



Investigating the Influence of Proficiency and Gender on the Use of Selected Test-Wisness Strategies in Higher Education

A. Majid Hayati

Dept. of English, College of Lit. & Humanities

Shahid Chamran University, Iran

E-mail: majid_hayati@yahoo.com

A. Nick Ghogh

Dept. of English, College of Lit. & Humanities

Shahid Chamran University, Iran

E-mail: abdolhaynick@yahoo.com

Abstract

This study investigated the relative frequency of seven TW strategies among Iranian EFL students to find out probable relationship(s) between test-wisness, proficiency and gender. To do so, out of 138 participants, a total number of 80 undergraduate EFL students from Shahid Chamran University were chosen and divided on the basis of their proficiency scores into four cells of 20 individuals (High/Low Proficient Male/Female). All of them were required to sit for the second exam after two weeks periods from the first experiment. They took a test of TW comprised of 50 items in which seven selected TW strategies were incorporated. Data analysis showed significant differences between the two groups, that is, high proficient groups outperformed their low proficient counterparts. In other words, more proficient students were more test-wise. Therefore, as proved by the results of the study, TW and proficiency of students are positively correlated. Another finding suggested that there is no significant difference between the two genders concerning TW strategies.

Keywords: Test-wisness, Validity, Language testing, Multiple-choice items, Standardized tests

1. Introduction

Testing is an indispensable element of our daily activities. Whether we realize it or not we are involved in hypothesis testing about every cognitive and affective effort we make. The motive behind testing is to make decision about a course of action, depending upon the significance to be attached to them. Language testing being a challenging field of testing in general is no exception. There have been recent developments in the field of language testing and applied linguistics preoccupying the minds of language testing experts during last decades. Recent theoretical advances mainly concerned with three issues, language proficiency, the effect of test-formats, and the test-takers' characteristics. Also there have been methodological advances which dealt with issues in psychometric testing, statistical analysis, and test-takers strategies in completing test tasks (Bachman, 1991). The third methodological advance, test-wisness (hereafter TW), has been focus by experts in the field to investigate the process or strategies that test-takers employ on a variety of test formats. Miller, Fuqua and Fagley (1990, p.204) state that "individual differences in TW would tend to decrease the validity of the test scores because scores would reflect test-taking skills in addition to knowledge of the subject matter being tested", hence reducing the content validity of the test. In the same line, the present study intends to investigate how Iranian EFL students manifest TW strategies in classroom settings. In particular, this study seeks to identify the relative frequency of seven TW strategies among high and low proficient male and female Iranian EFL students.

2. Review of the Literature

Test-wisness has widely been defined as an individual's ability to improve his or her test score by recognizing and utilizing cues in the test items, format or testing situation (Houston, 2005). Test-wisness is largely independent of the subject matter for which the items are supposed to measure (Millman, Bishop & Ebel, 1965). In other words, TW is the

ability to use special strategies to select the correct response in multiple choice tests, without necessarily knowing the content or skill that is being measured. Specifically, multiple-choice tests are more susceptible to test-wiseness cues, so it was expected that there will be a stronger relationship between test-taking skills and multiple-choice test performance than with constructed response test performance (Edwards, 2003). This may be logically expected, since multiple-choice items contain numerous components (e.g. a stem and four alternatives) where TW cues may occur. Secondly, since multiple-choice items are usually the most difficult to construct, they may be readily susceptible to all types of shortcomings including TW.

TW is therefore a source of test invalidity. Examinees vary in their knowledge and use of TW principles and unless this is controlled some candidates will have an unfair advantage over others (Allan, 1992). Scruggs and Mastropieri (1992) also found out that all students are not equal when using test-taking skills effectively. Moreover, Miller, Fuqua and Fagley (1990) state "individual differences in TW would tend to decrease the validity of the test scores because scores would reflect test-taking skills in addition to knowledge of the subject matter being tested, reducing content validity" (p. 204). This implies the conception that language test scores cannot be interpreted simplistically as an indicator of the particular language ability we want to measure, because these scores are affected to some extent by the characteristics and context of the test tasks, the characteristics of the test-taker and the TW strategies. Therefore, these issues endanger both the validity and reliability of the tests which in turn influence the process of decision making as the paramount goal of measurement, whereas the "Orthodoxy in language testing is the maintenance of balance between reliability and validity" (Davies, 2003, p.365; see also Brown, 1996; Gronlund and Linn, 1990).

The next step in examining TW is to determine where and when it may appear. The "where" part of this question refers to what types of tests and test items are susceptible to TW effects. As one may suspect, teacher-made tests frequently exhibit TW cues. The reasons for this occurrence should be sought in teachers. First of all, compared to the professional test-constructor, most teachers are relatively naïve of the TW principles. Secondly, teachers usually do not have the need, desire, or knowledge to determine such factors as test reliability, validity, item difficulty, and item discrimination.

Although one would guess that standardized tests would be relatively immune to TW, research indicates that this is not the case. Diamond, Ayler, Fishman and Green (1976) believed that some of students do poorly on standardized tests, because they do not know how to take tests. One implication could be that other students have somehow learned to figure out the answer to multiple-choice items about which they have no knowledge, by using skills which are usually under the rubric "test-wiseness". Bracely (2001, as cited in Deerman et al. 2008, p.62) states that "Standardized tests are unfair not only in terms of non-accommodation for diverse learning styles, but they do not take into account subgroups, such as racial and ethnic minorities, students with disabilities, students with limited language proficiency, and students from low socioeconomic groups" (see also Shavelson, Webb & Burstein 1986). Moreover, a number of researchers have shown that training in various TW skills will improve test scores on subsequent standardized testing (Bergman, 1980; Oakland, 1972; Petty & Harrell, 1977; Shuller, 1979). This fact that performance on these tests was affected by "general" aspects of TW, indicates that TW is a pervasive factor affecting different tests in a variety of ways.

Returning to the original question of the manifestation of TW, the "when" aspect refers to the age of the test-wise individual. Surprisingly enough, TW spans the broadest age range possible. At one end of the spectrum is the preschool child, and at the other end adults. Many of the researchers (Bergman, 1980; Diamond, et al. 1976; Oakland, 1972; Shuller, 1979) postulated that TW could be taught to preschool and elementary subjects. Moving up the age continuum, university students and adults have exhibited TW skills in a number of studies (Bajtelsmit, 1977; Callenbach, 1973; Diamond & Evans, 1972; Morse, 1998). Therefore it may be concluded, that TW abilities are characteristics of all age groups.

Literature in the EFL/ ESL fields produced no studies dealing directly with TW, and a few which dealt with the concept indirectly (Allan, 1992; Vattanapath & Jaiprayoon, 1999). However, there have been a lot of studies in the wider domain of psychological education (Millman et al., 1965; Morse, 1998; Sarnacki, 1979 among others). Diamond and Evans (1972), Diamond et al. (1976), and Morse (1998) investigated the relative difficulty of several TW strategies and the results indicated somewhat similar increasing difficulty order of the strategies as follows: (a) grammatical cue, (b) longer option, (c) absurd alternative, (d) item giveaway, (d) specific determiner, (e) alliterative association. Then it was observed that skills such as grammatical cue, longer option, and absurd option, were statistically and significantly easier than specific determiner and alliterative association.

In a recent study, Allan (1992, p.102) reported an inventory of 33 TW principles under ten most generalizable strategies for multiple-choice tests. He incorporated four TW subscales in the instrument and trialed on several groups of ESL students. The findings indicated that students are somewhat test-wise and use different TW strategies and these strategies are not equally easy to employ. That is, the four strategies were of the following order of increasing difficulty: (a) similar option, (b) grammatical cue, (c) item giveaway, and (d) stem option.

Literature on TW suggests various studies done in the domain of TW correlates. For instance, research has demonstrated a positive relationship between TW and intelligence (Diamond & Evan, 1972), but not as strong as may

have been expected. In a similar study Dunn and Goldstein (1959) obtained correlations of zero between intelligence and TW abilities. As a result, it was concluded that the ability to pick up TW cues may be demonstrated at all levels of intelligence.

A rather different variable that would be expected to correlate with TW in a positive manner is verbal achievement. Recognition of most TW cues is dependent upon skills such as knowledge of grammar, vocabulary, and sentence structure (Sarnacki, 1979). A number of studies found a significant positive correlation between TW and verbal skills (Diamond & Evans, 1972; Rowley, 1974).

Another possible variable to correlate with TW is test anxiety. Logical considerations would suggest a negative correlation between them. A number of studies proved this predication of correlation (Allen, 1972; Beidel & Turner, 1988; Koons & Vasey, 2000; Petty & Harrell, 1977).

Another variable to mention is race. In one study conducted by Edwards (2003), there was support for the hypothesis stating that test-taking skills were partially related to test performance. According to obtained results, there were subgroup differences on test-taking skills, and test-taking skills partially mediated the relationship between race and test performance. However, the strength of mediation was not sufficient to reduce subgroup differences on the multiple-choice test to the levels of subgroup differences observed on the constructed response test. Also, Houston (2005), in her research, found out that there were no significant differences between whites and African Americans on the pre-test Learning measure and the pre-test Behavior measure. While overall, training had a positive impact on subjects' abilities to identify the test-wiseness cues on the Learning measure with subjects showing a significant improvement, subjects showed only marginal improvements on the Behavior measure. In addition, rather than diminishing group differences, test-wiseness training appeared to have no significant race by training effect on the Learning measure and appeared to exacerbate the differences between whites and African Americans on the Behavior measure.

Mohamed, Gregory and Austin (2006) conducted a study to compare the test-taking skills and abilities (test-wiseness) of Canadian senior-level pharmacy students with those of international pharmacy graduates. For this purpose a 20-item test-wiseness questionnaire was developed and administered to 102 participants. According to the results of their study, mean test-wiseness scores indicated significant differences in performance between senior level pharmacy students and international pharmacy graduates. Test-wiseness deficiencies of international pharmacy graduates were particularly severe in domains requiring discerning use of English language.

In general, there have been few studies investigating the role of gender, as another aspect of TW correlation, in language tests. Slakter, Koehler, and Hampton (1970), examining the relationship between TW, grade level, and gender, found that although gender was not related to TW abilities, grade level was. As the grade level increased so did individual's performance on TW scales. Along the same line, Lo and Slakter (1973) examined the relationship between risk taking on objective examinations, TW, gender, and prior experience on examinations among Chinese and American students. They evidenced no relationship between gender and risk taking, test experience and risk taking, and TW and gender. However Chinese students were consistently lower on mean TW score than their American counterparts. Mottalebzade (1993) found little difference between male and female performance on grammar and vocabulary. However, on reading comprehension and cloze, male students significantly outperformed female ones. Contrary to this finding, Farhady (1982) found no significant difference between male and female students on language tests except listening comprehension in which male participants had a better performance over the female ones.

3. Research Questions

The present study intends to investigate how Iranian EFL students manifest TW strategies in classroom settings. In particular, this study seeks to identify the relative frequency of seven TW strategies among high and low proficient male and female Iranian EFL students. These strategies will be incorporated in a test of TW, which is developed by the researchers and validated by three experts in the field of language testing and teaching. In order to come up with conclusive and comprehensive results and to achieve the research purposes, the following questions are put forward.

- 1) Do Iranian EFL students possess TW strategies? If so, which ones are most prevalent?
- 2) How far and in what ways is TW linked with language proficiency?
- 3) Is there any relationship between TW and gender?

4. Research Hypotheses

To have a wholesome speculation and to have systematic investigation of the research questions, the following hypotheses are proposed:

H1: There is a positive relationship between TW and proficiency of the Iranian EFL students.

H2: There is a positive relationship between TW and gender of the Iranian EFL students.

5. Methodology

A number of methods have been designed to elicit and determine the difficulty and frequency within which TW strategies are employed among different age groups and across different language skills and components. These methods were applied by researchers and testing experts via using a variety of instruments and procedures namely: (1) Test of test – taking skills, (2) Passage independence test, (3) Direct interviews, (4) Use of test formats, (5) Evaluating the results of test – taking skills training (Scruggs & Mastropieri, 1992). Of all these methods and techniques employed to the assessment of TW strategies, test of test – taking skills has received a considerable attention for its economy and practicality. Justified on these characteristics, the researchers of the present study used this technique to investigate TW strategies among the participants.

5.1 Subjects

From among 138 (78 males and 60 females) participants, a total number of 80 undergraduate EFL students from Shahid Chamran University were chosen and divided on the basis of their proficiency scores into four cells of 20 individuals: High Proficient Male, Low Proficient Male, High Proficient Female, Low Proficient Female.

5.2 Instrument

In this research, the instrument used was a test of TW designed by the researchers of the study. It intended to measure seven TW strategies, each with 7 four – option multiple choice items. The seven strategies which were distributed among 50 items are as follows: 1) Stem option, 2) Grammatical cue, 3) Item giveaway, 4) Longer length option, 5) Option inclusion, 6) Similar option, 7) Specific determiner.

All of the 50 items were originally standardized in nature but after modification and for the present research purpose they were validated by three experts in the field of language teaching with an interest in language testing as well. They checked the items so that there was only one TW strategy per item, and that each item could be approached using only one strategy. Pilot items about which there was disagreement were either modified or dropped.

5.3 Procedures

Inspired by a pilot study which was done one year before the main experiment with 26 undergraduate EFL students from Shahid Chamran University, the researchers determined their procedure with some confidence. Ideas like total number of items, test administration time, number of strategies and participants were helpful in conducting the main experiment. In short, pilot study proved the feasibility of performing the main experiment.

Then, for the main study, a total number of 80 undergraduate EFL students from Shahid Chamran University majoring in English were chosen for the final experiment. They were screened and divided on the basis of their proficiency scores into four cells of 20 individuals. All of them were required to sit for the second exam after two weeks periods from the first experiment. They were gathered in the examination hall and supplied with a test of TW comprised of 50 items. All students were able to complete the test in 30-35 minutes. After the test completion some of the students voluntarily commented and verbalized many item flaws as well as strategies they used to answer. Their assertions were somewhat confirmed and evidenced in the data analysis.

After the test completion, answer sheets were collected and scored for the number of strategies the subjects employed correctly, without penalizing them for wrong answers. Then, the seven randomly distributed strategies were extracted from among 50 items by their pre – allocated item numbers. The scores were ranked on the number and frequency of strategies used by the participants. Moreover, high and low prevalent strategies in each group of the four cells of test-takers were roughly identified. To assess the exact number of TW strategies within and among the four groups as well as their correlations with proficiency and gender as independent variables, the data were yielded to a two way analysis of variance (ANOVA), t-test, and f-test. On the basis of these statistical procedures, meaningful differences were found among the four cells of participants which are discussed in the following part.

6. Data analysis and results

Mean scores and standard deviations for the performance on the test of TW for the four groups showed that high proficient male and female students had a better performance compared with their low proficient counterparts. Table 1 illustrates the results of one way ANOVA for the test of TW.

Insert Table 1 right about here

Although significant differences were observed among the four groups, they were not generalizable for all the TW strategies. That is, high proficient groups outperformed their low proficient counterparts only in three TW strategies (stem option, grammatical cue, and item giveaway) and these differences were not maintained for the other four strategies. More precisely, all groups performed almost equally in these four strategies. However, these results showed that TW and proficiency of students are positively correlated. In other words, more proficient students were more test-wise, too. Therefore, the first hypothesis is accepted at 0.05 level of probability. In brief, there is a significant

relationship ($P > 0.05$) between TW and proficiency of the Iranian EFL students.

Insert Table 2 right about here

As the results indicate, there is no significant difference between the two genders (Table 2). Moreover, this lack of difference is observed through all the strategies. Therefore, gender is not a defining factor regarding the TW. Hence, hypothesis 2 is safely rejected at 0.05 level of probability. That is, there is no significant relationship ($P < 0.05$) between TW and gender of the Iranian EFL students.

7. Discussion

One possible avenue of studying TW is to search for its possible correlates. This study was an attempt to investigate the correlations of TW and proficiency as well as TW and gender. The results indicated a positive correlation between TW and proficiency in which high proficient groups outperformed their low proficient counterparts in at least three TW strategies (grammatical cue, stem option, and item giveaway). The difference in implementing grammatical cue may be related to the amount and courses of grammar they have passed. Since high proficient groups were in higher levels of education, it seems logical to justify the obtained results by their greater exposure to grammar.

The difference in using stem option strategy may be sought in high cognitive maturation of the high proficient groups as a result of the greater exposure to test and test-taking experience. Because, contrary to the test-wise students, test-naïve ones looked for the answers only among the alternatives, while the former groups probably grasped the idea that understanding the stem-option relationship was the key and prerequisite to find the correct answer.

Another TW strategy in which proficient students had a better performance was item giveaway strategy. Test constructor may sometimes inadvertently give away the correct answer in another part of the test. This paves the way for an unfair situation in which test-wise students have advantage over other students because test-wise students have probably possessed a global and holistic view of test-taking routine. That is, they perceived the test items as a united and related entity, whereas test-naïve ones probably depended only on their knowledge of subject matter. Moreover, successful use of cues depends on previous test-taking experience.

However, these discrepancies among high and low proficient subjects were not prevailed through the other four strategies namely, longer option, option inclusion, similar option, and specific determiner. Implementing these strategies may necessitate high cognitive maturation on behalf of the students, which is appeared to be lacking in the test-taking power of the Iranian EFL students. These very findings are consistent with those obtained by Allan (1992), Diamond and Evans (1972), Diamond et al. (1976), and Morse (1998) in which the increasing mean P. value order of difficulty of the TW strategy was more or less alike as follows: stem option (5.64), item give away (5.35), grammatical cue (4.11), similar options (3.63), specific determiner (3.33), and longer length option (3.08). Thus, the present research along with other studies proves that TW strategies cannot be considered as a general trait and they are rather cue-specific. That is EFL students are equipped with some but not all TW strategies.

Another purpose of this study was to find if there is a relationship between TW and gender of the Iranian EFL students. Consistent with the results of previous studies, the present study found no significant relationship between TW and gender. At best, this equal performance of male and female students on the test of TW may possibly be attributed to their equal instructional and educational opportunities during their language learning career.

8. Conclusion

This study attempted to clarify the relationship between college students' proficiency and gender with their TW strategies. The results indicated a positive correlation between TW proficiency. That is, high proficient students outperformed their low proficient counterparts in TW. It shows that "Test-wise advice like "when in doubt, pick C", and "if you don't know, pick the longest answer" is still passed down from generation to generation of students" (Mohamed, Gregory and Austin, 2006). However, in order to make the evaluation system more valid and reliable, there should be ways to impede this sort of guessing. Therefore, the outcomes of the present research accompanied by those of aforementioned studies can be of value to various fields and professions dealing with evaluating process, teaching courses, curriculum development, teacher training courses, and test-construction processes. Appreciating the important role of TW in evaluation process, it is possible to get a more accurate picture of what we are measuring and move toward better testing tools, thereby an accurate decision making.

9. Limitations of the study

It goes without saying, however, that every study has its own particular limitations and this study is not an exception, too. The first limitation is concerned with the number of individuals who participated in the second phase of the experiment. To provide the obtained results with a high reliability, a wider range of test-taking population seems to be necessary. So, this requires another project with wider range of coverage and facilities. The second limitation of the study has to do with the fact that it is not known to what extent these results may apply to other test-formats or other test-taking circumstances such as standardized test, non-educational tests, etc. Therefore, other studies are required to

focus on the applicability of TW strategies in other test formats and specific tests such as matching, comprehension question, fill in the blanks, etc. However, it should be mentioned that these limitations did not have any influence on the obtained results. Rather, they are mentioned to prepare and propose other interested researchers to do future studies about the suggested issues.

References

- Allan, A. (1992). Development and validation of a scale to measure test-wiseness in EFL/ESL reading test takers. *Language Testing*, 9, 101-121.
- Allen, G. J. (1972). The behavioral treatment of test anxiety: recent research and future trends. *Behavior Therapy*, 3, 253-262.
- Bachman, L. E. (1991). What does language testing have to offer? *TESOL Quarterly*, 25, 671-704.
- Bajtelsmit, J. W. (1977). Test-wiseness and systematic desensitization programs for increasing adult test-taking skills. *Journal of Educational Measurement*, 14, 335-341.
- Beidel, D. C. & S. M. Turner. (1988). Comorbidity of test anxiety and other anxiety disorders in children. *Journal of Abnormal Child Psychology*, 16, 275-287.
- Bergman, I. B. (1980). The effects of providing test-taking instruction for various types of examinations to a selected sample of junior college students. *ERIC Document Reproduction Service No. ED 180566*.
- Brown, J. D. (1996). *Testing in Language Programs*. New Jersey: Prentice Hall.
- Callenbach, C. (1973). The effect of instruction and practice in content-independent test-taking techniques upon the standardized reading test scores of selected second-grade students. *Journal of Educational Measurement*, 10, 25-30.
- Davies, A. (2003). Three heresies of language testing research. *Language Testing*, 20/4, 355-368.
- Deerman, M., Fluker, C., Panik, E., Powell, J., Shelton, K., Uline, C. S. & Notar, C.E. (2008). Standardized Tests: Bellwether of Achievement? *Asian Social Science*, 4(5), 60-64.
- Diamond, J. J. & W. J. Evans. (1972). An investigation of the cognitive correlates of test-wiseness. *Journal of Educational Measurement*, 9, 145-150.
- Diamond, J. J., J. Ayres, R. Fishman & P. Green. (1976). Are inner city children test-wise? *Journal of Educational Measurement*, 14, 39-45.
- Dunn, T. F. & L. G. Goldstein. (1959). Test difficulty, validity and reliability as functions of selected multiple-choice item construction principles. *Educational and Psychological Measurement*, 19, 171-179.
- Edwards, B. (2003). An investigation of factors contributing to a reduction in race-based subgroup differences on a constructed response paper-and-pencil test of achievement. Unpublished Ph.D. dissertation. Office of Graduate Studies of Texas A & M University.
- Farhady, H. (1982). Measures of Language Proficiency from the learner's perspective. *TESOL Quarterly*, 16, 43-59.
- Gronlund, N. E. & R. L. Linn. (1990). *Measurement and Evaluation in Teaching*. New York: Macmillan Publishing Company.
- Houston, S. E. (2005). Test-wisness training: An investigation of the impact of test-wisness in an employment setting. *Unpublished Ph.D Dissertation*. The Graduate Faculty of The University of Akron.
- Koons, B. M. & M. W. Vasey. (2000). Considering gender differences in anxiety and its disorders across the life span: a construct-validation approach. *Applied and Preventive Psychology*, 9, 191-209.
- Lo, M. Y. & M. J. Slakter. (1973). Risk taking and test-wisness of Chinese students. *The Journal of Experimental Education*, 42, 56-59.
- Miller, P. M., D. R. Fuqua & N. S. Fagley. (1990). Factor structure of the Gibb experimental test of test-wisness. *Educational and Psychological Measurement*, 50, 203-208.
- Millman, J., C. H. Bishop & R. Ebel. (1965). An analysis of test-wisness. *Educational and Psychological Measurement*, 25, 707-726.
- Mohamed, A., P. A. M. Gregory & Austin, Z. (2006). Testwisness among international Pharmacy graduates and Canadian Senior Pharmacy students. *American Journal of Pharmaceutical Education* 70 (6) Article 131, 1-6.
- Morse, D. T. (1998). The relative difficulty of selected test-wisness skills among college students. *Educational and Psychological Measurement*, 58, 299-409.
- Oakland, T. (1972). The effect of test-wisness materials on standardized test performance of preschool disadvantaged children. *Journal of School Psychology*, 10, 355-360.

Petty, N. E. & E. H. Harrell. (1977). Effect of programmed instruction related to motivation, anxiety, and test-wisness on group IQ test performance. *Journal of Educational Psychology*, 69, 630-635.

Rowley, G. L. (1974). which examinees are favored by the use of multiple-choice tests? *Journal of Educational Measurement*, 11, 15-23.

Sarnacki, R. E. (1979). An examination of test-wisness in the cognitive test domain. *Review of Educational Research*, 49, 252-279.

Scruggs, T. E. & M. A. Mastropieri. (1992). *Assessing test-taking skills*. Cambridge: Brookline Books.

Shavelson, R. J., N. M. Webb & L. Burstein. (1986). Measurement of teaching. In M. C. Wittrock (ed.), *Handbook of Research on Teaching: a project of the American Educational Research Association*, 50-91. The American Educational Research Association, Macmillan Publishing Company.

Shuller, S. M. (1979). A large-scale assessment of an instructional program to improve test-wisness in elementary school students. *ERIC Document Reproduction Service No. ED 189143*.

Slakter, M. J., R. A. Koehler & S. H. Hampton. (1970). Grade level, gender, and selected aspects of test-wisness. *Journal of Educational Measurement*, 7, 119-122.

Vattanapath, R. & K. Jaiprayoon. (1999). An assessment of the effectiveness of teaching test-taking strategies for multiple-choice English reading comprehension tests. *Occasional Papers*, 8, 57-71.

Table 1. one way ANOVA for the test of TW

Sources	D.F	SS	MS	F	F
				Ratio	Prob.
Between Groups	3	945.3	315.10	15.13	0.00
Within Groups	76	1582.7	20.82	-	-
Total	79	2528.0	-	-	-

D.F= degree of freedom

F=statistic value

SS=sum of squares

MS=mean square

Table 2. t-test results for TW and Sex variables

Sex	Number of cases	Mean	SD	SEM	F Ratio	P
1	40	30.22	5.92	0.93	0.55	0.45
2	40	27.77	5.15	0.81	-	-

SD= standard deviation

P= probability ratio

SEM= standard error of measurement

Appendix: Test of TW

This test contains vocabulary items and structures that you may never have seen before. However, it is possible to answer the questions successfully by using skills and intuition.

Select the best answer out of a, b, c or d.

1. Impressionist artists tried to transitory visual impression of the real world.

*a) capture

b) prevailed

c) intrinsically

d) achievement

Grammatical cue strategy

2. Composer Philip Glass created single-handedly a new musical genre with both classical and Popular appeal. He was

a) brilliant novelist

b) a well-known painter

*c) a splendid musician

d) a popular anthropologist

Stem option strategy

3. He is fed up with those kinds of jobs, because they are

a) fascinating

b) attractive

c) amazing

*d) boring

Similar option strategy

4. A guyot is

a) a tall tree

b) a kind of boat

c) a beautiful wild flower

*d) an under water mountain that does not reach the surface of the sea

Longer option strategy

5. Pantean dessert is

a) never eaten for breakfast

*b) usually eaten after a meal

c) always eaten at night

d) seldom eaten by children

Specific determiner strategy

6. The author of that book implies that in 1960's Dr. King wasto few people in Montgomery

a) primarily

b) gratitude

*c) known

d) survive

Grammatical cue strategy

7. In the 1860, Paul Tradson, a Danish surgeon, reported that damage to specific part of spinal cords was resulted to extreme difficulty in body's movement. This disorder is known as.....

a) Phineas disability

*b) Tradson's disorder

c) Mark's disease

d) Morgan's aphasia

Stem option strategy

8. Of all the language courses, French is

a) the least difficult to learn

b) the easiest to study

c) learned effortlessly

*d) mentally challenging

Similar option strategy

9. Tarestosia is

*a) the study of universe as a whole and its nature

b) a sport

c) the art of writing

d) a kind of illness

Longer option strategy

10. His speech was laconic. It means

a) he spoke very much

*b) he expressed much in a few words

c) his speech was bitter

- d) he spoke critically Specific determiner strategy
11. "Sleep Learning" has become an..... topic of study in recent years.
- a) fascinating
- b) devastating
- *c) intriguing
- d) pleasing Grammatical cue strategy
12. Granite crystals are large enough to be seen with
- *a) naked eyes
- b) big telescopes
- c) precise spectacles
- d) geometrical devices Stem option strategy
13. Paul Tradson was a
- a) painter
- b) musician
- c) novelist
- *d) surgeon Item giveaway strategy
14. Someone suffering from insomnia has trouble sleeping during the day and night.
- a) seldom
- *b) usually
- c) rarely
- d) infrequently Similar option strategy
15. "Procrastinate" means
- a) to play
- b) to make fun of somebody
- *c) to put off doing something until a future time
- d) to be famous Longer option strategy
16. Dr. King met a few people in Montgomery in..... .
- a) 1970's
- b) 1980's
- c) 1975
- *d) 1960's Item giveaway strategy
17. Philip Glass created
- *a) a new musical genre
- b) anthropological conventions
- c) attractive plays
- d) amazing works of painting Item giveaway Strategy
18. Football coach told the players to be abstemious. He meant the players should
- a) not eat anything offered to them
- *b) be moderate in eating and drinking
- c) eat only after the game
- d) drink a lot of water Specific determiner strategy
19. There are a lot of men working there, but only one of them..... .
- *a) ranches cattle in the farm
- b) help me to cure our illness

- c) communicate with divers
 d) defend against invaders Grammatical cue strategy
20. Frederica Von Stade has sung in opera houses throughout the U.S.A and abroad. He is a well-known
 a) artist
 b) actor
 *c) singer
 d) householder Stem option strategy
21. In Mackinnox time.....
 a) the day is longer than night
 b) the earth orbits very quickly
 c) the weather becomes very cold
 *d) the day and night have equal length Specific determiner strategy
22. Granite can be found
 a) naturally in the mountains
 *b) in the form of the crystals
 c) in small and tiny sizes
 d) in every mine fields Item giveaway strategy
23. He has got a mild sore throat, so he coughs at night.
 *a) often
 b) constantly
 c) always
 d) unceasingly Similar option strategy
24. Grandma Moses, popular painter, spent her life in a Little community.
 a) enormous
 b) unusual
 c) isolated
 *d) tranquil Grammatical cue strategy
25. Everyone agrees that she is similitude of her mother. It means she her mother .
 a) works for
 b) is influenced by
 *c) looks like
 d) fights Stem option strategy
26. The Taylor returned the lawn mower to their neighbors
 a) immediately
 b) in time
 *c) not promptly
 d) right away Similar option strategy
27. Insomnia is a kind of
 a) flower
 b) disease
 *c) sleep learning
 d) music Item giveaway strategy
28. The idiom “to turn the table” means
 a) to be happy

- b) to get angry
- *c) to change a situation to your own advantage
- d) to complain Longer option strategy
29. A person of mediocre abilities or attainments is the one
- *a) who is not very good or very bad in his attainments
- b) who is always successful in his work
- c) whose abilities are amazing and fascinating
- d) who is never able to do a work Specific determiner strategy
30. "Labyrinth" is
- a) an authorized person
- b) a soft job
- c) a hasty action
- * d) an arrangement of winding passages Longer option strategy
31. Benjamin Franklin was a/an
- a) scientist and philosopher
- b) inventor and craftsman
- c) humanitarian and essayist
- *d) all of the above Option inclusion strategy
32. Gold, silver, copper and platinum are found in
- a) mines
- b) rocks
- * c) nature
- d) mountains Option inclusion strategy
33. It is very discourteous to during one's conversation.
- a) devastating
- *b) intrude
- c) demolished
- d) frivolously Grammatical cue strategy
34. Symbiosis means
- a) absence of government control
- b) having an insatiable appetite
- *c) close association of two organisms resulting in advantage to both
- d) style of hair arrangement Longer option strategy
35. Although tornadoes occur in many regions of the world, they are mostly prevalent in
- *a) the United States
- b) California
- c) Texas
- d) Denver Option inclusion strategy
36. Scientists have established that man was present in the United States as early as
- *a) 50,000
- b) 25,000
- c) 75,000
- d) 70,000 Option inclusion strategy
37. Many animals use odors for

- a) identification
- b) sexual attraction
- c) territorial marking

*d) alarm and other purposes

Option inclusion strategy

38. The most spectacular recurring comet to be seen in historic times is, named after the English astronomer Edmund Halley, who discovered its periodicity in 1705.

- a) Evan's star
- *b) Halley's comet
- c) The Galaxy
- d) Milky Way

Stem Option strategy

39. Grandma Moses was a popular

- a) politician
- b) writer
- c) musician
- *d) painter

Item giveaway strategy

40. Liquid or gas pressure is exerted

- a) only in one direction
- b) in three directions
- *c) equally in all directions
- d) in two direction

Option inclusion strategy

41. Pinalorous land is a kind of area found in some countries.

- a) vast
- *b) dried
- c) unlimited
- d) boundless

Similar option strategy

42. Estuary is a/an

- a) waterfall
- b) polluted area
- c) area of body
- *d) mouth of a river where it mixes with the sea

Longer option strategy

43. Since Elizabeth Barret Browning's never approved of her marrying Robert Browning, the couple eloped to Italy where they lived and wrote. Mr. and Mrs. Browning are famous as..... nowadays.

- *a) writers
- b) lecturers
- c) philosophers
- d) humanitarian

Stem option strategy

44. Because of the extreme pressure underwater, divers are sluggish.

- a) never
- *b) often
- c) seldom
- d) always

Specific determiner strategy

45. He received a letter from his friend.

- a) impeccable
- b) obsolete

