Item Sequence on Test Performance: Easy Items First?

KAMAL HEIDARI SOURESHJANI
Islamic Azad University, Shahrekord Branch, Iran

Bio Data:
Kamal Heidari Soureshjani holds M.A in TEFL from Shiraz University and is a Young Researchers Club Member. He taught English courses and IELTS at different institutes in Shiraz and is presently the academic member of Azad University, Shahrekord branch. He has also published papers in journals including IJLS, TPLS, JLTR, US-China, Iranian-EFL-journal.

Abstract
Among several factors affecting the performance of testees on a test is the sequence of the test items. The present study served as an attempt to shed light on the effect of test item sequence on Iranian EFL learners' performance on a test of grammar. To achieve such a purpose, 70 language learners of English at Pooyesh Institute in Shiraz (the capital of Fars province, Iran) were given two tests (one with easy to difficult items and the other just the reverse) were designed. The results of the study revealed that the sequence of items affect foreign language learners' performance. That is, those taking easy to difficult test outperforming students taking the difficult to easy test. The study also bears a set of implications.

Keywords: test specifications, test method facets, identical items, test battery

Introduction

It is a truism that the performance of testees on a test is affected by numerous factors. These assorted factors have widely been taken into consideration and both theoretical and practical studies have been carried out with regard to these factors (e.g., Bachman & Palmer, 1981, 1982; Brutsch, 1979; Clifford, 1978, 1981; Shohamy, 1984). It is also known that every test construction includes two major points: one is the "trait," meaning the knowledge which is to be measured, and the other is the "method" referring to the procedure by which we assess the trait. In order to assess a given trait, many different methods may be used and as a result, each of them affects the trait in a different way which finally affects the performance of test takers and their scores. In fact, it could be concluded that, as Shohamy (1984) rightly stated, a test is considered as a good one if the method has little effect on the trait. To put it another way, if students' performance on a test is the result of the trait being

measured rather than the testing method, that test is considered to be a good testing tool.

In addition, test specifications have always been the concern of test developers; test specifications is mandatory and necessary for test designers, test users, test evaluators and all individuals who are, in one way or another, involved in designing and using the test for decision making. Moreover, as Alderson et al. (1995) argue, test specifications act a blueprint to the design, administration, and use of the tests contributing to their content validity and more importantly to their construct validity. Therefore, scholars in the field of testing English as a second or foreign language have appreciated the role of test methods on learners' performance on a given test. Therefore, many scholars in the field, now and then, have proposed some descriptions of the characteristics of language testing methods over the years. Carroll (1986) (cited in Bachman,1990) provided an extensive treatment of test method facets in language tests in which he described four general types of language test tasks in which he made a distinction between tasks stimulus and response characteristics. Moreover, Clark (1972) in Bachman (1990) discusses the notion of "test modalities". He proposed a dichotomy of stimulus-response in which test items, spoken or written, acts as stimulus to which learners provided response, their performance on test.

The other point that needs to be asserted here is that teachers, within their teaching process, need to obtain information about the students to assess their achievement and improve their teaching by applying the results. To use language tests for these purposes and to make decisions, the quality of the information upon which the decisions are based must be reliable and relevant. Also, scores from language tests need to be generalizable, and decision makers need to be able to make fair decisions. Thus, achieving all these purposes necessitates taking into account of all the factors which may somehow have an impact on the performance of test takers. Bachman's test method framework (Bachman, 1990) could be a good model in considering and studying such factors. It is in fact an extension of Carroll's and Clark's frameworks, including more recent views of the nature of language. In his model, Bachman describes five major categories of test facets viz. the testing environment, the test rubric, the nature of the input of the test, the nature of the expected response, and the relationship between input and response (Bachman, 1990,p. 119). Furthermore, one of the test method facets is the sequence of parts, sub-part of the rubrics of the test, which is the concern of the present study. The sequence in which the different parts of a given test are presented may influence testees' performance. Therefore, this study aims to examine the effect of the sequence of parts on test takers' performance on a test of grammar.

Background to the study

With regard to what was mentioned in the previous section, it is clear that test method factors have some sorts of bearings on test taking process and therefore must be taken into consideration by test constructors while preparing a test. Bachman (1990) stated that test method facets may influence not only the test takers' performance, but their abilities may also be affected by these test method factors. In this section of the paper, first the focus will be on the variables mentioned in the

Bachman's model (1990) which may influence the test-takers' scores. And then some general points about the various factors which may affect the test takers' performance are mentioned. Based on the framework proposed by Bachman (1990) on different categories of test method facet, there are numerous factors which have an impact on testees' performance on a test. Each of these factors involves some subcategories which must be taken into considerations in test construction and test interpretation. One of such factors is the influence of "test format". Whether test constructors use "multiple-choice", "true-false", "open-ended" or other testing formats in their tests, may influence the test takers' performance. (e.g., Alderson, 2000; Bachman & Palmer, 1996; Buck, 2001). As In'nami and Koizumi (2009) argued, since none of the test formats is perfect to function well in every context, test constructors must first look into the characteristics of each test format and then make the best selection. With regard to the test format, most of the studies focused on the two commonly-used forms: Open-ended and Multiple-choice forms. Shohamy (1984), for example, asserted that in second language reading, multiple-choice formats are easier than open-ended formats. On the contrary, Elinor (1997) found that these two formats could be of similar difficulty.

"Test rubric" is also an influential test method facet. Test rubric is concerned with the principles specifying how the testees are supposed to proceed on a given test. Language of test instructions is part of test rubric that has been a concern to test constructors. The effect of language of instruction has been studied by some researchers. Vygotsky (1969) for example, argued that there was a relationship between the language of test instructions and test-takers' performance. Jekayinfa (1987) also in another study examined the role of competence in the language of test instructions as a predictor of performance in secondary school history and concluded that there was a positive correlation between competence in the language of test instructions and academic achievement in history.

Moreover, a bulk of research has investigated the effect of test method facet on test takers' performance on a given test. Shohamy (1984), as an example, examined the effect of various test methods, namely multiple-choice and open-ended questions measuring reading comprehension. Results of her study revealed that each of the test facets produced different degrees of difficulty for subjects and that each of the variables viz. method, text, language had a significant effect on students' performance on the test of reading comprehension especially with low-level students suggesting further support of the role of language proficiency on test takers' performance on reading comprehension test. Kobayashi (2002) also addressed the effects of test method facets such as text organization and response format. He found that text organization and test format had a significant impact on the Japanese university students' performance of reading comprehension tests, and with an interaction between the two variables. His study further revealed that more proficient learners performed better in summary writing and open-ended questions with clearly organized texts. Moreover, Brown (2003) studies variations among interviewers in the ways they elicit demonstrations of communicative ability and its impact on candidate performance and, hence, raters' perceptions of candidate ability. Analysis of verbal reports suggested that differences in interviewers result in different impressions of the candidates' ability. In one interview, for instance, one

interview was considered to be more effective and willing as a communicator than in the other.

Furthermore, Jafarpur (2003) explored the relative effect of test developer on the performance of test takers using multiple choice reading comprehension tests that had no specifications. He concluded that there may be a facet of test constructor. That is, as it was revealed in his study that the characteristics of the test method restrict the learners' responses, the expected response becomes part of the test method. Lumley and O'Sullivan (2005) also investigated the role of interaction of variables such as the task topic and the gender of the person presenting the items and the gender of test takers on a tape-mediated test of speaking ability. They founded out small effects for some, but not all; interactions with limited evidence for the role of the gender of the interlocutor and a slightly more significant effect of the task topic.

The item stem can also influence test performance. Buck (1990, 1991) and Sherman (1997) examined the effects of item stem preview on test taker performance through a comparison between the mean scores of the groups who previewed item stems and those who did not. Interestingly, neither study found any significant effect for item stem preview on test taker performance or item difficulty. Buck (1991) then asserted that the lack of attention to the questions or a low level of interest in the content of the text was the cause for this negative result. Buck finally concluded that knowing the question may not motivate listening as much as test developers hope. In another study Freedle and Fellbaum (1987) suggested that the advantages of previewing item stems and answer options may be only restricted to more advanced listeners, putting less able listeners at a comparative disadvantage.

Kim (2009) investigated the impact of native and non-native language teachers' judgments on oral performance through mixed methods approach. 12 Canadian teachers (native teachers) and 12 Korean teachers (non-native teachers) were compared in terms of their internal consistency, severity, and evaluation criteria. Results of the study revealed almost acceptable levels of internal consistency and similar severity in both groups. However, the difference between groups was significant with regard to their evaluation criteria. Moreover, verbal protocol analysis indicated that native teachers were more detailed and elaborate than non-native teachers in their judgments in areas such as pronunciation, specific use of grammar and the accuracy of the information being transferred. Moreover, In'nami and Koizumi (2009) conducted a meta-analysis on the impacts of test facets, namely multiple-choice and open-ended questions on performance on L_1 reading, L_2 reading, and L_2 listening. In general, they found multiple-choice formats easier than open-ended questions in both L_1 reading and L_2 listening while no impact of test formats was found in L_2 reading.

Despite the comprehensiveness of the above-cited works, however, no research has specifically investigated the impact of different sequence of items of a test on testees' performance especially in an EFL context. In other words, no study has examined whether sequencing identical items from easy to difficult versus arranging those parts from difficult to easy has an impact on learners' performance on a grammar test.

All in all, the study tries to address the following research question:

• Does sequencing items of a grammar test from easy to difficult versus difficult to easy affect performance of Iranian foreign language learners of English on that test?

This is a significant factor as knowing that the sequence of items of any test influences the performance of testees can be a great help to present the test items in a way where testees perform the test at their best possible performance.

Method

Participants
Altogether 70 students learning English as a foreign language in one well-known English language institutes in Shiraz, Iran took part in the study. They were taking a rotation class; in the sense that almost half of the students attended classes in the morning and the other half in the evening. They ranged from 17 to 23 in age and were all taking intermediate-proficiency level courses in the institute. By the way, they, who were both male and female, were selected by utilizing a convenient sampling.

Instruments
To gather the required data, a test of 39 identical items taken from Fowler and Coe (1976) Nelson test battery corresponding to the subjects' level of proficiency (i.e. intermediate) was administered to estimate the relative difficulty of each item. Then, identical tests (each with 20 items) were designed with different sequence of items based on calculated difficulty of each item. That is, one test was designed with easy to difficult items(ED, hereafter) and the other test just the reverse, i.e. difficult to easy items (DE, hereafter). In order to control the effect of practice, the tests were administered with an interval of two weeks in between. Therefore, the pilot test and the two tests with different sequence of items incorporate the instruments of this study. It should also be noted that 30 learners took the pilot test and from which it became clear that the reliability of the test was almost .69. As to its validity, the tests were examined and confirmed by a set of related professors from Shiraz state university.

Data Collection Procedure
The participants sat twice for the designed tests. In the first administration, the ED test was distributed among them; and in the second administration, the DE version was handed in to take. Besides, the testees in each of the administrations were given enough time to take the test. The main reason for not allocating any specific time limit was to remove or at least minimize the affective factors effect (like stress).

Data Analysis
Having collected the intended data, in order to analyze them, SPSS package in general; and descriptive statistics, pair t-test, and also a wilcoxon test in particular were run. In other words, first a pair t-test was run to see if there is any significant difference between the two test performances of the same participants. Second, in the case of observing a significant difference, descriptive statistics and especially

mean differences are employed. Finally, to ensure the assumption of normality, a Wilcoxon test was run.

Results and Discussion

Having collected the data and analyzed them, now in this section of the study the results of the study are presented and discussed. As it was mentioned in the beginning of the study, the main research question the study explores is

- Does sequencing items of a grammar test from easy to difficult versus difficult to easy affect performance of Iranian foreign language learners of English on that test?

Table 1 presents the results of the paired-sample t-test for the study. As it is conspicuous from the table, since the reported p. value is less than .05 it can be concluded that there is a significant difference between the administration of the two administration tests (p = .24< .05).

Table 1
Pair Sample Test of the Administrations

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviatio	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Pai EDadministration r 1 ion DEadministration ion	.1. 23	21. 23	.. 24	.74	1.72	2.58	8 89	.. 24

Now to see in which type of test testees outperform, Table 2 revealing the paired sample descriptive statistics of the two administrations (ED and DE tests) are presented.

Table 2
Paired Samples Statistics of the Two administrations

	Mean	N	Std. Deviation	Std. Error Mean
Pair 1 ED administration	13.69	70	3.69	.24
DE administration	10.66	70	1.60	.15

By looking at the column titled as Mean, it is understood that the administration mean for the ED item test (13.69) is greater than that of the

administration in which DE item test (10.66) was used; and as a result, higher scores have been obtained. Now in order to see whether there is any significant difference between the two performances paired-sample test results are presented. The table also indicates that the standard deviation for ED is almost twice as much as that of DE (3.69 and 1.60, respectively) indicating a larger disparity among scores on ED item test administration.

The point needs to be mentioned now is that the above test performed under the assumption of normality. However, there may be some reasons for the violation of the assumption such as the small number of participants in DE and ED tests. Therefore, along with the paired samples t-test, Wilcoxon signed ranks test for which the normality assumption is not required was also run. The results, shown in table 3, were consistent with the results in table 2 further indicating the difference between performances on DE test and ED test was significant.

Table3
Wilcoxon Test Result for DE and ED tests

Z	Sig.	
DE-ED	-2.12	.033*
Asymp. Sig. (2-tailed)	*significant at P<0.05	

Overall, the findings confirm that test method facets -sequence of parts -can have an impact on learners' performance on a test of grammar and can affect students' scores on the test. It was shown that subjects taking the ED test found it easier to answer the questions; we also found that some subjects who scored high on the pilot test had a weak performance on the DE test further confirming the hypothesis that sequence of parts on a test can affect test takers' scores on a test. As to related studies, Shohamy (1984) suggested the methods that proved a bit difficult and produced lower scores may be used for norm-referenced tests in which the purpose is to maximize the inter-individual differences in scores. However, when the purpose of decision making is to estimate the level of mastery/non-mastery of students (i.e. criterion-referenced tests) tests resulting in higher scores may be recommended.

Moreover, this study once more proves the mandatory nature of test specifications in the processes of designing, administering, and using tests for decision making which involves individuals including item writers, test developers, test users, publications, and the whole society at large especially with high stake tests. This pertains to the value implications of tests which can determine in a way or another future of an individual. One may lose his/her job, for instance, or may gain a job opportunity for which he/she is not qualified as a result of taking the test. All in all, the value of test specifications or rather test method facets should be recognized in designing, administering and using a test should be recognized. In other words, when designing a test its purpose, audience, the items which it aims to test and its possible social consequences should be taken into great consideration.

Conclusion

The overall purpose of this study was to examine whether the order in which the items of a test are presented have any impact on the performance of testees. Specifically, does bringing first the easy items and then difficult ones or vice versa make any difference in the performance of testees or not.

The results of the study highlight the importance of test method facets and more importantly the need for writing test specifications in the process of designing, administering and using a test. All in all, this study showed that sequence of parts in a test of grammar can have an effect on students' performance and can differentiate among their scores. That is, learners who took the ED test outperformed those who took The DE test.

The point worthy of mentioning regarding this finding of the study is that the outperformance of language testees on the ED test in comparison to DE test may be due to a host of factors one of which is test items. To put clearly, as we noted in the review of the literature section, there are factors like content knowledge factor, affective factors (like stress), random factors (such as fatigue), and many other factors all may bear some influences on the performance of testees on the test. However, it should also be said that although the above-cited factors may influence the learners' performance, it can be claimed that these factors have a little contribution and the major factors influencing greatly their performance are factors such as language knowledge, test-method factors especially the order of test items. Besides, the order of items of a test may in itself inhibit or promote other factors like stress, motivation, etc. each affecting the test performance adversely or in a positive way. It means that ED test items may, for instance, encourage and motivate language testees to take the test with more care and interest. DE tests, in contrast, may cause demotivation, stress, and a set of other negative traits in testees and consequently, they may underperform the test.

References

Alderson, J. C. (2000). *Assessing reading.* Cambridge: Cambridge University Press.

Alderson, J. C., Clapham, C., & Wall, D. (1995). *Language test construction and evaluation.* London: CUP.

Bachman, L. (1990). *Fundamental considerations in language testing.* London: OUP.

Bachman, L. F., & Palmer, A. S. (1981). The construct validation of the FSI oral interview. *Language Learning 31*(1), 67-86.

Bachman, L. F., & Palmer, A. S. (1982). The construct validation of some component of communicative proficiency. *TESOL Quarterly 16*(4), 449-65.

Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice.* Oxford: Oxford University Press.

Brown, A. (2003). Interviewer variation and the co-construction of speaking proficiency. *Language Testing, 20*, 1-25.

Brutsch, S. M. (1979). *Convergent-discriminant validation of prospective teacher proficiency in and written French by means of the MLA cooperative language proficiency test, French direct proficiency tests for teachers (TOP and TWP), and self-ratings.* Unpublished doctoral dissertation, University of Minnesota.

Buck, G. (1990). *Testing second language listening comprehension.* Unpublished doctoral dissertation, University of Lancaster.

Buck, G. (1991). The testing of listening comprehension: An introspective study. *Language Testing, 8*, 67-91.

Buck, G. (2001). *Assessing listening.* Cambridge: Cambridge University Press.

Clifford, R. T. (1978a). Reliability and validity of language aspects contributing to oral proficiency of prospective teachers of German. *Clark*, 191-209.

Clifford, R. T. (1981). Convergent and discriminant validation of integrated and unitary language skills: the need for a research model. *Palmer et al.*, 62-70.

Elinor, S. H. (1997). *Reading native and foreign language texts and tests: The case of Arabic and Hebrew native speakers reading L1 and English FL texts and tests.* Education Resources Information Centre. Retrieved from http://www.eric.ed.gov/ERICWebPortal/search/detailmini.jsp?_nfpb=true &_&ERICExtSearch_SearchValue_0=ED412746&ERICExtSearch_SearchType_ 0=no&accno=ED412746

Fowler, S. W., & Coe, N. (1976). *Nelson English language tests.* London: Butler & Tanner Ltd.

Freedle, R., & Fellbaum, C. (1987). An exploratory study of the relative difficulty of TOEFL's listening comprehension items. In Freedle, R., & Duran, R. (Eds.), *Cognitive and linguistic analyses of test performance.* New Jersey: Ablex Publishing Corporation.

In'nami, Y., & Koizumi, R. (2009). A meta-analysis of test format effects on reading and listening test performance: Focus on multiple-choice and open-ended formats. *Language Testing, 26*, 219-244.

Jafarpur, A. (2003). Is the test constructor a facet? *Language Testing, 20*, 57-87.

Jekayinfa, A. A. (1987). Competence in the language of instruction as a predictor of performance in secondary school history. *Language Testing, 46I*, 43-65.

Kim, Y. (2009). An investigation into native and non-native teachers' judgments of oral English performance: A mixed methods approach. *Language Testing, 26*, 187-217.

Kobayashi, M. (2002). Method effects on reading comprehension test performance: Text organization and response format. *Language Testing, 19*(2), 193-220.

Lumley, T., & O'Sullivan, B. (2005). The effect of test-taker gender, audience and topic on task performance in tape-mediated assessment of speaking. *Language Testing, 22*, 415-437.

Sherman, J. (1997). The effect of question preview in listening comprehension tests. *Language Testing, 14*, 185-213.

Shohamy, E. (1984). Does the testing method make a difference? The case of reading comprehension. *Language Testing,* 147-61.

Vygotsky, L. S. (1969). *Thought and language.* Cambridge, MA: MIT Press.

Appendix 1. Pilot Test

Choose the item that best fits the sentence.
Example: can this camera …… good photos?
A: make   B: to take   **C: take**   D: to take
As you can see the best choice is __take__ which is choice __C.__
1 – Call me when you get………
A: to home   B: at home   C: in home   D: home
2 – When………. a game of football?
A: had you last   B: did you last have
C: last had you   D: did you have last
3 - I can't find the book………..
A: nowhere   B: everywhere   C: anywhere D: somewhere
4- A: Do you like cakes?   B: I don't like ……..
A: all them   B: them all   C: every   D: everyone
5 – I …….. meet my grandfather every day.
A: used to   B: wanted   C: liked   D: needed
6 – You don't need your hat.
A: Put it off!   B: Take it off!   C: Put off it! D: Take off it!
7 – Her dress is…………… yours.
A: the same as   B: the same that   C: similar than D: similar that
8 – I haven't got a chair………..
A: to sit   B: sitting   C: to sit on   D: for sitting
9 – A typist's job isn't always easy………..
A: for to do   B: for doing   C: of doing   D: to do
10 – Mina likes going to the theater and………..
A: so do I   B: so go I   C: so I like   D: so I am
11 – Mary brought us……….. cheese.
A: enough of   B: too many   C: some more   D: a lot
12 - ………… like ice cream.
A: Every children   B: Every child   C: All of children   D: All children
13 – The girl ……….. house he visited was Mina.

A: of which   B: of whom   C: which   D: whose

14 – He had had ………. hard time last night that he slept all the morning.

A: such a   B: such   C: such an   D: very

15 – When you go abroad, do you ………take your passport?

A: have to   B: ought to   C: need   D: must

16 – A: Sarah never speaks Japanese.   B: …………Reza.

A: So does   B: Nor does   C: So is   D: Neither is

17 – Don't leave your shoes on the table.

A: Put them off!   B: Take them off!

C: Put off them!   D: Take off them!

18 – A: ……… to drink cola?   B: Sure, thanks.

A: May you   B: Would you like   C: Should you   D: Can you

19 – Tell Homa come back………tomorrow, Ali Shouted at her father!

A: home   B: in home   C: to home   D: at home

20 – To drive a car safely it is essential……….. good brakes.

A: with having   B: to have   C: having   D: have

21- I looked everywhere but I could find my book……………

A: nowhere   B: anywhere   C: everywhere   D: somewhere

22 – She cut the cloth with …………. scissors.

A: a couple   B: a pair of   C: two   D: a

23 – I didn't know where………..

A: the cinema to be   B: was the cinema

C: is the cinema   D: the cinema was

24 – It is the…………….. interesting film I've ever seen.

A: more   B: most   C: less   D: little

25 – We haven't got a record player. Let's………… Ali's recorder.

A: borrow   B: to borrow   C: borrows   D: borrowing

26 – I was very hungry but I didn't know where a restaurant………..nearby.

A: was   B: is   C: to be   D: be

27 - Let me…………… you what happened to me yesterday.

A: to tell   B: tells   C: tell   D: telling

28 – That's the man………. killed my cat.

A: whose dog   B: which dog   C: the dog of whom   D: whom dog

29 – It was ………big job that he had to ask for help.

A: a so   B: such a   C: a very   D: such an

30 - ……… I repeat the question?

A: Shall   B: Will   C: Do you want   D: Would

31– Mummy, can I have …………. butter ?

A: some more   B: much   D: enough of   D: a lot

32 – To travel from Iran to Foreign countries you ………… a passport.

A: mustn't have   B: haven't got   C: needn't   D: don't need

33 - It is very important……………. brakes when we want to drive a car.

A: to check   B: checking   C: with checking   D: check

34– It is usually noisy here………… Saturday morning.

A: on   B: at   C: in   D: of

35 – Amir always arrives at school……… 8:30'.

A: at   B: on   C: of   D: in

36 – When………….. English?

A: has he begun to study   C: has he begun study

B: did he begin to study   D: did he begin study

37 – Ali likes walking in the country and …………

A: also does John   B: so does John   C: John likes also   D: so John likes

38 – Who was the first person……….today?

A: spoke to you   B: you spoke to

C: you spoke   D: whom you spoke

39 – Reza ………………. go to cinema, but now, he goes everyday!

A: use to   B: didn't use to   C: used to   D: is used to

Appendix 2. Test with Easy to Difficult Items

Please answer the questions in the allotted time: **(25 minutes)**

Choose the item that best fits the sentence.

Example: If you…………. up, you will arrive late.

A: didn't hurry   **B: don't hurry**   C: won't hurry   D: hurry

As you can see the best choice is **don't hurry** which is choice **B.**

1 - I can't find the book………..

A: nowhere   B: everywhere   C: anywhere   D: somewhere

2 – I …….. meet my grandfather every day but now I can't because I am too busy.

A: used to   B: wanted   C: liked   D: needed

3 – Mina likes going to the theater and………..

A: so do I   B: so go I   C: so I like   D: so I am

4 – A: ……… to drink cola?   B: Sure, thanks.

A: May you   B: Would you like   C: Should you   D: Can you

5– It is the…………….. interesting film I've ever seen.

A: more   B: most   C: less   D: little

6 – Amir always arrives at school……… 8:30'.

A: at   B: on   C: of   D: in

7– Call me when you get………

A: to home   B: at home   C: in home   D: home

8– Her dress is…………… yours.

A: the same as   B: the same that   C: similar than   D: similar that

9 – When you go abroad, do you ……….take your passport?

A: have to   B: ought to   C: need   D: must

10 – You don't need your hat.

A: Put it off!   B: Take it off!   C: Put off it!   D: Take off it!

11– He had had ………. hard time last night that he slept all the morning.

A: such a   B: such   C: such an   D: very

12– The girl ……….. house he visited was Mina.

A: of which   B: of whom   C: which   D: whose

13 – I didn't know where………..

A: the cinema to be   B: was the cinema

C: is the cinema   D: the cinema was

14 - Let me…………… you what happened to me yesterday.
A: to tell   B: tells   C: tell   D: telling
15- A: Do you like cakes?   B: I don't like ……..
A: all them   B: them all   C: every   D: everyone
16– Arash brought us……….. cheese.
A: enough of   B: too many   C: some more   D: a lot
17 - It is very important…………… brakes when we want to drive a car.
A: to check   B: checking   C: with checking   D: check
18- I can't think of …………… furniture in my house that I bought new.
A: a   B: a piece of   C: piece of   D: one
19 - I haven't got a chair………..
A: to sit   B: sitting   C: to sit on   D: for sitting
20 – When………. a game of football?
A: had you last   B: did you last have
C: last had you   D: did you have last

Appendix 3. Test with Difficult to Easy Items

Please answer the questions in the allotted time: (**25 minutes**)
Choose the item that best fits the sentence.
Example: If you…………. up, you will arrive late.
A: didn't hurry   **B: don't hurry**   C: won't hurry   D: hurry
As you can see the best choice is **don't hurry** which is choice **B.**
1 – Don't leave your shoes on the table.
     A: Put them off!   B: Take them off!
     C: Put off them!   D: Take off them!
2– A: Sarah never speaks Japanese.   B: …………Reza.
A: So does   B: Nor does   C: So is   D: Neither is
3 - ………… like ice cream.
A: Every children   B: Every child   C: All of children   D: All children
4 – A typist's job isn't always easy………..
A: for to do   B: for doing   C: of doing   D: to do
5 – She cut the cloth with …………. scissors.
A: a couple   B: a pair of   C: two   D: a
6 – To travel from Iran to Foreign countries you ………… a passport.
A: mustn't have   B: haven't got   C: needn't   D: don't need
7– That's the man………. killed my cat.
A: whose dog   B: which dog   C: the dog of whom   D: whom dog
8– It was ………big job that he had to ask for help.
A: a so   B: such a   C: a very   D: such an
9 – Mummy, can I have …………. butter ?
A: some more   B: much   D: enough of   D: a lot
10 - ……… I repeat the question?
A: Shall   B: Will   C: Do you want   D: Would
11 – Tell Homa come back………tomorrow, Ali Shouted at her father!
A: home   B: in home   C: to home   D: at home

12 – We haven't got a record player. Let's………… Ali's recorder.
A: borrow   B: to borrow   C: borrows   D: borrowing
13 – I was very hungry but I didn't know where a restaurant……… nearby.
A: was   B: is   C: to be   D: be
14– When………….. English?
A: has he begun to study   C: has he begun study
B: did he begin to study   D: did he begin study
15 – Reza ………………. go to cinema, but now, he goes everyday!
A: use to   B: didn't use to   C: used to   D: is used to
16 - I like English………than other languages.
A: most   B: more   C: least   D: little
17– It is usually noisy here………… Saturday morning.
A: on   B: at   C: in   D: of
18 – To drive a car safely it is essential……….. good brakes.
A: with having   B: to have   C: having   D: have
19– He's already about………. his father.
A: so tall than   B: as tall than   C: as tall as   D: so tall as
20 - I looked everywhere but I could find my book……………
A: nowhere   B: anywhere   C: everywhere   D: somewhere