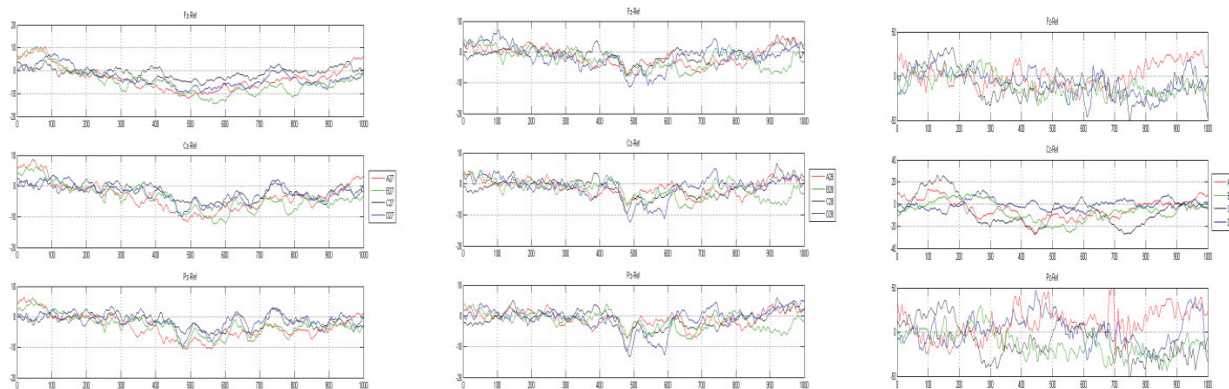
Figure 19 ERPs of "identifying underlying concept" items in 3 channels

Figures 18 and 19 show the ERPs of items of "finding main ideas" and "identifying the underlying concept" sections.

Figures 20 and 21 are the ERP bar graphs which show the significant difference between the item discrimination of "finding main idea" (1) and "identifying underlying concept" (2) sections.

Results of the analysis for the sixth research question (Is there any significant difference in item discrimination of "finding main ideas" section and "identifying relationships between the main ideas" section based on ERP?) indicated that the sixth null-hypothesis was confirmed.

p=0.04 p≤0.05 from Cz channel for the P600 component. Moreover, p=0.04 , p≤0.05 was observed from Fz channel for P600 component and P=0.01 was reported from Pz channel for P600 component.. Table 8 displays the reported P-values.

Also, figures 24 and 25 show the ERPs of items of the two sections (finding main idea and drawing logical inferences).

Figures 26, 27, and 28 are the ERP bar graphs which show the significant difference between the item discrimination of "finding main idea" (1) and "drawing logical inference" (2) sections, regarding P600 component.
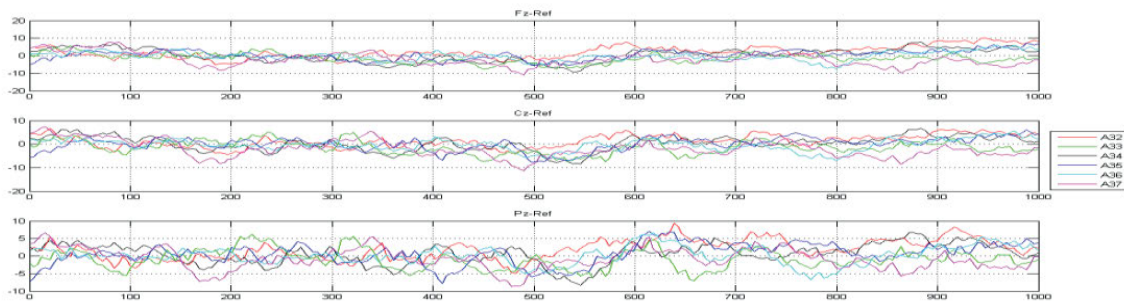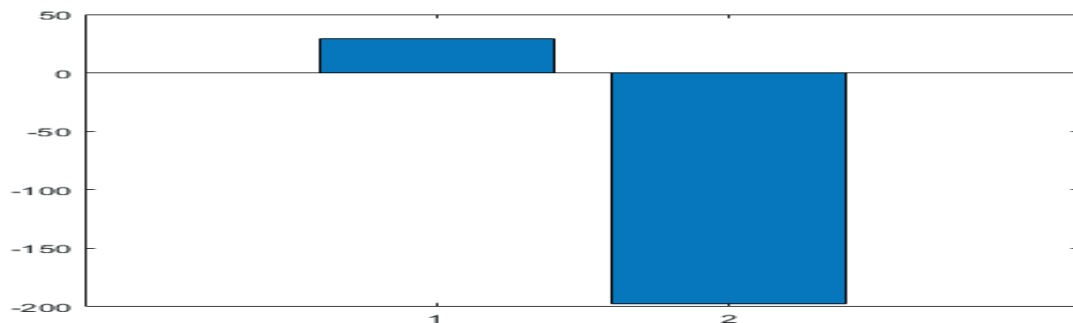
Table 8. P-values of "finding main ideas" and "drawing logical inferences" sections

Independent sample T-test P-value		
	N400	**P600**
Cz		0.04
Fz		0.04
Pz		0.01

*p≤0.05

and "drawing logical inferences" sections. The independent T-test reported p=0.05, p≤0.05 from Pz channel for the N400 component. Table9 displays the reported P-values.

Also, figure 29 shows the ERPs of "identifying the underlying concept" items.

Furthermore, figure 30 shows the ERPs of items of the "identifying relationship between main ideas".

Figure 31 is the ERP bar graph which shows the significant difference between the item discrimination of "identifying the underlying concept" (1) and "identifying
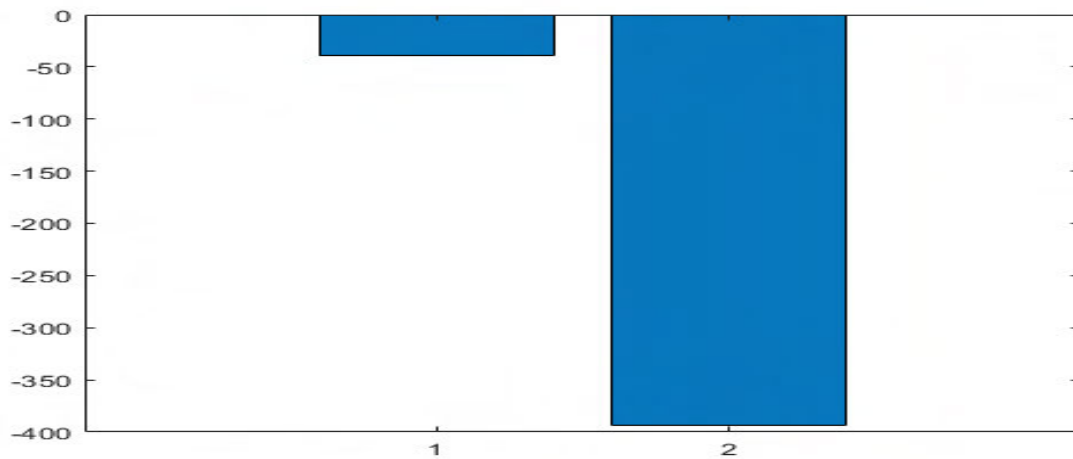
Figure 20 ERP bar graphs of item discrimination of "finding main idea" and "identifying underlying concept" sections in Pz channel of N400
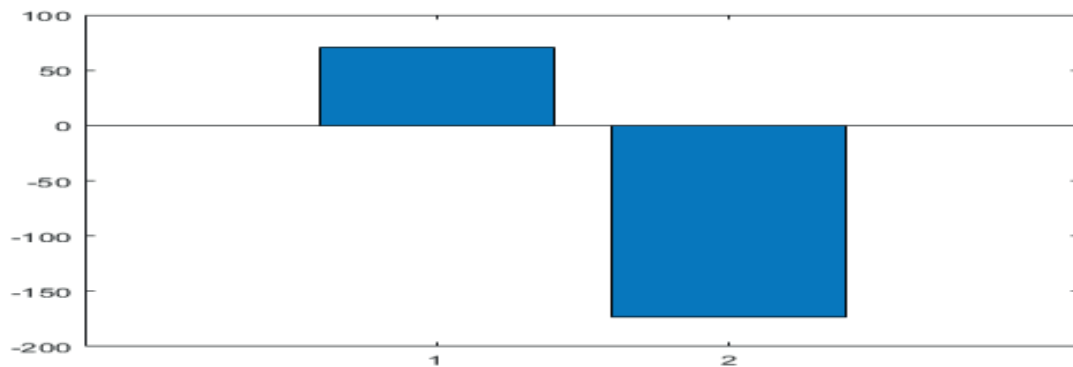
Figure 21 ERP bar graphs of item discrimination of "finding main idea" and "identifying underlying concept" sections in Cz channel of P600

Figure 27 and 28 display ERP bar graphs of item discrimination of "main idea" and "drawing logical inference" sections in Cz and Fz channels for P600

Also, figure 28 shows ERP bar graph of item discrimination of "main idea" and "drawing logical inference" sections in Pz channel for P600.

Results of the analysis for the eighth research question (Is there any significant difference in item discrimination of "identifying the underlying concept" section and "identifying relationships between the main ideas" section based on ERP?) indicated that the eighth null-hypothesis was rejected. Thus, there was a significant difference in item discrimination of "finding main idea"

Table 9. P-values of "identifying the underlying concept" and "identifying relationships between the main ideas" sections

Independent sample T-test P-value		
	N400	**P600**
Cz		
Fz		
Pz	0.05	

*p≤0.05
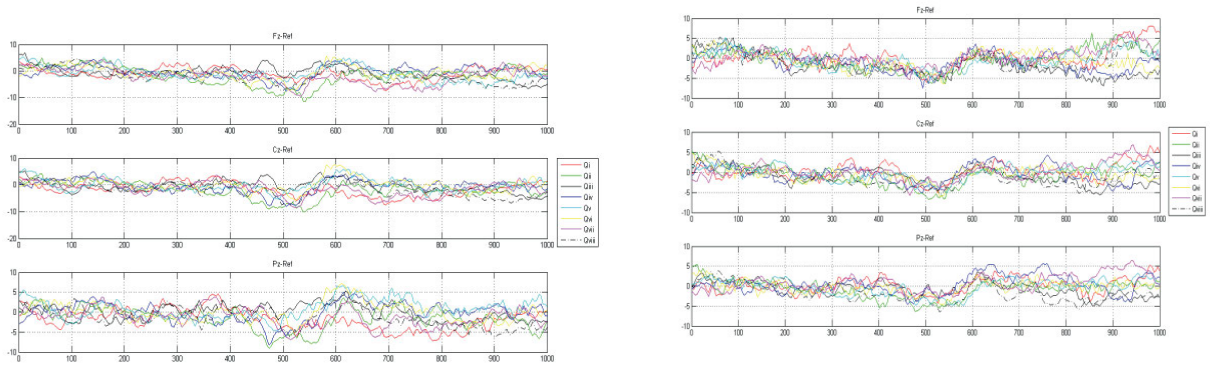
Figure 22 ERPs of "finding main idea" items in 3 channels
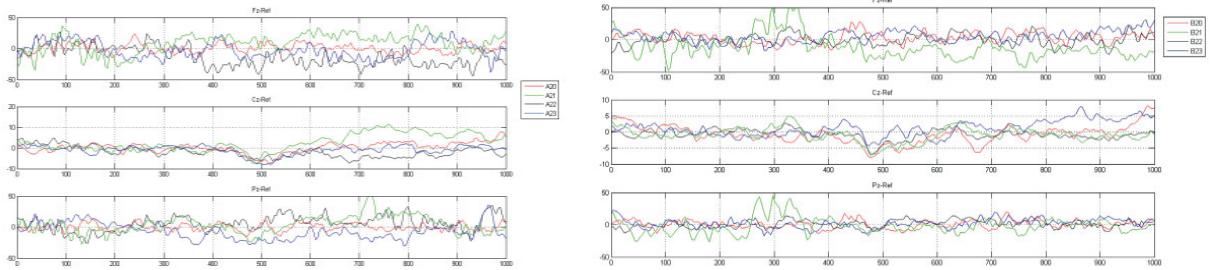
Figure 23 ERPs of "identifying relationship between main ideas" items in 3 channels
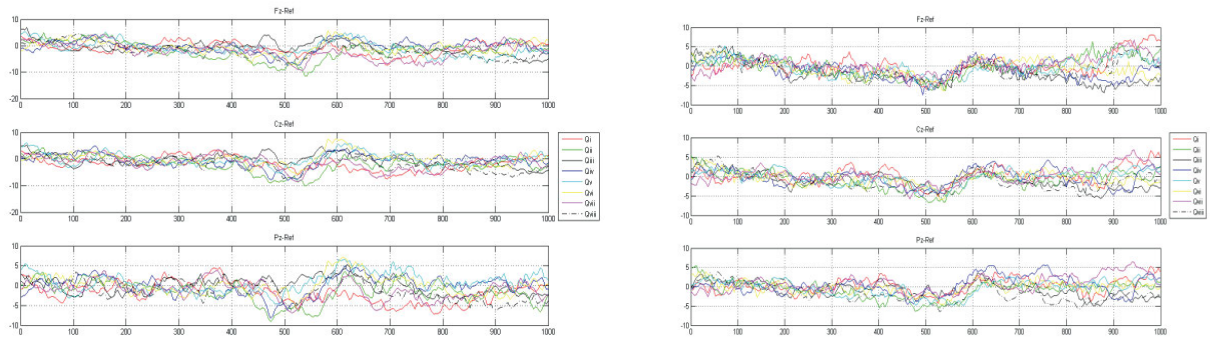
Figure 24 ERPs of "finding main idea" items in 3 channels

Figure 25 ERPs of "drawing logical inferences" items in 3 channels
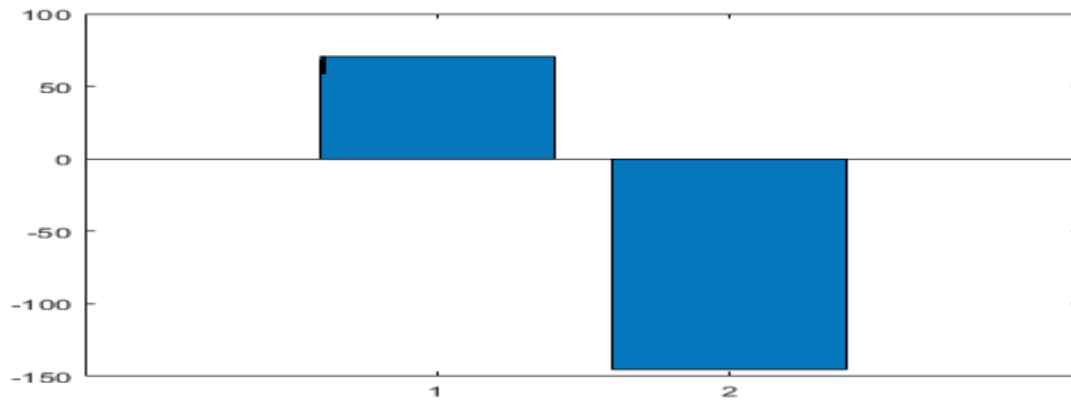
Figure 26 ERP bar graph of item discrimination of "main idea" and "drawing logical inference" sections in Cz channel for P600
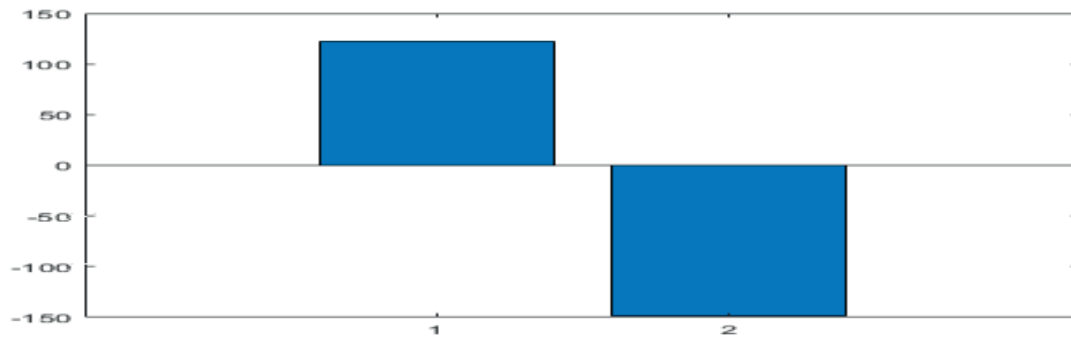

Figure 27 ERP bar graphs of item discrimination of "main idea" and "drawing logical inference" sections in Fz channel for P600
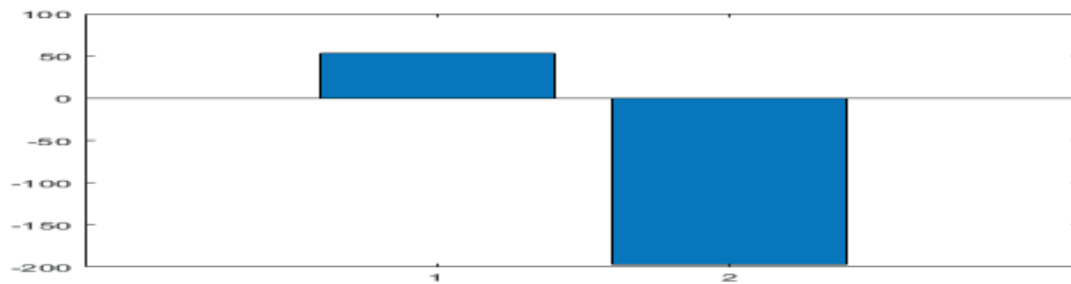
Figure 28 ERP bar graphs of item discrimination of "main idea" and "drawing logical inference" sections in Pz channel for P600

relationship between main ideas" (2) sections, regarding P600 component.

Results of the analysis for the ninth research question (Is there any significant difference in item discrimination of "identifying the underlying concept" section and "drawing logical inferences" section based on ERP?) indicated that the ninth null-hypothesis was confirmed. Although figures 32 and 33 show that there were differences in item discrimination of "identifying the underlying concept" items and "drawing logical inferences" items , the item discrimination of the two sections was not significant and no significant p-value was reported.

Results of the analysis for the tenth research question (Is there any significant difference in item discrimination of "identifying relationship between the main ideas" section and "drawing logical inferences" section based on ERP?) indicated that the tenth null-hypothesis was rejected. Thus, there was a significant difference in item discrimination of "identifying relationship between the main ideas" and "drawing logical inferences" sections. The independent T-test reported p=0.03, p≤0.05 from Fz channel for the P600 component. Moreover p=0.03, p≤0.05 was observed from Pz channel forN400 component. Table 10 shows the reported p-values from independent T-test.

Table 10. P-values of "identifying relationship between the main ideas" and "drawing logical inferences" sections

Independent sample T-test	P-value	
	N400	**P600**
Cz		
Fz		0.03
Pz	0.03	

*p≤0.05

Figures 34 and 35 display the ERPs of items of "identifying relationship between the main ideas" and "drawing logical inferences" sections.

Figure 35 shows the ERPs of "drawing logical inferences" items in 3 channels (Fz, Cz, and Pz).

Figures 36 and 37 are the ERP bar graphs which show the significant difference between the item discrimination of "identifying the relationship between paragraphs" (1) and "drawing logical inference" (2) sections, regarding N400 and P600 components.

Figure 37 presentsthe ERP bar graphs which shows the differences between the item discrimination of "identifying the relationship between paragraphs" and "drawing logical inference" sections.

3. RESULTS

Regarding the ERP recordings from 3 channels of (Fz, Cz, and Pz), the N400 and P600 components were elicited during the reading comprehension task. The N400 is elicited through semantic anomalies (Tanaka, et al., 2019) and the P600 component is elicited by semantic anomalies (Roberts, Alonso, Pliatsikas, Rothman, 2018).

The results show that there was a significant difference in item discrimination of "following instruction" and "drawing logical inferences" sections. In other words, "following instruction" section has a more difficult comprehension level in comparison with "drawing logical inference" section, regarding semantic issues.

Although "following instruction" section showed more challenge in semantics based on N400 elicitation,
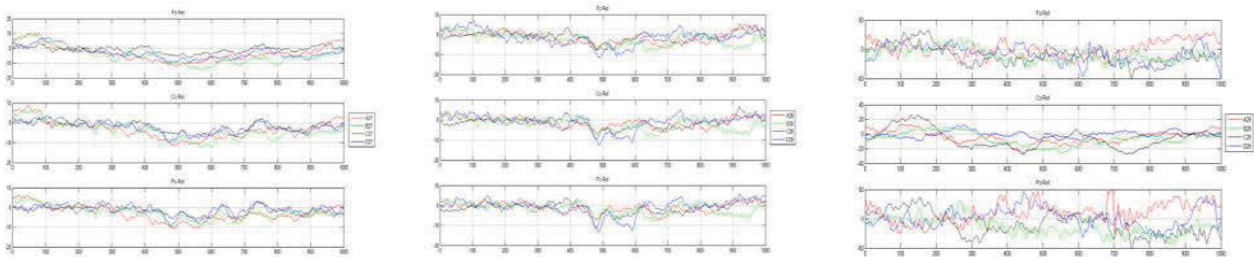
Figure 29 ERPs of "identifying underlying concept" items in 3 channels
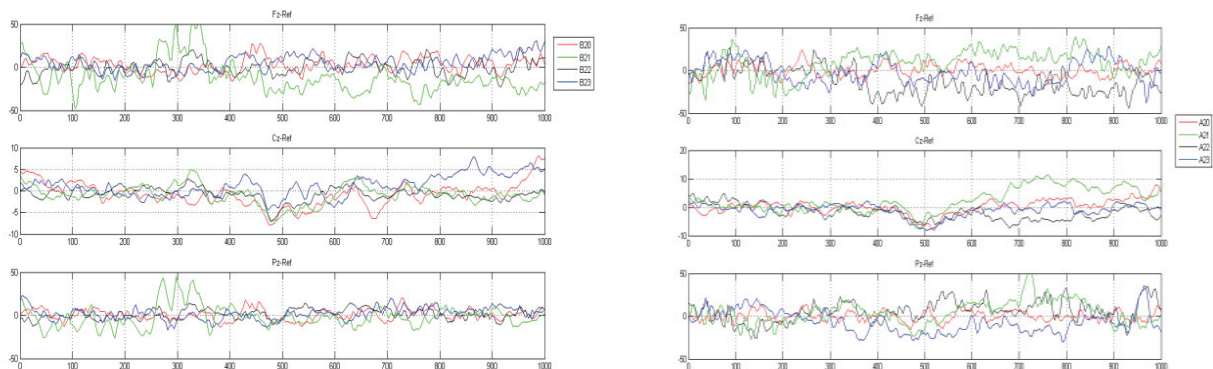
Figure 30 ERPs of "identifying relationship between main ideas" items in 3 channels
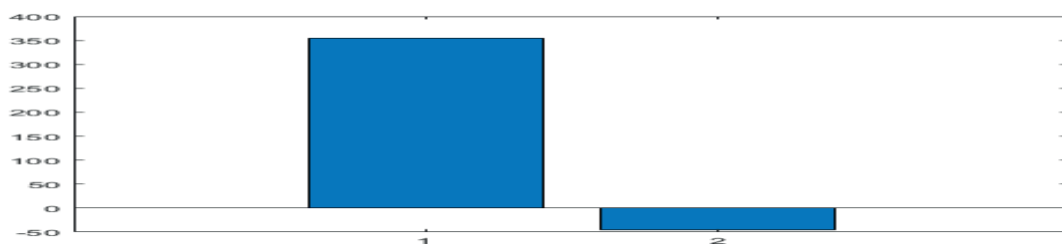
Figure 31 ERP bar graphs of item discrimination of "identifying the underlying concept" and "identifying relationship between main ideas " sections in Pz channel for N400
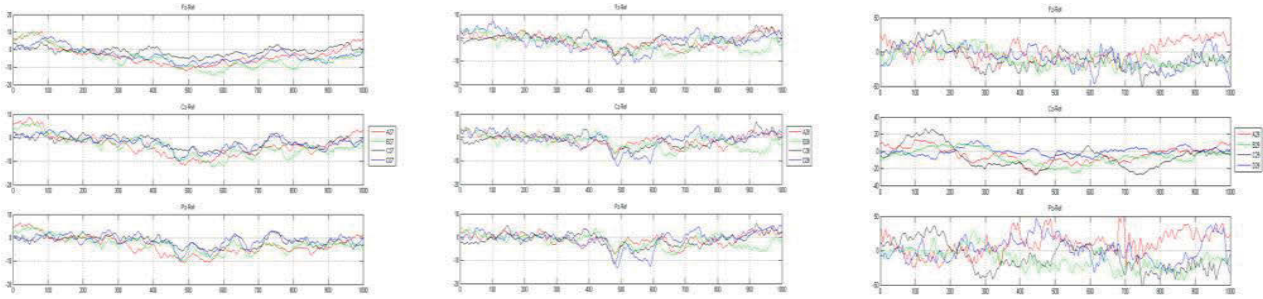
Figure 32 ERPs of "identifying underlying concept" items in 3 channels
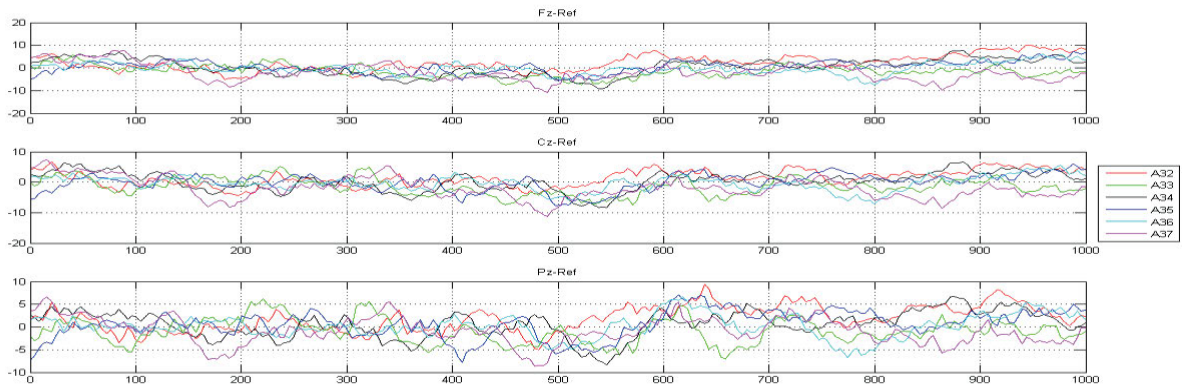
Figure 33 ERPs of "drawing logical inferences" items in 3 channels
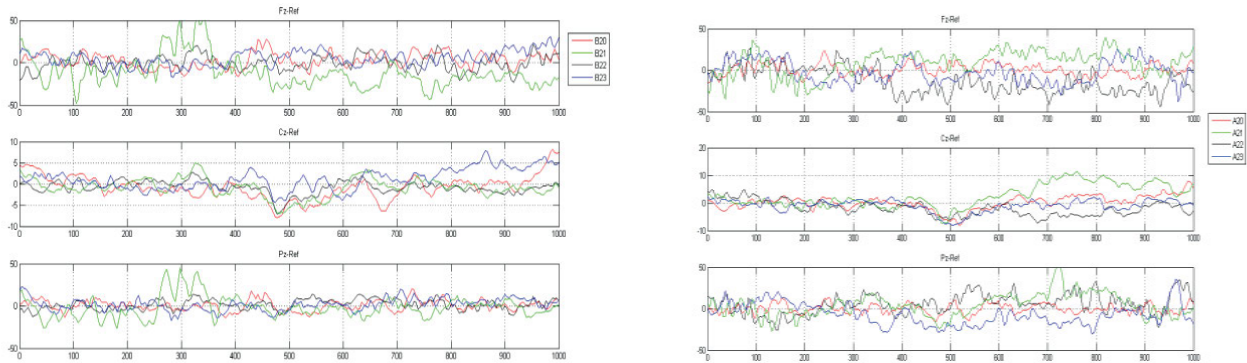
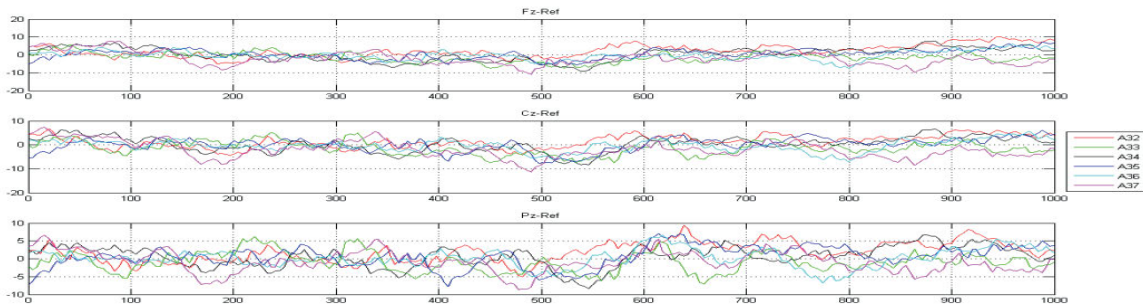**Figure 34 ERPs of "identifying relationship between main ideas" items in 3 channels**
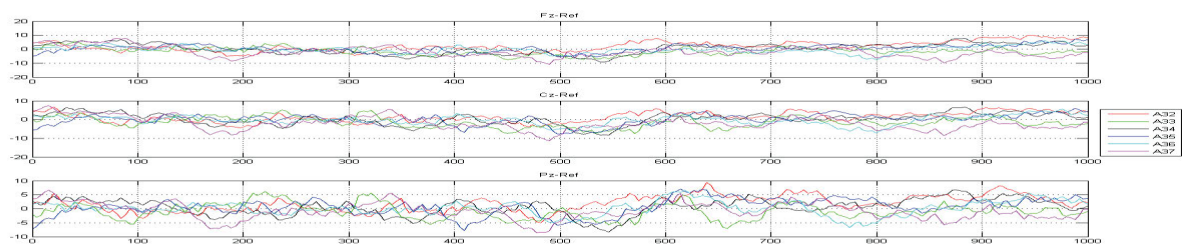
Figure 35 ERPs of "drawing logical inferences" items in 3 channels
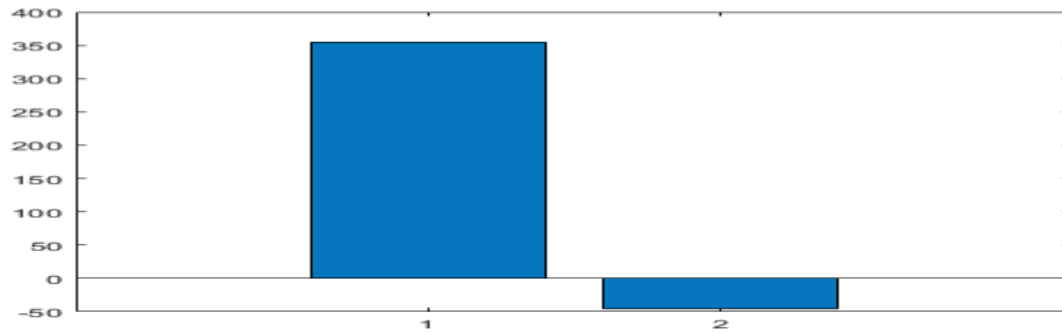
Figure 36 ERP bar graphs of item discrimination of "identifying the relationship between main ideas" and "drawing logical inferences " sections in Pz channel for N400
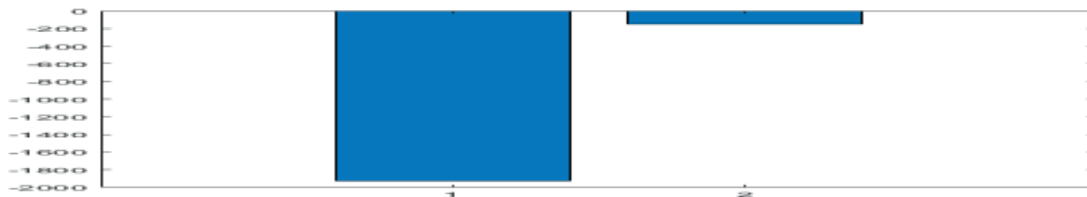
Figure 37 ERP bar graphs of item discrimination of "identifying the relationship between main ideas" and "drawing logical inferences " sections in Fz channel for P600

it did not show the same result in syntactic perspectives based on the results of P600 elicitation. ERP bar graphs showed that "drawing logical inference" has a more challenging level in syntactic level in comparison with "following instruction", regarding P600 component. This finding is in contrast with Soemer and Schiefele's (2019) study which showed that "drawing logical" section is more difficult than "following instruction" section in terms of syntax (P600) and semantics (N400).

Also, the results showed that there was a significant difference in item discrimination of "following instruction" and "identifying relationship between main ideas" sections. It can be implied that "following instruction" section has a more difficult level in comparison with "identifying relationship between main ideas" section, regarding semantics. Also, "following instruction" section syntactic level was higher than "identifying the relationship with main ideas" section. This finding is in agreement with Moore, Morton, and Price's (2012) study, which was conducted to investigate the construct validity of IELTS academic reading test. In that study, the researchers reported that "following instruction" section (True-false format) engagement level is at both global and local level and test takers should engage at different textual levels. The test takers who attempt to answer "following instruction' section items should have a good grammatical and semantic knowledge as well.

This study also reported that the engagement level of "identifying relationship between main ideas" section in format of matching information is mostly at local level and needs less mental engagement. For instance, test takers are supposed to just find a name of a scholar in order to match it with the list of ideas. Thus, "following instruction"

section is more challenging than "identifying relationship between main ideas" section.

Also, results showed that there was a significant difference in item discrimination of "identifying main ideas" and "drawing logical inference" sections. It may be implied that "main idea" section is more difficult than "drawing logical inference" section, regarding the reanalysis of sentences and grammatical issues because as discussed previously, P600 is elicited by syntactic stimulus.

This finding of the study is in agreement with Moore, Morton, and Price's (2007) study which reported that "identifying main idea" section in format of matching the list of headlines to paragraphs requires high level of understanding the semantic and pragmatic knowledge of the text. This section engages the test taker in different syntactic structures, more idiomatic phrases, and inexplicit sentences. Thus, it needs more mental activity than "drawing logical inference" section.

Also, the results showed that there was a significant difference in item discrimination of "finding main idea" and "identifying underlying concept" sections. It can be implied that the former has a more difficult comprehension level compared to the latter in terms of both semantics and syntax.

The findings are also in line with Moore, Morton, and Price's (2012) study which reported that "finding main idea" section involves both local and global level of comprehension. On the other hand, "identifying underlying concept" with multiple choice format has limitations including the potential for guessing the correct option.

Moreover, the results showed that there was a significant difference in item discrimination of "identifying

the relationship between main ideas" and "drawing logical inference" sections. Thus, the comprehension difficulty of "identifying the relationship between main ideas" is more than "drawing logical inference", regarding semantics and syntax. This finding is in contrast with Day and Park's (2005) study findings, which was conducted to investigate the difficulty level of different types of reading comprehension questions. Results of Day and Park's (2005) study showed that "identifying relationship between main ideas" section is easier than "drawing logical inference" section and students may have a difficult time in order to answer the "drawing logical inference" section.

4. CONCLUSION

According to the findings of the present study, it can be concluded that by focusing on ERP components, item discriminations of reading comprehension sections can be observed. Item discrimination, which is based on neurological model and ERP components, could facilitate the process of developing amore valid reading comprehension question taxonomy from easy to difficult. The use of reading comprehension question taxonomies is very considerable in the field of testing. The results of the current research will be a stimulus for test developers and theoreticians to change their perspectives from traditional to neurological methods of evaluating item discrimination.

Moreover, this research confirmed that Barrett's reading comprehension taxonomy and IELTS reading comprehension taxonomy have some shortcomings in categorizing the reading comprehension questions based on difficulty level. According to the results of this study, there are potentially helpful implications for test designers, IELTS book writers, IELTS test designer's team, and English teachers, in this regard. By using the results of the current study, test designers can design placement tests or other test types which are categorized from easy to difficult as far as the reading comprehension questions are concerned.

Moreover, the second group who can directly benefit from the results of the current research is the IELTS designers' team. This study suggests some modification to the IELTS reading comprehension section based on the ERP components, which illustrated its item discrimination in reading comprehension sections. As scholars including Tabullo et al. (2019) investigated the relationship between ERP components and reading comprehension, they conclude that reading comprehension scores of the participants correlated with their ERPs. Thus, IELTS designers can use the results of this study to have well-designed reading comprehension items and promote the validity of the test. Well-designed reading comprehension items are used to improve understanding of texts and can be the best predictor of test takers' reading comprehension level (Bachman, Lyle, & Palmer, 1996; Day & Park, 2005).

Finally, this study deals with some limitations; however, these limitations can also open new ideas for more research. Due to the limited sample size, the research cannot be generalized to other populations. Due to the ERP limitations which do not show reactions that are related to creativity, the format of reading comprehension questions were chosen to be matching headlines, following instructions, and multiple choice and the researcher omitted 'complete with only one word section' of the reading test. Because this section deals with creativity of the brain, so ERP cannot be measured as Raymond, Delahunty, and Seery (2018) believed that "the precise neurological substrates that underlie creativity are yet unknown".

Another limitation of this study entails the number of items in reading comprehension section of IELTS exam. Due to the time consuming process of recording EEG (each item was shown 10 times) in order to extract ERP, participants would get tired, so this fatigue leads to malfunction; consequently, the researcher decreased the number of items in reading comprehension section and for each section of that a limited number of items (19 items) were selected.

Due to the restricting criterion for participating in this research (having the IELTS band score of 6 to 8) the number of participants was10 initially. However, due to the time consuming process (around 3 hours) of the present research the number of participants decreased reliable data from 10 to 9-5 in each section of the test due to the presence or absence of artifacts. Also, There has been much research in ERP field which were conducted with few participants, for instance, a study by Shahsavar, Ghoshuni, and Talaei (2018) was conducted with 7 participants, Ghoshuni, Firoozabadi, and Khalilzade, (2012) with 10 participants, Liu, et al.,(2019) with 8 participants. Moreover, there was research by Hekmatmanesh, et al., (2019) which was conducted by 9 participants. Another limitation is that gender of test takers could not be controlled. The researcher also delimited the study in different aspects. The age range of the participants was decided to be between 26 and 36 years old and participants had to be in healthy and normal neurological status. Another delimitation of this study was the participants' band score of the IELTS exam. Participants with the band score of 6-8 were allowed to participate in this experiment, because ERP is better extracted from the average and upper intermediate candidates in order to show the item discrimination of reading comprehension items in the IELTS exam. Also, in the current study, the researcher only considered the stem and categories of the reading comprehension questions and the researcher did not investigate the format of the reading comprehension questions.

REFERENCES

Ackermann, R. (2019). *Nonreductive inference*. UK: Routledge.

Al Dahhan, N. Z., Kirby, J. R., Brien, D. C., & Munoz, D. P. (2017). *Eye movements and articulations during a letter naming speed task: Children with and without dyslexia*. Journal of Learning disabilities, 50(3), 275-285. DOI: 10.1177/0022219415618502

Alderson, J. (2000). *Assessing Reading*. Cambridge UK: Cambridge University Press.

Aldunate, N., López, V., Cornejo, C., Moënne-Loccoz, C., & Carr, D. (2019). *Analytical and Holistic Approaches Influence the Semantic*

Integration: Evidence from the N400 effect. Revista signos: estudios de lingüística, 52(100), 217-241.DOI: 10.4067/S0718-0934201 9000200217

Aleksandrov, A. A., Dmitrieva, E. S., Volnova, A. B., Knyazeva, V. M., Polyakova, N. V., Ptukha, M. A., & Gainetdinov, R. R. (2019). *Effect of alpha-NETA on auditory event related potentials in sensory gating study paradigm in mice*. Neuroscience letters, 712, 134470. DOI: 10.1016/j.neulet.2019.134470

Bachman, L. F., Lyle, F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford: Oxford University Press.

Beker, K., Jolles, D., Lorch, R. F., & van den Broek, P. (2016). *Learning from texts: activation of information from previous texts during reading*. Reading and writing, 29(6), 1161-1178. DOI: 10.1007/ s11145-016-9630-3

Bergin, T. (2018). *An Introduction to Data Analysis: Quantitative, Qualitative and Mixed Methods*. UK: SAGE Publications Limited.

Boksem, M. A., Meijman, T. F., & Lorist, M. M. (2005). *Effects of mental fatigue on attention: an ERP study*. Cognitive brain research, 25(1), 107-116. DOI: 10.1016/j.cogbrainres.2005.04.011

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2008). *An alternative perspective on "semantic P600" effects in language comprehension*. Brain research reviews, 59(1), 55-73. DOI: 10.1016/j.brainres-rev.2008.05.003

Brouwer, H., & Crocker, M. W. (2017). *On the Proper Treatment of the N400 and P600 in Language Comprehension*. Frontiers in psychology, 8. DOI: 10.3389/fpsyg.2017.01327

Brown, S., & Knight, P. (2012*). Assessing learners in higher education*. UK: Routledge.

Chang-Jun, S. H. I. (2017). *The Element of Critical Thinking in TEM-4 Reading Test-Based on a Comparative Study of the Cognitive Levels of Questions in TEM-4 and IELTS Reading Tests.* Journal of Harbin University, (9), 22.

Cheyette, S. J., & Plaut, D. C. (2017). *Modelling the N400 ERP component as transient semantic over-activation within a neural network model of word comprehension.* Cognition, 162, 153-166. DOI: 10.1016/j.cognition.2016.10.016

Courteau, É., Martignetti, L., Royle, P., & Steinhauer, K. (2019). *Eliciting ERP components for morphosyntactic agreement mismatches in grammatical sentences*. Frontiers in Psychology, 10, 1152. DOI: 10.3389/fpsyg.2019.01152

Creswell, J. W. (2002). *Educational research: Planning, conducting, and evaluating quantitative (pp. 146-166)*. Upper Saddle River, NJ: Prentice Hall

Danuwijaya, A. A. (2018). *Item analysis of reading comprehension test for post-graduate students.* English Review: Journal of English Education, 7(1), 29-40. DOI: 10.25134/erjee.v7i1.1493

Day, R. R., & Park, J. S. (2005). *Developing Reading Comprehension Questions. Reading in a foreign language*, 17(1), 60-73.

De Ayala, R. J. (2013). *The theory and practice of item response theory*. NY: Guilford Publications.

Delahunty, T., Seery, N., & Lynch, R. (2018). *Exploring the use of electroencephalography to gather objective evidence of cognitive processing during problem solving.* Journal of Science Education and Technology, 27(2), 114-130. DOI: 10.1007/s10956-017-9712-2

Draschkow, D., Heikel, E., Vo, M. L. H., Fiebach, C., & Sassenhagen, J. (2019*). A single stage of semantic processing of scenes: MVPA evidence for the identity of N300 and N400*. Psyarxiv Preprints, 1-27. DOI:10.1016/j.neuropsychologia.2018.09.016

Dudschig, C., Mackenzie, I. G., Maienborn, C., Kaup, B., & Leuthold, H. (2019). *Negation and the N400: investigating temporal aspects of negation integration using semantic and world-knowledge violations*. Language, Cognition and Neuroscience, 34(3), 309-319. DOI: 10.1080/23273798.2018.1535127

Epstein, M. L., Lazarus, A. D., Calvano, T. B., Matthews, K. A., Hendel, R. A., Epstein, B. B., & Brosvic, G. M. (2002*). Immediate feedback assessment technique promotes learning and corrects inaccurate first responses*. The Psychological Record, 52(2), 187-201. DOI: 10.1007/ BF03395423

Faust, M. (Ed.). (2015). *The handbook of the neuropsychology of language* . UK: John Wiley & Sons.

Friederici, A. D. (2004). *Event-related brain potential studies in language*. Current neurology and neuroscience reports, 4(6), 466-470. DOI: 10.1007/s11910-004-0070-0

Gebodh, N., Esmaeilpour, Z., Adair, D., Chelette, K., Dmochowski, J., Woods, A. J.,& Bikson, M. (2019). *Inherent physiological artifacts in EEG during tDCS*. NeuroImage, 185, 408-424. DOI: 10.1016/j. neuroimage.2018.10.025

Goçer, A. (2014). *The Assessment of Turkish Written Examination Questions Based on the Text in Accordance with the Barrett's Taxonomy*. Online Submission, 3, 1-16.

Gouvea,A., Philips,C., Kazanina,N.,& Poeppel, D. (2010). *The Linguistic Processes Underlying the P600*. Language and Cognitive Processes 25(2), 149-188. DOI: 10.1080/01690960902965951

Hashemi, A.,& Daneshfar, S, (2018). *A review of the IELTS Test: Focus on Validity , Reliability, and Washback*. IJELTA (Indonesian Journal of English Language Teaching and Applied linguistics. 3(1), 39-52. DOI: 10.21093/ijeltal.v3i1.123

Hekmatmanesh, A., Wu, H., Li, M., Nasrabadi, A. M., & Handroos, H. (2019). *Optimized Mother Wavelet in a Combination of Wavelet Packet with Detrended Fluctuation Analysis for Controlling a Remote Vehicle with Imagery Movement: A Brain Computer Interface Study*. In New Trends in Medical and Service Robotics (pp. 186- 195). Springer, Cham. DOI: 10.1007/978-3-030-00329-6_22

Henson, R., Dibello, L.,& Stout, B. (2018). *A Generalized Approach to Defining Item Discrimination for DCMs*. Measurement: Interdisciplinary Research and Perspectives 16(1), 18-29. DOI: 10.1080/ 15366367.2018.1436855

Herten, M., Chwilla, D. J., & Kolk, H. H. (2006). *When heuristics clash with parsing routines: ERP evidence for conflict monitoring in sentence perception*. Journal of cognitive neuroscience, 18(7), 1181-1197. DOI: 10.1162/jocn.2006.18.7.1181

Hyatt, D.,& Brooks, G.(2009). *Investigating Stakeholders' Perceptions of IELTS as an Entry Requirement for Higher Education in the U.K*. IELTS Research Reports, 10(1), 17-68. DOI: 10.1080/0309877X .2012.684043

Iralde, I.,& Allain,P. (2019).*Text memory and aging: Effect of reading perspective on recall of semantically related information*. European Review of Applied Psychology 69(3), 101-110. DOI: 10.1016/j. erap.2019.05.001

Isaacs, T.,& Trofimovich, p. (2012). *Deconstructing Comprehensibility: Identifying the Linguistic Influences on Listeners' L2 Comprehensibility Ratings*. Cambridge University Press 34(3)

Jared, D., Jouravlev, O., & Joanisse, M. F. (2017). *The effect of semantic transparency on the processing of morphologically derived words: Evidence from decision latencies and event-related potentials*. Journal of Experimental Psychology: Learning, Memory, and Cognition, 43(3), 422–450. DOI: 10.1037/xlm0000316

Kaan, E., Harris, A., Gibson.,& Holcomb, P. (2000).*The P600 as an index of syntactic integration difficulty.* Language and Cognitive Processes 15(2). DOI: 10.1080/016909600386084

Kane, M.,& Bejar, I. (2014). *Cognitive Frameworks for Assessment, Teaching, and Learning: A Validity Perspective*. Psicologia Educativa, 20, 117-123. DOI: 10.1016/j.pse.2014.11.006

Khairani H., & Shamsuddin, H. (2016).*Improving Psychological Well--being among Undergraduates: How Creativity in Learning Can Contribute?* Malaysian Journal of Communication, 35(2). DOI: 10.17576/JKMJC-2019-3502-21

Kovalenko, M. (2018). *The Validation Process in the IELTS Reading Component: Reading Requirements for Preparing International Students*. Journal of Language and Education, 4(1),63-78. DOI: *10.17323/2411-7390-2018-4-1-63-78*

Kozleski, E. B. (2017). *The uses of qualitative research: Powerful methods to inform evidence-based practice in education.* Research and Practice for Persons with Severe Disabilities, 42(1), 19-32. DOI: 10.1177/1540796916683710

Kropotov, J. D. (2010). *Quantitative EEG, event-related potentials and neurotherapy*. USA: Academic Press.

Labate, D., La Foresta, F., Mammone, N., & Morabito, F. C. (2015). *Effects of artifacts rejection on EEG complexity in Alzheimer's disease. In Advances in Neural Networks: Computational and*

Theoretical Issues (pp. 129-136). Springer, Cham. DOI: 10.1007/ 978-3-319-18164-6_13

Lewis, T. T., Yang, F. M., Jacobs, E. A., & Fitchett, G. (2012). *Racial/ ethnic differences in responses to the everyday discrimination scale: a differential item functioning analysis.* American journal of epidemiology, 175(5), 391-401. DOI: 10.1093/aje/kwr287

Lhatoo, S D., Kahane, P. & Lüders, H. O. (Eds). (2018). *Invasive Studies of the Human Epileptic Brain: Principles and Practice. Oxford*: Oxford University Press.

Loeches, M., Ouyang, G., Rausch, P., Stürmer, B., Palazova, M., Schacht, A., & Sommer, W. (2017). *Test–retest reliability of the N400 component in a sentence-reading paradigm.* Language, Cognition and Neuroscience, 32(10), 1261-1272. DOI: 10.1080/23273798.2017 .1330485

Lord, K. M. (2015). *Determining the main idea: Instructional strategies that work.* Kappa Delta Pi Record, 51(3), 138-142. DOI: 10.1080/ 00228958.2015.1056669

Luck, S. J., & Kappenman, E. S. (Eds.). (2011). *The Oxford handbook of event-related potential components.* Oxford university press.

Mahaney, T. (2018). *How and why it Works. USA: change your mind*. https://www.changeyourmind.com/ wp-content/uploads/2019/ 02/ How-It-Works-Research-latest.pdf. DOI: 10.2478/squa-2018-0005

Meyer, A. (2017). *A biomarker of anxiety in children and adolescents: A review focusing on the error-related negativity (ERN) and anxiety across development.* Developmental cognitive neuroscience, 27, 58-68. DOI: 10.1016/j.dcn.2017.08.001

Moctezuma, L. A., & Molinas, M. (2020). *EEG Channel-selection method for epileptic-seizure classification based on multi-objective optimization.* Frontiers in Neuroscience, 14, 593. DOI: 10.3389/fnins .2020.00593

Molfese, D. L., Molfese, V. J., & Pratt, N. L. (2007). *The use of eventrelated evoked potentials to predict developmental outcomes.* Infant EEG and event-related potentials, 199-225. DOI: 10.4324/ 9780203759660

Moore, T., Morton, J., & Price, S. (2012). *Construct validity in the IELTS academic reading test: A comparison of reading require-ments in IELTS test items and in university study. In Studies in language testing* (pp. 120-211). Cambridge University Press Cam-bridge.

Mori, Y. (2000). Urquhart, S., & Weir, C.(1998). *Reading in a second language: Process, product, and practice.* Studies in second language acquisition, 22(4), 590-591.

Morrow Jr, J. R., Mood, D., Disch, J., & Kang, M. (2015). *Measurement and Evaluation in Human Performance, 5E.* Human Kinetics.

Murphy, P. (2018). *How Multiracial Individuals Are Addressed in Diversity Courses in Counselor Master's Degree Programs: A Mixed Methods Content Analysis (PhD dissertation)*, Auburn University: Alabama.

Nidal, K., & Malik, A. S. (Eds.). (2014). *EEG/ERP analysis: methods and applications.* NY: Crc Press.

Nieuwland, M., Barr, D., Bartolozzi, F., Busch-Moreno, S., Donaldson, D., Ferguson, H. J., ... & Ito, A. (2019). *Dissociable effects of prediction and integration during language comprehension: Ev-idence from a large-scale study using brain potentials.* Proceedings of the Royal Society B: Biological Sciences, 375 (1791). DOI: 10.1098/ rstb.2018.0522

O'Reilly, M., & Lester, J. N. (2017). *Examining mental health through social constructionism: The language of mental health*. USA: Springer. DOI: 10.1007/978-3-319-60095-6

O'Sullivan, B. (2018*). IELTS (International English Language Testing System). The TESOL Encyclopedia of English Language Teaching*, 1-8. DOI:10.1002/9781118784235.eelt0359

O'Brien, E. J., & Cook, A. E. (2016). *Coherence threshold and the continuity of processing: The RI-Val model of comprehension.* Discourse Processes, 53(5-6), 326-338. DOI: 10.1080/0163853X .2015.1123341

Osterhout, L., & Holcomb, P. J. (1992). *Event-related brain potentials elicited by syntactic anomaly.* Journal of memory and language, 31 (6), 785-806. DOI: 10.1016/0749-596X(92)90039-Z

Patel, G., Arkin, S., Ruiz-Betancourt, D., Jamerson, E., Sanchez-Peña, J., & Javitt, D. C. (2019). F189. *Deficits and Compensation in Attentional Networks in Schizophrenia.* Biological Psychiatry, 85(10), S286-S287. DOI: 10.1016/j.biopsych.2019.03.726

Patton, M. Q. (2002). *Two decades of developments in qualitative inquiry: A personal, experiential perspective.* Qualitative social work, 1(3), 261-283. DOI:10.1177%2F1473325002001003636

Pavlov, N. (2017). *Optimizing Methods with MATLAB*. Retrieved from https://www.mathworks.com/ matlabcentral/ profile/authors/ 4710580-pavlov.

Prebianca, G. V. V. (2019). *Individual differences in main idea identification and text summarization in EFL reading comprehen-sion: an exploratory study*. Leitura, 1(39), 199-222. DOI: 10.28998/ 0103-6858.2007v1n39p199-222

Prichard, C., & Trowler, P. (2018). *Realizing qualitative research into higher education.* UK: Routledge.

Payne, B. R., Stites, M. C., & Federmeier, K. D. (2019). *Event-related brain potentials reveal how multiple aspects of semantic processing unfold across parafoveal and foveal vision during sentence reading*. Psychophysiology, 56(10), e13432. DOI: 10.1111/psyp.13432

Rabovsky, M., Hansen, S. S., & McClelland, J. L. (2017). *Neural responses decrease while performance increases with practice: A neural network model.* In CogSci. DOI: 10.1101/546325

Ramonda, K., & Sevigny, P. (2019). *Graded reader comprehension questions and item discrimination analysis. ELT journal, 73*(3), 265–274. DOI: 10.1093/elt/ccy062

Raney, G. E., & Bovee, J. C. (2016). *Reading integration in bilingual speakers. In Methods in Bilingual Reading Comprehension Research* (pp. 157-181). Springer, New York, NY. DOI: 10.1007/978-1-4939-2993-1_7

Rastelli, S. (2018). *Neurolinguistics and second language teaching: A view from the crossroads.* Second Language Research*, 34*(1), 103-123. DOI: 10.1177%2F0267658316681377

Riccio, C. A., Sullivan, J. R., & Cohen, M. J. (2010). *Neuropsychological assessment and intervention for childhood and adolescent disorders*. UK: John Wiley & Sons.

Riddle, J. P. (2019). *A Method for Moving from the Main Idea of a Biblical Text to Sermon Divisions* (Doctoral dissertation), South--eastern Baptist Theological Seminary: North Carolina.

Roberts, B. W., Kilgannon, J. H., & Trzeciak, S. (2018*). Response by Roberts et al to Letters Regarding Article,"Association between Early Hyperoxia Exposure after Resuscitation from Cardiac Arrest and Neurological Disability: Prospective Multi-center Protocol--Directed Cohort Study".* Circulation*, 138*(24), 2864-2865. DOI: 10.1161/CIRCULATIONAHA.117.032054

Romeo, T., Otgaar, H., Smeets, T., Landström, S., & Jelicic, M. (2019). *The memory-impairing effects of simulated amnesia for a mock crime.* Applied Cognitive Psychology, 33(6), 983-990. DOI: 10.1002 /acp.3508

Rouet, J. F., & Potocki, A. (2018). *From reading comprehension to document literacy: learning to search for, evaluate and integrate information across texts/De la lectura a la alfabetización docu-mental: aprender a buscar, evaluar e integrar información de diversos textos.* Infancia y Aprendizaje, 41(3), 415-446. DOI: 10.1080/02103702.2018.1480313

Salo, K. S. T., Mutanen, T. P., Vaalto, S. M., & Ilmoniemi, R. J. (2020). *EEG artifact removal in TMS studies of cortical speech areas.* Brain topography, 33(1), 1-9. DOI: 10.1007/s10548-019-00724-w

Sassenhagen, J., & Fiebach, C. J. (2019). *Finding the P3 in the P600: Decoding shared neural mechanisms of responses to syntactic violations and oddball targets*. NeuroImage, 200(15).425-436. DOI: 10.1016/j.neuroimage.2019.06.048

Shahsavar, Y., Ghoshuni, M., & Talaei, A. (2018). *Quantifying clinical improvements in patients with depression under the treatment of transcranial direct current stimulation using event related potentials.* Australasian physical & engineering sciences in medicine, 41(4), 973-983. DOI: 10.1007/s13246-018-0696-x

Scheffler, A. W. (2019). *Modelling Region-Referenced Longitudinal Functional Electroencephalography Data* (Doctoral dissertation), University of California: Los Angeles.

Sim, S. M., & Rasiah, R. I. (2006). *Relationship between item difficulty and discrimination indices in true/false-type multiple choice questions of a para-clinical multidisciplinary paper*. Annals-Academy of Medicine Singapore, *35*(2), 67. PMID: 16565756.

Small, S. L., & Hickok, G. (2016). *The neurobiology of language*. In Neurobiology of Language (pp. 3-9). USA: Academic Press. DOI: 10.1016/B978-0-12-407794-2.00001-8

Soemer, A., & Schiefele, U. (2019). *Text difficulty, topic interest, and mind wandering during reading*. Learning and Instruction, *61*, 12-22. DOI: 10.1016/j.learninstruc.2018.12.006

Steinhauer, K., Connolly, J. F., Stemmer, B., & Whitaker, H. A. (2008). *Event-related potentials in the study of language*. Concise encyclopaedia of brain and language, 91-104.

Tabullo, Á. J., Shalom, D., Sevilla, Y., Gattei, C. A., París, L., & Wainselboim, A. (2019). *Reading Comprehension and Predictability Effects on Sentence Processing: An Event-Related Potential Study*. Mind, Brain, and Education, 1-19. DOI: 10.1111/mbe. 12205

Tanaka, H., Watanabe, H., Maki, H., Sakriani, S., & Nakamura, S. (2019). *EEG-based Single Trial Detection of Language Expectation Violations in Listening to Speech*. Frontiers in computational neuroscience, 13, 15. DOI: 10.3389/fncom.2019.00015

Van den Broek, P., Beker, K., & Oudega, M. (2015). *Inference generation in text comprehension: Automatic and strategic processes in the construction of a mental representation*. Inferences during reading, 94-121. Cambridge University Press.

Veeravagu, J., Muthusamy, C., Marimuthu, R., & Michael, A. S. (2010). *Using Bloom's Taxonomy to Gauge Students' Reading Comprehension Performance/Utiliser La Taxonomie De Bloom Pour Evaluer Les Performances De Comprehension Ecrite Des Eleves*. Canadian Social Science, 6(3), 205.

Walliman, N. (2017). *Research methods: The basics*. UK: Routledge.

Ward, J. (2015). *The student's guide to cognitive neuroscience*. UK: Psychology Press.

Weir, C., Hawkey, R., Green, A., Unaldi, A., & Devi, S. (2009). *The relationship between the academic reading construct as measured by IELTS and the reading experiences of students in their first year of study at a British university*. International English Language Testing System (IELTS) Research Reports 2009: Volume 9, 97.

Westby, C. (2019). *Executive Function in Children with Language Impairment*. Word of Mouth, 30(3), 5-7. DOI: 10.1177/10483950 18821153a

Zhu, R., Pan, Y. H., Sun, L., Zhang, T., Wang, C., Ye, S., & Zhang, W. (2019). *ADAMTS18 Deficiency Affects Neuronal Morphogenesis and Reduces the Levels of Depression-like Behaviors in Mice*. Neuroscience, *399*, 53-64. DOI: 10.1016/j.neuroscience.2018.12.025

Appendix A (Full form of IELTS when appears for the first time)

Reading

Researchers like Colton don't believe it is right to measure machine creativity directly to that of humans who 'have had millennia to develop our skills'. Others, though, are fascinated by the prospect that a computer might create something as original and subtle as our best artists. So far, only one has come close. Composer David Cope invented a program called Experiments in Musical Intelligence, or EMI. Not only did EMI create compositions in Cope's style, but also that of the most revered classical composers, including Bach, Chopin and Mozart. Audiences were moved to tears, and EMI even fooled classical music experts into thinking they were hearing genuine Bach. Not everyone was impressed however. Some, such as Wiggins, have blasted Cope's work as pseudoscience, and condemned him for his deliberately vague explanation of how the software worked. Meanwhile, Douglas Hofstadter of Indiana University said EMI created replicas which still rely completely on the original artist's creative impulses. When audiences found out the truth they were often outraged with Cope, and one music lover even tried to punch him. Amid such controversy, Cope destroyed EMI's vital databases.

But why did so many people love the music, yet recoil when they discovered how it was composed? A study by computer scientist David Moffat of Glasgow Caledonian University provides a clue. He asked both expert musicians and non-experts to assess six compositions. The participants weren't told beforehand whether the tunes were composed by humans or computers, but were asked to guess, and then rate how much they liked each one. People who thought the composer was a computer tended to dislike the piece more than those who believed it was human. This was true even among the experts, who might have been expected to be more objective in their analyses.

Where does this prejudice come from? Paul Bloom of Yale University has a suggestion: he reckons part of the pleasure we get from art stems from the creative process behind the work. This can give it an 'irresistible essence', says Bloom. Meanwhile, experiments by Justin Kruger of New York University have shown that people's enjoyment of an artwork increases if they think more time and effort was needed to create it. Similarly, Colton thinks that when people experience art, they wonder what the artist might have been thinking or what the artist is trying to tell them. It seems obvious, therefore, that with computers producing art, this speculation is cut short — there's nothing to explore. But as technology becomes increasingly complex, finding those greater depths in computer art could become possible. This is precisely why Colton asks the Painting Fool to tap into online social networks for its inspiration: hopefully this way it will choos themes that will already be meaningful to us.

READING PASSAGE 3

You should spend about 20 minutes on **Questions 27–40**, which are based on Reading Passage 3 below.

Artificial artists

Can computers really create works of art?

The Painting Fool is one of a growing number of computer programs which, so their makers claim, possess creative talents. Classical music by an artificial composer has had audiences enraptured, and even tricked them into believing a human was behind the score. Artworks painted by a robot have sold for thousands of dollars and been hung in prestigious galleries. And software has been built which creates art that could not have been imagined by the programmer.

Human beings are the only species to perform sophisticated creative acts regularly. If we can break this process down into computer code, where does that leave human creativity? 'This is a question at the very core of humanity,' says Geraint Wiggins, a computational creativity researcher at Goldsmiths, University of London. 'It scares a lot of people. They are worried that it is taking something special away from what it means to be human.'

To some extent, we are all familiar with computerised art. The question is: where does the work of the artist stop and the creativity of the computer begin? Consider one of the oldest machine artists, Aaron, a robot that has had paintings exhibited in London's Tate Modern and the San Francisco Museum of Modern Art. Aaron can pick up a paintbrush and paint on canvas on its own. Impressive perhaps, but it is still little more than a tool to realise the programmer's own creative ideas.

Simon Colton, the designer of the Painting Fool, is keen to make sure his creation doesn't attract the same criticism. Unlike earlier 'artists' such as Aaron, the Painting Fool only needs minimal direction and can come up with its own concepts by going online for material. The software runs its own web searches and trawls through social media sites. It is now beginning to display a kind of imagination too, creating pictures from scratch. One of its original works is a series of fuzzy landscapes, depicting trees and sky. While some might say they have a mechanical look, Colton argues that such reactions arise from people's double standards towards software-produced and human-produced art. After all, he says, consider that the Painting Fool painted the landscapes without referring to a photo. 'If a child painted a new scene from its head, you'd say it has a certain level of imagination,' he points out. 'The same should be true of a machine.' Software bugs can also lead to unexpected results. Some of the Painting Fool's paintings of a chair came out in black and white, thanks to a technical glitch. This gives the work an eerie, ghostlike quality. Human artists like the renowned Ellsworth Kelly are lauded for limiting their colour palette — so why should computers be any different?

24

---

if allowed to fester.' For Eastwood, the central feature of boredom is a failure to put our 'attention system' into gear. This causes an inability to focus on anything, which makes time seem to go painfully slowly. What's more, your efforts to improve the situation can end up making you feel worse. 'People try to connect with the world and if they are not successful there's that frustration and irritability,' he says. Perhaps most worryingly, says Eastwood, repeatedly failing to engage attention can lead to a state where we don't know what to do any more, and no longer care.

E Eastwood's team is now trying to explore why the attention system fails. It's early days but they think that at least some of it comes down to personality. Boredom proneness has been linked with a variety of traits. People who are motivated by pleasure seem to suffer particularly badly. Other personality traits, such as curiosity, are associated with a high boredom threshold. More evidence that boredom has detrimental effects comes from studies of people who are more or less prone to boredom. It seems those who bore easily face poorer prospects in education, their career and even life in general. But of course, boredom itself cannot kill — it's the things we do to deal with it that may put us in danger. What can we do to alleviate it before it comes to that? Goetz's group has one suggestion. Working with teenagers, they found that those who 'approach' a boring situation — in other words see that it's boring and get stuck in anyway — report less boredom than those who try to avoid it by using snacks, TV or social media for distraction.

F Psychologist Francoise Wemelsfelder speculates that our over-connected lifestyles might even be a new source of boredom. 'In modern human society there is a lot of overstimulation but still a lot of problems finding meaning,' she says. So instead of seeking yet more mental stimulation, perhaps we should leave our phones alone, and use boredom to motivate us to engage with the world in a more meaningful way.

Reading

Why being bored is stimulating — and useful, too

This most common of emotions is turning out to be more interesting than we thought

A We all know how it feels — it's impossible to keep your mind on anything, time stretches out, and all the things you could do seem equally unlikely to make you feel better. But defining boredom so that it can be studied in the lab has proved difficult. For a start, it can include a lot of other mental states, such as frustration, apathy, depression and indifference. There isn't even agreement over whether boredom is always a low-energy, flat kind of emotion or whether feeling agitated and restless counts as boredom, too. In his book, *Boredom: A Lively History*, Peter Toohey at the University of Calgary, Canada, compares it to disgust — an emotion that motivates us to stay away from certain situations. 'If disgust protects humans from infection, boredom may protect them from "infectious" social situations,' he suggests.

B By asking people about their experiences of boredom, Thomas Goetz and his team at the University of Konstanz in Germany have recently identified five distinct types: indifferent, calibrating, searching, reactant and apathetic. These can be plotted on two axes — one running left to right, which measures low to high arousal, and the other from top to bottom, which measures how positive or negative the feeling is. Intriguingly, Goetz has found that while people experience all kinds of boredom, they tend to specialise in one. Of the five types, the most damaging is 'reactant' boredom with its explosive combination of high arousal and negative emotion. The most useful is what Goetz calls 'indifferent' boredom: someone isn't engaged in anything satisfying but still feels relaxed and calm. However, it remains to be seen whether there are any character traits that predict the kind of boredom each of us might be prone to.

C Psychologist Sandi Mann at the University of Central Lancashire, UK, goes further. 'All emotions are there for a reason, including boredom,' she says. Mann has found that being bored makes us more creative. 'We're all afraid of being bored but in actual fact it can lead to all kinds of amazing things,' she says. In experiments published last year, Mann found that people who had been made to feel bored by copying numbers out of the phone book for 15 minutes came up with more creative ideas about how to use a polystyrene cup than a control group. Mann concluded that a passive, boring activity is best for creativity because it allows the mind to wander. In fact, she goes so far as to suggest that we should seek out more boredom in our lives.

D Psychologist John Eastwood at York University in Toronto, Canada, isn't convinced. 'If you are in a state of mind-wandering you are not bored,' he says. 'In my view, by definition boredom is an undesirable state.' That doesn't necessarily mean that it isn't adaptive, he adds. 'Pain is adaptive — if we didn't have physical pain, bad things would happen to us. Does that mean that we should actively cause bad things would happen to us. ... if boredom has evolved to help us survive, it can still be toxic

21

Test 1

READING PASSAGE 2

You should spend about 20 minutes on **Questions 14–26**, which are based on Reading Passage 2 on pages 21 and 22.

Questions 14–19

Reading Passage 2 has six paragraphs, **A–F**.

Choose the correct heading for each paragraph from the list of headings below.

Write the correct number, **i–viii**, in boxes 14–19 on your answer sheet.

List of Headings
i The productive outcomes that may result from boredom
ii What teachers can do to prevent boredom
iii A new explanation and a new cure for boredom
iv Problems with a scientific approach to boredom
v A potential danger arising from boredom
vi Creating a system of classification for feelings of boredom
vii Age groups most affected by boredom
viii Identifying those most affected by boredom

14 Paragraph A

15 Paragraph B

16 Paragraph C

17 Paragraph D

18 Paragraph E

19 Paragraph F

Later, a Travel Planner feature was added, which allowed visitors to click and 'bookmark' places or attractions they were interested in, and then view the results on a map. The Travel Planner offered suggested routes and public transport options between the chosen locations. There were also links to accommodation in the area. By registering with the website, users could save their Travel Plan and return to it later, or print it out to take on the visit. The website also had a 'Your Words' section where anyone could submit a blog of their New Zealand travels for possible inclusion on the website.

The Tourism New Zealand website won two Webby awards for online achievement and innovation. More importantly perhaps, the growth of tourism to New Zealand was impressive. Overall tourism expenditure increased by an average of 6.9% per year between 1999 and 2004. From Britain, visits to New Zealand grew at an average annual rate of 13% between 2002 and 2006, compared to a rate of 4% overall for British visits abroad.

The website was set up to allow both individuals and travel organisations to create itineraries and travel packages to suit their own needs and interests. On the website, visitors can search for activities not solely by geographical location, but also by the particular nature of the activity. This is important as research shows that activities are the key driver of visitor satisfaction, contributing 74% to visitor satisfaction, while transport and accommodation account for the remaining 26%. The more activities that visitors undertake, the more satisfied they will be. It has also been found that visitors enjoy cultural activities most when they are interactive, such as visiting a *marae* (meeting ground) to learn about traditional Maori life. Many long-haul travellers enjoy such learning experiences, which provide them with stories to take home to their friends and family. In addition, it appears that visitors to New Zealand don't want to be 'one of the crowd' and find activities that involve only a few people more special and meaningful.

It could be argued that New Zealand is not a typical destination. New Zealand is a small country with a visitor economy composed mainly of small businesses. It is generally perceived as a safe English-speaking country with a reliable transport infrastructure. Because of the long-haul flight, most visitors stay for longer (average 20 days) and want to see as much of the country as possible on what is often seen as a once-in-a-lifetime visit. However, the underlying lessons apply anywhere — the effectiveness of a strong brand, a strategy based on unique experiences and a comprehensive and user-friendly website.

Reading

Questions 32–37

Drawing logical inferences

Complete each sentence with the correct ending, **A–G** below.

Write the correct letter, **A–G**, in boxes 32–37 on your answer sheet.

32 Simon Colton says it is important to consider the long-term view when D

33 David Cope's EMI software surprised people by A

34 Geraint Wiggins criticised Cope for not E

35 Douglas Hofstadter claimed that EMI was C

36 Audiences who had listened to EMI's music became angry after G

37 The participants in David Moffat's study had to assess music without B

List of Ideas

A generating work that was virtually indistinguishable from that of humans.

B knowing whether it was the work of humans or software.

C producing work entirely dependent on the imagination of its creator.

D comparing the artistic achievements of humans and computers.

E revealing the technical details of his program.

F persuading the public to appreciate computer art.

G discovering that it was the product of a computer program.

Questions 20–23

Underlying Concept = Aim of the writer

Look at the following people (Questions 20–23) and the list of ideas below.

Match each person with the correct idea, **A–E**.

Write the correct letter, **A–E**, in boxes 20–23 on your answer sheet.

20　Peter Toohey

21　Thomas Goetz

22　John Eastwood

23　Francoise Wemelsfelder

multiple

A20	B20
A21	B21
A22	B22
A23	B23

List of Ideas

A　The way we live today may encourage boredom.　(1) *peter t...*

B　One sort of boredom is worse than all the others.　(2) *peter t...*

C　Levels of boredom may fall in the future.　(3) *peter toohey...*

D　Trying to cope with boredom can increase its negative effects.

E　Boredom may encourage us to avoid an unpleasant experience.

Reading

Questions 8–13   following instruction

Do the following statements agree with the information given in Reading Passage 1?

In boxes 8–13 on your answer sheet, write

TRUE   if the statement agrees with the information
FALSE   if the statement contradicts the information
NOT GIVEN if there is no information on this

8   The website www.newzealand.com aimed to provide ready-made itineraries and packages for travel companies and individual tourists.

9   It was found that most visitors started searching on the website by geographical location.

10   According to research, 26% of visitor satisfaction is related to their accommodation

11   Visitors to New Zealand like to become involved in the local culture. *True*

12   Visitors like staying in small hotels in New Zealand rather than in larger ones.

13   Many visitors feel it is unlikely that they will return to New Zealand after their visit.

Test 1

READING

READING PASSAGE 1

You should spend about 20 minutes on **Questions 1–13**, which are based on Reading Passage 1 below.

Case Study: *Tourism New Zealand* website

New Zealand is a small country of four million inhabitants, a long-haul flight from all the major tourist-generating markets of the world. Tourism currently makes up 9% of the country's gross domestic product, and is the country's largest export sector. Unlike other export sectors, which make products and then sell them overseas, tourism brings its customers to New Zealand. The product is the country itself – the people, the places and the experiences. In 1999, Tourism New Zealand launched a campaign to communicate a new brand position to the world. The campaign focused on New Zealand's scenic beauty, exhilarating outdoor activities and authentic Maori culture, and it made New Zealand one of the strongest national brands in the world.

A key feature of the campaign was the website www.newzealand.com, which provided potential visitors to New Zealand with a single gateway to everything the destination had to offer. The heart of the website was a database of tourism services operators, both those based in New Zealand and those based abroad which offered tourism services to the country. Any tourism-related business could be listed by filling in a simple form. This meant that even the smallest bed and breakfast address or specialist activity provider could gain a web presence with access to an audience of long-haul visitors. In addition, because participating businesses were able to update the details they gave on a regular basis, the information provided remained accurate. And to maintain and improve standards, Tourism New Zealand organised a scheme whereby organisations appearing on the website underwent an independent evaluation against a set of agreed national standards of quality. As part of this, the effect of each business on the environment was considered.

To communicate the New Zealand experience, the site also carried features relating to famous people and places. One of the most popular was an interview with former New Zealand All Blacks rugby captain Tana Umaga. Another feature that attracted a lot of attention was an interactive journey through a number of the locations chosen for blockbuster films which had made use of New Zealand's stunning scenery as a backdrop. As the site developed, additional features were added to help independent travellers devise their own customised itineraries. To make it easier to plan motoring holidays, the site catalogued the most popular driving routes in the country, highlighting different routes according to the season and indicating distances and times.

16