ASIAN EFL JOURNAL
The EFL Professional's Written Forum

Exploring the Effects of Learner Training on Motivation

Nalan Bayraktar Balkır and Ece Zehir Topkaya

Çanakkale Onsekiz Mart University, Turkey

Bio Data

Nalan Bayraktar Balkır has taught EFL for over nine years in İstanbul and Çanakkale in Turkey. At present she is teaching at the Preparatory School of English at Çanakkale Onsekiz Mart University. She holds an MA in ELT from the same university. Her main research interests include psychology of learning, learner-training and vocabulary teaching.

Ece Zehir Topkaya is an assistant professor in the Department of English Language Teaching at Çanakkale Onsekiz Mart University, Turkey. She holds a PhD (TEFL) from Dokuz Eylül University, Turkey. Her research interests include psychology of learning and pre- and in-service teacher education.

Abstract

This study was implemented in order to investigate the effects of a learner training programme on learners' motivation in learning English. The study, which adopted the pre-experimental study design, was carried out with 30 participants studying in Compulsory and Voluntary English Preparatory Programme at Çanakkale Onsekiz Mart University, Turkey in 2005-2006 academic year. The data were collected by means of quantitative and qualitative research techniques. Pre-tests and post-tests (questionnaires) given before and after learner training sessions were quantitative in nature; and a qualitative approach was adopted for the follow-up interview. The results of quantitative findings revealed that there was not a significant difference in learners' overall motivation after the treatment. However, the analysis of qualitative findings from the interview indicated that a reasonable level of increase in learners' motivation occurred. Furthermore, learners' metacognition turned out to be developed. Finally, no notable differences between intrinsic and extrinsic motivation, and between female and male students' motivation were detected. This study concludes that learner training activities have resulted in a moderate increase in learners' motivational level and enhanced metacognition. In the light of these findings, this study draws attention to the importance of learner training in foreign language learning and points to some methodological and pedagogical implications. Finally, it offers some suggestions for further research.

Key words: learner training, metacognition, motivation, intrinsic/extrinsic motivation

1. Introduction

The increasing realisation of the fact that learning is a life-long process has led to the need to employ a learning-to-learn approach in most educational settings in recent decades. A learning-to-learn approach is closely related to learner training which is to do with the attempt to equip learners with necessary and essential skills and knowledge in a particular subject area so that they could manage their further learning more effectively. Learner training also aims to build and heighten learners' awareness of themselves as learners and the language learning process. In other words, the overall aim of learner training is to contribute to both learners' cognitive and affective development. Foreign language learning is one of the subject areas in which learners are expected to learn how to learn in order to be more self-sufficient in their language learning process. In this respect, foreign language teachers have another main duty of teaching learners to learn how to learn as well as teaching them the content of the target language.

The results of most studies indicate that the implementation of learner training or a learning-to-learn approach has been very beneficial and rewarding in terms of many aspects of the language learning process such as learner beliefs, attitudes, autonomy, perceptions, self-confidence, self-esteem, motivation etc. The purpose of this article is to report on a learner training study and its effects on participants' motivation.

2. Review of literature

In this section, a brief description of the key concepts related to the study such as learner training, motivation, and the link between the two will be presented in turn. In addition, the findings or the results of some of the related studies in the literature will be mentioned.

2.1. Learner Training

The term *learner training* has been defined in quite similar terms by different scholars in the relevant literature. According to Hedge (1993, p. 92), for instance, learner training is to do with "a set of procedures or activities which raise learners' awareness

of what is involved in the process of learning a second language, which encourage learners to become more involved and responsible for their own learning, and which help learners to develop and strengthen their strategies for language learning". The term is often used interchangeably with *learner development* which is defined as "cognitive and affective development involving increasing awareness of oneself as a learner and an increasing willingness and ability to manage one's own learning" (Sheerin, 1997, pp.59-60). Both of the terms are concerned with the attempt to help learners maximise their learning efficiency.

Learner training involves two main components: *raising learners' awareness* and *helping them acquire a set of skills* for more effective learning (Hedge, 2000). In other words, typical learner training programmes aim at enhancing learners' *metacognition* which is defined as "an awareness of one's own mental processes and an ability to reflect on how one learns, in other words, knowing about one's knowing" (Williams and Burden, 1997, p.148).

There are different ways or methods to implement learner training in foreign language learning settings. According to Wenden (1998a), for example, there are two main methods for implementing learner training: indirect methods and direct methods. *Indirect methods* use inductive or discovery learning through which "learners reflect upon past learning experiences or acquired knowledge in order to seek insight into their approach to learning and their beliefs" (Wenden, 1998a, pp.16-17). The common elements in the frameworks of schedules using indirect methods are setting goals and objectives, having discussions on metacognitive knowledge and reflecting on past experience (Cotterall, 2000; Dickinson, 1993; Wenden, 1998b). *Direct methods* for implementing learner training, on the other hand, involve deductive and didactic learning activities. These activities are mainly implemented for cognitive strategy instruction (Wenden, 1998a, p.19). The instruction is in the form of explicit and informed teaching of cognitive strategies. Direct and indirect methods are usually used in combination in most of the learner training programmes and textbooks designed for learner training; and many practitioners also support the use of different methods in combination (e.g. Benson, 2001; Logan & Moore, 2004; McCarthy, 1998;

Wenden, 1998a).

Most researchers who have conducted learner training studies appear to have similar concerns about the principles for implementing learner training programmes. Firstly, learners should be assisted to gain a sense of awareness of themselves as learners and the language learning process. Next, a supportive environment should be provided, which combines the principles of goal-setting, developing language learning strategies, undertaking self-assessment, building an awareness of learning styles, encouraging self-confidence and motivation, raising metacognitive awareness, cooperating with the teacher and peers, and reflecting on one's progress (e.g. Cotterall, 1999; Dickinson, 1993; Esch, 1997; Finch, 1998; Gower, Philips & Walter, 1995; Lee, 1998; Nunan, 2002; Sheerin, 1997; Victori & Lockhart, 1995; Wenden, 1998b; Yang, 1998).

General conclusions that the researchers have drawn from learner training studies reveal that having been exposed to learner training and awareness building activities, students get significant gains in terms of cognitive and affective development. In conclusion, learner training is a valuable practice which is likely to encourage learners gradually to become more active, successful, responsible, autonomous and motivated in their learning process.

2.2. Motivation

While the term *motivation* is often assumed to be a rather abstract concept (Dörnyei, 2001), it could be basically defined as "the extent to which you make choices about (a) goals to pursue and (b) the effort you will devote to that pursuit" (Brown, 2001, p. 72). As a more detailed definition suggests, motivation is "a state of cognitive and emotional arousal, which leads to a conscious decision to act, and which gives rise to a period of sustained intellectual and/or physical effort, in order to attain a previously set goal (or goals)" (Williams & Burden, 1997, p. 120).

It could be inferred from the definitions above that the occurrence of motivation is affected by a number of different factors, and goal-setting is an indispensable part of motivation. In addition, sustaining effort is as significant as initiating motivation.

Therefore, when attempting to enhance learners' motivation in learning a foreign language, different aspects or stages of motivation and factors affecting it should be taken into account.

In the literature on motivation, the concept has been classified in different ways. The most well-known classification of motivation has been made in cognitive terms as whether it is internal or external to the learner. In this respect, intrinsic motivation is to do with "behaviour performed for its own sake in order to experience pleasure and satisfaction" (Dörnyei, 2001, p.27) while extrinsic motivation is about "performing a behaviour as a means to an end, that is, to receive some extrinsic reward or to avoid punishment" (Dörnyei, 2001, p.27). In other words, intrinsically motivated people accomplish a task for an internal interest whereas extrinsically motivated people perform an activity because of an interest external to the activity (Dickinson, 1995, p.169).

There is the common belief among many scholars that intrinsically motivated learners will reach more effective learning outcomes and attain mastery in a foreign language (e.g. Dickinson, 1995; Ersöz, 2004; Mynard, 1999; Williams & Burden, 1997). However, it is also suggested that intrinsic and extrinsic motivation are actually quite interrelated and most behaviours are affected by a combination of both types of drives (Williams & Burden, 1997, p.123). Brown (2001) also points out that some research studies indicate that extrinsic drives could have a considerable effect on intrinsic motivation. A research study conducted by Salehi (2005, p.5) seems to support the latter view on motivation by indicating that there is a high correlation between extrinsic and intrinsic motivation. Therefore, it is stated that these two distinct types of motivation are not actually mutually exclusive.

2.3. The link between learner training and motivation

It has been previously mentioned that learner training schedules aim to enhance learners' cognitive and affective development through an attempt to build learners' awareness of themselves as learners and the language learning process, and also assist them to attain a set of skills for more effective learning. The common aim of such

research studies is to investigate the effects of learner training on different constructs such as learner autonomy, academic success, beliefs about and attitudes towards language learning, motivation, etc.

Even though there is a small number of studies exploring the effects of learner training programmes on motivation, the results of several studies point to some link between learner training and motivation. For example, the results of the self-directed programme conducted by Victori & Lockhart (1995, p.228) pointed to the improvement in learners' tolerance of ambiguity in language learning/use contexts, increased motivation, and improved self-esteem as language learners. Furthermore, the results of Lee's (1998) study revealed that more enthusiastic students attained more gains from the self-directed learning programme than less enthusiastic students did. Similarly, Mynard's (1999) study indicated that more able learners tended to have higher intrinsic motivation and a higher internal locus of control than less able ones.

Without a doubt, variables such as learner characteristics, teacher characteristics, the practices followed during learner training, etc. all interact differently in different contexts. Therefore, to investigate the topic of learner training from as many angles as possible in varied ways will add new insights to our understanding of the issue better. In this respect, the attempts taken in this particular study to implement such a training program and its results will hopefully contribute to the existing literature.

3. Research Problem and Questions

The present study aimed to investigate the effects of learner training sessions on participants' motivation in learning English. In addition, any possible significant differences between female and male learners' motivational level in learning English before and after learner training sessions were investigated. Finally, it was also intended to explore any significant differences between intrinsic and extrinsic motivation before and after the treatment. With these aims in mind, the following research questions were put forward:

RQ 1: What is learners' motivational level in learning English before and after

treatment (learner training sessions)?

RQ 2: Is there a significant difference in learners' motivational level after treatment in terms of gender?

RQ 3: Is there a significant difference between learners' extrinsic and intrinsic motivation before and after treatment?

RQ 4: How do learners evaluate learner training sessions?

4. Methodology

This section describes the methodology of the study, particularly design, participants and setting, limitations, instruments, and procedures.

4.1 Design of the Study

In this study, pre-experimental research design was utilised. To specify, only one intact group participated in the study; and that group was administered a pre-test prior to the treatment (learner training programme). Following the treatment, a post-test was conducted again to find out whether the treatment caused any significant differences in participants' motivational level in learning English. In addition, interviews were conducted with a group of participants in order to collect qualitative data.

4.2 Participants and Setting

The study was conducted in the Compulsory and Voluntary English Preparatory Programme at Çanakkale Onsekiz Mart University. The students enrolled in this programme were going to study in different departments such as 'Tourism and Hotel Management', 'Fine Arts', 'Drama', 'Physics', 'Public Relations and Advertising', 'Foreign Trade and European Union' and 'Accounting' after a two-semester-long general English course. The main reason for conducting the study in this setting was its convenience to the researcher since she worked at the same institution at the time of the research. Thus, arranging the appropriate time and conditions for the implementation of the study was easier. The study covered a period of five weeks

during the fall semester of 2005-2006 academic year.

The study was implemented with 30 subjects from one whole group. 17 of the subjects were female and 13 were male. Their age ranged from 18 to 22. All of the subjects were native speakers of Turkish with an elementary level of English. 28 of the subjects were compulsory and 2 were voluntary programme students. As for the interview part of the study, 9 volunteers were chosen. 6 of the interviewees were female and 3 male.

4.3 Limitations

This study had several limitations:

- Since this study is small-scale, the conclusions are not easily generalised, i.e. the results of the study only concern the subjects of this study.
- The learner training schedule began in the middle of the first school term. It could have been more fruitful if it had started at the beginning of the term.
- The study ended in the last week of the first term when the learners were rather exhausted by being overloaded with a number of exams, projects and assignments. This fact might have affected learners' responses in the post-test.

4.4 Instruments

In order to investigate the research questions stated previously, the researchers developed a *questionnaire on motivation*. This instrument was developed by adapting various items about motivation from different sources. Most of the items were taken and adapted from Mynard's (1999) questionnaire and Demir's (2004) Attitude-Motivation Scale. The remaining items were generated in the light of theoretical aspects of motivation by the researchers. After deciding which statements to include in the questionnaire, the statements were translated from English into Turkish. Back-translation of the statements was undertaken by two English instructors. Next, three experts at the department of ELT at the faculty of Education in Çanakkale Onsekiz Mart University were requested to check and evaluate the questionnaire in terms of face and content validity, wording and the clarity of the items. Finally,

necessary alterations were done on the questionnaire items in the light of the experts' suggestions and comments.

The questionnaire on motivation, which involved 34 close-ended items originally, was built on a five-point Likert scale having the options of (1) I strongly agree, (2) I agree, (3) I am not sure, (4) I disagree, (5) I strongly disagree The analysis of the data collected from the piloting of questionnaires signalled that some of the items needed to be removed in order to ensure that data collection was reliable. After the necessary elimination, the final version of the questionnaire included 19 items which were divided into the two dimensions of extrinsic motivation and intrinsic motivation (see Appendix 1). Then, the remaining 19 items were analyzed statistically. The Cronbach-alpha value of 12 items belonging to *intrinsic motivation* was found highly reliable (α= ,86) for data collection. The Cronbach-alpha value of items related to extrinsic motivation was found to be ,69, which is accepted to be moderately reliable according to the literature (Şencan, 2005).

4.5 Procedures

This section includes a description of the procedures conducted for data collection and analysis.

4.5.1. Pre-Experimental Study

Prior to the study, the participants were informed about the purpose, content, length, time, language and procedures of the study. Then, they were asked whether they would like to participate in the study, and all of them agreed to take part.

The study was carried out in a five-week-long period. Before the training sessions started, a motivation pre-test was administered. Altogether 6 sessions were held in the participants' own classrooms after regular class hours. Each session focusing on a different aspect of foreign language learning lasting approximately 50 minutes. Participants' mother tongue, Turkish, was used during the sessions since the general aim of the study was to help learners learn how to learn and in relation to this, investigate their motivational level. In addition, learners did not seem to be confident

enough to speak in English about complicated matters as their level of English was just elementary at the time of the research. The pre-experimental design of the study is summarised and the topic of each session is presented in the following figure:

Figure 1: The pre-experimental design of the study

Pre-treatment	**TREATMENT (Learner training and awareness building activities)**						**Post-treatment**
Questionnaire on motivation	1st session	2nd session	3rd session	4th session	5th session	6th session	Questionnaire on motivation
	Features of good language learners	I. What kind of a learner am I? II. Determining strengths and weaknesses	Needs analysis and goal setting	I. Learning styles II. Language learning strategies	I. Teacher and learner roles II. Self-assessment	Group and whole class activities on foreign language learning experience and process	
	Diary	Diary	Diary	Diary	Diary	Diary	

A typical learner training session started with greeting and establishing rapport and continued with lead-in and main activities related to the topic. In addition, different instruments such as scales and inventories related to the aim of each session were utilised in every session. Then, the participants reflected on their experience from the session by answering these two questions: 1) What have I learned from this learner training session? 2) How and where can I use what I have learned? The aim of this procedure was to increase participants' awareness about the topic and also make them personalise the content. At the end of each session, the researcher and the participants reviewed what they covered in the session (see Appendix 2 for an example session).

In addition to the learner training sessions, the participants were asked to keep a weekly report named 'My English Diary' during the study. The aim of this diary-keeping activity was to back up the sessions through learners' regular reflection on their weekly language learning experience so that their awareness of language learning process would be heightened. These reports, which the participants were

required to fill in outside the class during a school week, involved questions concerning issues such as what they learned, the usefulness of classroom activities, their performance in the class, their studies with homework assignments, the problems they encountered and their suggestions to solve these problems, etc. The forms containing the questions mentioned above completed by the participants were collected at the beginning of each week by the researcher. Following the last session, the questionnaire on motivation was administered as the post-treatment instrument.

4.5.2. Interview

The second phase of the study involved semi-structured interviews by which it was intended to collect qualitative data. The interviews were conducted on the last day of the treatment with nine participants. The participants, who volunteered, were informed about the purpose, duration and conditions of the interview. In addition, their permission was asked for tape-recording their interviews. The interviewer asked questions from general to specific to the participants. Later on, the interviews were transcribed by the researcher to be analysed.

4.5.3. Procedures for data analysis

The data obtained from the pre-experimental study were entered onto the computer and analysed with several statistical procedures like descriptive statistics, Paired Samples T-test and Independent Samples T-test on SPSS (ver. 15.0). The data collected from the interviews were evaluated both qualitatively and quantitatively; the findings were presented; and necessary interpretations were provided.

5. Findings

Findings from the questionnaire and the interview will be presented in the following sections.

5.1. Findings from the questionnaire

In order to find out whether a significant change occurred in learners' motivational level in learning English, firstly descriptive statistics of pre-test motivation was carried out and mean values were calculated. The mean value of pre-test total motivation was found to be 4,31 (see Table 1).

Table 1: Total mean values of motivation pre-test

	N	Mean	SD
Pre-test intrinsic motivation	30	4,34	,46
Pre-test extrinsic motivation	30	4,27	,42
Pre-test total motivation	30	**4,31**	,39

The results of descriptive statistics indicate that learners' overall intrinsic motivation appears to be fairly high (mean: 4,34). This result reveals that the majority of the participants are quite aware of the significance of English in their future lives. Moreover, they choose the particular department they enrolled in knowing that there is an English preparatory class. As it had been previously observed students in these groups were quite motivated to learn English.

To understand the learners' ideas in relation to each specific item, the results of the descriptive statistics are presented below in tables 2 and 3 in two parts corresponding to the two dimensions of motivation: intrinsic and extrinsic motivation.

Table 2: Descriptive statistics of learners' intrinsic motivation in learning English before treatment

Items of questionnaire on motivation	N	Mean	SD
3. I would like to be able to speak English.	30	4,83	,38
2. I would like to visit an English-speaking country.	30	4,77	,57
6. I want to do well in English class.	30	4,67	,66
19. It makes me happy to think that I learn English.	30	4,57	,57
7. I would like to meet English-speaking people.	30	4,53	,63
13. I would like to learn English even if I didn't have to.	30	4,50	,63
18. Learning English is important for my personal development.	30	4,33	,66
1. I like English.	30	4,20	,92
12. When I learn new things in English, I feel satisfied.	30	4,20	,92
4. I enjoy English lessons.	30	4,10	,99
15. I find learning English enjoyable.	30	3,80	1,03
5. I find English interesting.	30	3,53	1,04

As the table above shows, while they find English somewhat interesting (mean: 3,53) and enjoyable (mean: 3,80), it could be noticed that their motivation to learn English (item 12, mean: 4,20; item 18: 4,33; item 19, mean: 4,57; item 6, mean: 4,67) seems to be higher. Especially, item 3 with the top mean value of 4,83 is a sign of learners' high motivation to learn English.

Table 3: Descriptive statistics of learners' extrinsic motivation in learning English before treatment

Items of questionnaire on motivation	N	Mean	SD
8. I will need to know English in the future.	30	4,87	,35
9. It will be important for me to know English in the future.	30	4,87	,35
11. I need to learn English for my future career.	30	4,87	,35
14. I want only to survive the English lesson.	30	4,37	,67
16. If I had the choice I'd give up learning English.	30	4,23	,94
10. The main reason I learn English because I have to.	30	3,73	1,34
17. If I learn to speak English, other people will respect me more.	30	2,97	1,16

Participants' *extrinsic* motivation is also quite high (mean: 4,27) although it is slightly lower than the mean of intrinsic motivation. When the mean values of the items of extrinsic motivation are studied, it is noticed that all of the first three highest scoring items (mean: 4,87) involve statements related to the importance of English with respect to learners' future employment needs.

In order to find out whether the difference between total mean values of extrinsic and intrinsic motivation of pre-test was significant, a Paired-Samples T-test was implemented (see Table 4).

Table 4: Paired-Samples T-test results for the differences betweenextrinsic and intrinsic motivation before treatment

	N	Mean	SD	t	df	Sig.
Intrinsic motivation	30	4,34	,46	,806	29	,427
Extrinsic motivation	30	4,27	,42			

While the mean value of extrinsic motivation (mean: 4,27) is slightly lower than the

mean value of intrinsic motivation (mean: 4,34), the difference is not statistically significant (p>.05). Next, the difference between male and female learners' motivational level before the treatment was sought through an Independent Samples T-test (see table 5).

Table 5: Independent Samples T-test results for gender differences in learners' motivational level before treatment

Gender	N	Mean	SD	t	df	f	Sig.
Female	17	4,34	,36	,496	28	333	,624
Male	13	4,27	,44				

Table 5 indicates that there is not a significant difference between female and male participants' motivation to learn English prior to the treatment (p>.05).

With the aim of finding out the outcomes of post-test motivation, firstly mean values were calculated and mean of post-test total motivation was found to be 4.22 (see table 6).

Table 6: Total mean values of motivation post-test

	N	Mean	SD
Post-test intrinsic motivation	30	4,21	,46
Post-test extrinsic motivation	30	4,23	,42
Post-test total motivation	30	**4,22**	,42

Then, descriptive statistics of intrinsic motivation and extrinsic motivation was conducted (see tables 7 and 8).

Table 7: Descriptive statistics of learners' intrinsic motivation in learning English after treatment

Items of questionnaire on motivation	N	Mean	SD
3. I would like to be able to speak English.	30	4,80	,48
6. I want to do well in English class.	30	4,60	,67
2. I would like to visit an English-speaking country.	30	4,60	,86
7. I would like to meet English-speaking people.	30	4,53	,86
13. I would like to learn English even if I didn't have to.	30	4,40	,89
19. It makes me happy to think that I learn English.	30	4,40	,89
12. When I learn new things in English, I feel satisfied.	30	4,33	,88
18. Learning English is important for my personal development.	30	4,20	,91
1. I like English.	30	4,07	1,01
4. I enjoy English lessons.	30	3,73	,91
15. I find learning English enjoyable.	30	3,40	,93
5. I find English interesting.	30	3,40	1,07

The results of descriptive statistics show that participants' *intrinsic* motivation could be said to be high as in the pre-test, though it was observed that a slight decrease occurred in the mean values (pre-test intrinsic motivation mean: 4,34 and post-test intrinsic motivation mean: 4,21). Similar to the descriptive statistics of the pre-test, the lowest scoring intrinsic motivation items are again the ones involving the statements about the nature of English and the study of English language (items 5 and 15, mean: 3,40; item 4, mean: 3,73; item 1, mean: 4,07). However, they seem to be quite motivated again to learn English. For example, item 30 with the mean value of 4,40 reveals that they would still like to learn English even if it was not compulsory. Similarly, item 19 (mean: 4,40) indicates that they are intrinsically motivated since they become happy when they realise that they learn English. Once again, the mean values of the first four highest scoring items point to the learners' willingness to be able to learn English due to the intrinsic factors.

Table 8: Descriptive statistics of learners' extrinsic motivation in learning English after treatment

Items of questionnaire on motivation	N	Mean	SD
11. I need to learn English for my future career.	30	4,90	,31
9. It will be important for me to know English in the future.	30	4,87	,35
8. I will need to know English in the future.	30	4,87	,35
16. If I had the choice I'd give up learning English.	30	4,23	,90
14. I want only to survive the English lesson	30	4,17	,83
10. The main reason I learn English because I have to.	29	3,45	1,40
17. If I learn to speak English, other people will respect me more.	30	3,07	1,20

The results from descriptive statistics show that learners' extrinsic motivation is considerably high, especially in terms of their future professional concerns again (item 11: 4,90; items 9 and 8, mean: 4,87). On the other hand, participants are not concerned with how others will approach them if they learn English (item 17, mean: 3,07).

After the results of descriptive statistics of post-test motivation were presented and interpreted, a Paired-Samples T-test was carried out so as to discover whether

there is a significant difference between total mean values of extrinsic and intrinsic motivation (see Table 9).

Table 9: Paired-Samples T-test results for the differences between extrinsic and intrinsic motivation after treatment

	N	Mean	SD	t	df	Sig.
Intrinsic motivation	30	4,21	,49	-,241	29	,811
Extrinsic motivation	30	4,23	,38			

The values in Table 9 indicate that the difference between the mean values of extrinsic motivation (mean: 4,23) and intrinsic motivation (mean: 4,21) is not statistically significant (p>.05). With the purpose of finding out the difference between male and female learners' motivational level after treatment, the data were subjected to an Independent Samples t-test (see Table 10).

Table 10: Independent Samples T-test results for gender differences in learners' motivational level after treatment

Gender	N	Mean	SD	t	df	f	Sig.
Female	17	4,13	,42	-1,248	28	,001	,223
Male	13	4,33	,41				

Table 10 indicates that there is not a significant difference between female (mean: 4,13) and male participants' (mean: 4,33) motivation in the post-test (p>.05). In order to find out how significant the difference is between pre-test and post-test total motivation, a Paired-Samples T-test was carried out (see Table 11).

Table 11: Paired-Samples T-test results for the differences between pre-test and post-test total motivation

	N	Mean	SD	t	df	Sig.
Pre-test total motivation	30	4,31	,39	-,914	29	,368
Post-test total motivation	30	4,22	,42			

The values in Table 11 show that there is not a significant difference between participants' motivation before and after the treatment (p>.05). In other words, no

statistically meaningful differences were detected in learners' motivational level after treatment.

5.2. Findings from the interview

In order to analyse the findings from the interviews conducted with nine participants, the interviews were first transcribed and then the findings were interpreted. As mentioned in the methodology of the study, the interviewees were volunteers. Although academic success level was not taken into consideration when selecting the interviewees, it was noticed that their level ranged from 'very high' to 'low' according to the end-of-term grades (The average is '60' out of '100', which is the passing grade in English Preparatory Programme). Three of the participants (33. 3%) had 'very high' level of academic success; four participants' (44. 4 %) academic success level was 'high'; and the academic success level of two participants (22. 2 %) was 'low'.

The interviewees were asked general and specific questions. The *first* question was the most general one: '**What do you think about the activities done in the sessions?'** On the whole**,** the participants had quite positive opinions about the activities. 7 participants out of 9 stated that they found the activities very useful. For example, P-9 said that:

> "Firstly, the activities have been very useful in terms of getting to know myself. Furthermore, I have learned a lot of things about what I can do to learn better."

P-5 also reported:

> "We learned the importance of learning. Therefore, I think the activities have been very useful and to speak frankly, I liked them very much."

Another participant stated that:

> "The activities were very positive that I had the chance to question myself."

The *second* interview question was: '**How have the activities contributed to you**

personally?' The responses to this question exhibit variance. Actually, each interviewee responded to this question quite differently (see Table 12).

Table 12: Interviewees' responses to the second interview question

Participant	Participant's answer
P-1	"I became more aware of things about language learning and got more concentrated. I also learned how to learn more easily."
P-2	"I learned to give more importance to the homework assignments. I used to think that the teachers have to teach us but now I think that we are also responsible for our own learning."
P-3	"I had already known most of these things but you reminded them to me again."
P-4	"These activities reminded me how necessary learning a foreign language is. I became aware of the fact that English is more important than I was told before."
P-5	"Now I think that I have more responsibility for my learning. And I also think about how I should be in the lessons. Once again, I can say that I have gained the feeling of responsibility."
P-6	"I noticed that I could describe people accurately after we had analyzed three types of students in one of the sessions."
P-7	"I became more aware of how to learn English better."
P-8	"I learned to overcome my prejudices about my inability to learn English. In a way, my self-confidence boosted. Moreover, I became more aware of the indispensable and important place of English in my life."
P-9	"I learned what kind of a learner I am."

When Table 12 is examined, it is noticed there is a marked increase in learners' overall metacognition. In terms of motivation, while no significant differences in learners' motivation were found out via statistical procedures, qualitative findings from the interview reveal that some of the participants' motivation seemed to be enhanced. For example, two participants' (P-4 and P-8) motivation to learn English could be said to have increased as both of them stressed that they realised how important English was in their life. Furthermore, P-9 said that:

> "Now I believe that learning English is very important to me because I will extremely need it for my job in the future."

In addition, P-6 stated that:

> "My friend and I have made a decision: from now on we are going to watch films in English more often. We find it very enjoyable to be able to understand some of the words."

The *third* interview question was: **'Have the activities been useful to you? (If so, which one do you think has been the most useful?)'** All of the participants agreed that the activities had been very useful to them; and they specified which activities have particularly helpful to them. Table 13 shows how each participant answered this question.

Table 13: Interviewees' responses to the third interview question

Participant	Participants' answer
P-1	Goal setting and learning styles
P-2	Teacher and learner roles
P-3	Features of good language learners
P-4	All of the activities
P-5	Teacher and learner roles
P-6	Learning styles and character analysis
P-7	Learning styles
P-8	Learning styles and strategies
P-9	What kind of a learner am I? and learning styles

Next, the *fourth* question '**What were your opinions related to the process of learning English before the learner training sessions and what about your present opinions?'** was asked to the interviewees: The interviewees' responses to this question also show variation. For example, P-1 said that:

> "Before these sessions, I used to think that learning English takes a very long time and is more complex but now I know that if I set my goals and plan my studies, the time spent for learning could be reduced. I also gained an increasing awareness about learning English."

P-2 emphasised that she became aware of her responsibility in her learning and also the importance of studying English:

> "I learned to attach more importance to English and that I need to study through both writing and reading after I learned that I was a visual learner. I also learned to take more responsibility for my learning."

Furthermore, P-3 reported that he learned many things about different aspects of language learning. P-4 pointed out that he felt as if the thoughts in his mind came alive when the discussions were being held. In other words, he remembered what he knew previously. Similar to P-1, P-5 also emphasised that he started to think that the time allocated for learning English could be shortened if one knows how to plan his studies carefully. Likewise, P-6 made it clear that if one works hard enough, she can achieve anything. She also added that since she realised that it had been her fault when her performance got poorer, she started to put more effort into her work to improve her performance. Moreover, P-7 noted down that:

> "I think that I did not know how to study to learn more effectively. However, with the help of these activities, I learned something new each week which increased my knowledge about how to study English."

Participants 8 and 9 responded to the fifth question in a similar way. They stated that they had decided how to study English. They also added that they had not been aware of many things about foreign language learning before the sessions. P-9 emphasised that:

> Now I can say that I attend this course more willingly; and I would choose to learn English even if it wasn't compulsory for me."

The *fifth* and last interview question was: **'Do you have a final comment you would like to add?'** All of the participants evaluated the study quite positively. Three of the participants stated that the study became very effective. Two participants said that the activities were quite enlightening. Three participants pointed out that the study was very useful. One participant indicated that the study was very necessary. The seventh participant, for example, had the following comments on the overall study:

> "These activities certainly need to have been done in terms of getting to know ourselves and knowing how to learn English. Because having the desire to learn something is not sufficient alone; we should also be aware of and have knowledge about how to learn most effectively in order to be able to do something. I think these activities have contributed to this aim a lot."

P-4 also had positive opinions about the study:

> "To me, this kind of activities should not only be done for your research study but also for a longer period of time. I hope they will continue in the second term, too. Personally, I had a lot of fun."

To sum up, while no statistically significant differences were found in participants' motivational level in learning English before and after treatment, findings of the qualitative data indicate that their motivational level increased to a certain extent. In addition, participants' metacognitive awareness appeared to be much greater.

6. Discussions of findings

This section involves the discussions of the findings from the overall study.

6.1. Discussions of findings from the quantitative data

The results of the quantitative findings are three-fold: First, one aim was to find out learners' motivational level before the learner training sessions started. The findings indicated that learners' motivation to learn English appeared to be really high at the beginning of the study. Next, although no significant differences were observed between intrinsic and extrinsic motivation, it was noticed that the prevailing source for this high motivation was learners' future employment concerns. This is not surprising since attending English Preparatory Programme was compulsory for most of the participants (28 out of 30). In terms of gender differences with respect to motivation prior to the treatment, female and male students' motivational level turned out to be quite close to each other.

Second, it was found that learners' motivational level was again quite high following the treatment. Similar to the motivation pre-test, no considerable differences were found between intrinsic and extrinsic motivation; while it was noticed again that learners seemed to be willing to learn English particularly because of external factors related to their future professional careers. However, learners also

appeared to be intrinsically motivated as a high scoring item indicated that they would like to study English even if it was not compulsory. In this respect, it could be compared to Salehi's (2005) study in which he suggests that there is not a clear-cut boundary between intrinsic and extrinsic motivation as the findings of his study pointed to a high and meaningful correlation between these two types of motivation. Finally, no meaningful differences between females' and males' motivational level were detected. Similar results were also reached by Dursun (2007) and Salem (2006), who have reported that no significant differences were found in learners' motivational level with respect to gender in their study.

Third, the results of statistical findings indicated that no significant differences were found in learners' overall motivation after learner training and awareness building sessions. This result might be attributed to the general observation that learners' motivation is usually higher in the beginning of the course but it tends to decrease by the end of the term as they are involved in so many engaging activities and exams throughout the course. Since the researcher was the teacher of the group, towards the end of the sessions, fatigue and frustration were extensively observed among the participants.

6.2. Discussions of findings from the qualitative data

As for the results of the last research question, the findings obtained from the interviews indicated that all of the interviewees appeared to have positive opinions regarding the learner training sessions. Responses of the interviews particularly indicated that their metacognitive awareness enhanced considerably. In contrast to the statistical findings which revealed that no significant changes occurred in participants' overall motivational level, the qualitative findings signified a relative amount of increase in their motivation as several of the interviewees reported that they started to perceive English as having a more important place in their lives; attend the course more willingly, etc.

When learners were asked to express their further opinions about the sessions, it was observed that all of them found the activities quite useful. The most favourable

activities of the participants were the ones related to improvement of metacognition such as learner styles, teacher and learner roles and character analysis.

7. Conclusions and implications

The main aim of the present study was to seek any possible significant increases in learners' motivational level after a learner training programme. The statistical results showed that no considerable increase was detected in learners' overall motivation. However, according to the qualitative findings, a reasonable level of increase in participants' motivational level was found. In addition, the results of the qualitative data indicate that learners' metacognitive awareness has increased. Moreover, the findings pointed to no significant differences between intrinsic and extrinsic motivation neither before nor after the treatment. Finally, female and male students turned out to have quite close motivational level in learning English.

The results of the present study have several important implications regarding learners, teachers, and methodological issues. Considering the fact that motivation is a crucial construct shaping learners' behaviours and performance, and determination to achieve goals, it could be suggested that learner training aiming to enhance learners' motivation with a focus on the development of learners' metacognition should be incorporated into every subject area and introduced preferably from the very beginning of education. However, the timing and the duration of the learner training should be more cautiously organised and, thus, be better observed over time since behaviour changes occur in long periods**.** Similarly, the intensity of the teaching programme should also be taken into consideration when designing learning training sessions in order to minimise likely negative effects of such tiring schedule on learners' motivation. As for foreign language teacher education, the training of pre-service language teachers should be equally concerned with teaching how to learn apart from teaching how to teach the content area.

This study also has several methodological implications. For this particular research, the training sessions were designed as separate sections apart from regular language classes. However, incorporating learning training into teaching materials and

course content could yield better gains in the long term (Cotterall, 2000; Wenden, 1987). In order to make learner training an essential part of a course, learners could be given questionnaires focusing on different aspects of foreign language learning such as attitudes, learning styles and strategies, needs analysis and goal-setting, motivation and so on throughout the course. In addition, peer or whole-class discussions, learner journals, counselling sessions, and diary-keeping could be carried out so that learners could find the opportunity to reflect on various aspects of learning and also benefit from others' ideas and experiences. Furthermore, learners could be assisted and guided to assess and evaluate their own performance and progress, set realistic goals and also determine the ways of achieving them. To sum up, learners' metacognition could be developed well in a number of ways so as to encourage them to become more motivated and effective learners.

The present study calls for further research for examining the construct of motivation more thoroughly with respect to learner training. The fact that there was not a noteworthy change in the level of participants' motivation at the end of the training sessions raises the ultimate question as to whether such training is necessary or should be given at all. To respond to this question, several perspectives should be considered tentatively. Firstly, in this particular study, the initial level of students' motivation prior to the training sessions was already high. Therefore, to make generalisations according to the present results might cause misunderstanding. For this reason, this study could be replicated with a larger number of participants, preferably having a low level of motivation and the results could be compared. Secondly, as is well known, motivation is affected by several factors including personal ones such as age; academic success level; educational, cultural and family background, etc. and several contextual ones such as timing, duration of the training, intensity of teaching programme and so on. As one of these, metacognition, i.e. planning learning, setting goals, evaluating self-progress, etc. is certainly helpful and the findings from the follow-up interview of the study are also in support of this fact. However, training learners to become aware of themselves as learners and learning process should be taken as part of a whole scheme of language learning. In other

words, all the other factors, or at least a majority of them, such as effective teachers, effective materials, and effective instruction should be provided so that motivation in learning can be greatly enhanced. Therefore, future research studies might look into the correlations, if any, between learning training, motivation and the factors mentioned above by controlling them more cautiously.

8. References

Benson, P. (2001). *Teaching and researching autonomy in language learning***.** Essex: Pearson Education Limited.

Boydak, A. (2001). *Öğrenme stilleri*. İstanbul: Beyaz Yayınları.

Brown, H. D. (2001). *Teaching by principles: An interactive approach to language pedagogy.* New York: Longman.

Cotterall, S. (1999). Key variables in language learning: What do learners believe about them*?*. *System, 27*, 493-513.

Cotterall, S. (2000). Promoting learner autonomy through curriculum: Principles for designing language courses. *ELT Journal 54*(2), 109-117.

Demir, B. (2004). *An investigation into the effects of motivational factors and of primary school students on learning English as a foreign language.* Unpublished. *attitudes* M.A thesis. Çanakkale: Çanakkale Onsekiz Mart University.

Dickinson, L. (1993). Talking shop: Aspects of autonomous learning. *ELT Journal 47*(4), 330-336.

Dickinson, L. (1995). Autonomy and motivation: a literature review. *System 23*(2), 165-174.

Dörnyei, Z. (2001). *Teaching and researching motivation*. Essex: Longman.

Dursun, E. (2007). *An investigation into reasons of gender differences in foreign language learning process at university level prep classes.* Unpublished M.A. thesis.

Çanakkale: Çanakkale Onsekiz Mart University.

Ersöz, A. (2004) Key to effective teaching: Intrinsic motivation. In G. Durmuşoğlu-Köse (Ed.), *Proceedings of the 5th International Inged-Anadolu ELT Conference* (pp.51-54). Eskişehir, Turkey.

Esch, E. M. (1997). Learner training for autonomous language learning. In P. Benson & P. Voller (Eds.), *Autonomy and independence in language learning* (pp.164-175). London: Longman..

Finch, A. E. (1998). Designing and using a learner journal for false beginners: Self-assessment and organization of learning. *Self-Access Language Learning Newsletter,* Retrieved June 25, 2006 from http://lc.ust.hk/HASALD.

Gower, R., Phillips, D. & Walters, S. (1995). *Teaching practice handbook*. Oxford: Heinemann.

Hedge, T. (1993). Key concepts in ELT. *ELT Journal 47*(2), 92

Hedge, T. (2000). *Teaching and learning in the language classroom.* Oxford: Oxford University Press.

Lee, I. (1998). Supporting greater learner autonomy in language learning. *ELT Journal 52*(4), 282-291.

Logan, S. & Moore, N. (2004). Implementing learner training from a teacher's perspective. In H. Reinders (et al) (Eds.), *Supporting Independent Learning in the 21st Century: Proceedings of the Independent Learning Association Conference*, the University of Melbourne, Australia.

McCarthy, C. P. (1998). Learner training for learner autonomy on summer language courses. *The Internet TESL Journal 3*(7). Retrieved August 12, 2006 from http://www.aitech.ac.jp/~iteslj/

Mynard, J. (1999). *Motivation for learning English among first year female university students in Abu Dhabi*. Unpublished Ph. D. thesis. Exeter: The University of Exeter.

Nunan, D. (2002). Learner strategy training in the classroom: an action research study. In J. C. Richards & W. A. Renandya (Eds.), *Methodology in language teaching* (pp.133-143). Cambridge: Cambridge University Press.

Salehi, M. (2005). The relationship between intrinsic motivation and learner

autonomy. *Language Forum 31*(2), 1-17.

Salem, N. M. (2006). *The role of motivation, gender and language learning strategies in EFL proficiency*. Unpublished MA thesis. Lebanon: American University of Beirut.

Sheerin, S. (1997). An exploration of the relationship between self-access and independent learning. In P. Benson & P. Voller (Eds.), *Autonomy and independence in language learning* (pp. 54-65). London: Longman.

Şencan, H. (2005). *Sosyal ve davranışsal ölçmelerde geçerlik ve güvenirlik***.** Ankara: Seçkin Yayınevi.

Victori, M. & Lockhart, W. (1995). Enhancing metacognition in self-directed language learning. *System, 23*(2), 223-234.

Wenden, A. (1987). Incorporating learner training in the classroom. In A. Wenden and J. Rubin (Eds.), *Learner strategies in language learning* (pp. 159-168). New York: Prentice Hall International.

Wenden, A. L. (1998a). Learner training in foreign/second language learning: a curricular perspective for the 21st century. *ED 416 67.*

Wenden, A. L. (1998b). Metacognitive knowledge and language learning. *Applied Linguistics, 19*(4), 515-537.

Williams, M. & Burden, R. L. (1997). *Psychology for language teachers.* Cambridge: Cambridge University Press.

Yang, N. (1998). Exploring a new role for teachers: promoting learner autonomy. *System, 26,* 127-135.

Appendix 1: Dimensions of motivation and corresponding questionnaire items

Dimensions of motivation	**Item Numbers**
Extrinsic motivation	8. I will need to know English in the future. 9. It will be important for me to know English in the future. 10. The main reason I learn English because I have to. 11. I need to learn English for my future career. 14. I want only to survive the English lesson. 16. It is not worth learning English because it takes a long time. 17. If I had the choice I'd give up learning English.
Intrinsic motivation	1. I like English. 2. I would like to visit an English-speaking country. 3. I would like to be able to speak English. 4. I enjoy English lessons. 5. I find English interesting. 6. I want to do well in English class. 7. I would like to meet English-speaking people. 12. When I learn new things in English, I feel satisfied. 13. I would like to learn English even if I didn't have to. 15. I find learning English enjoyable. 18. Learning English is important for my personal development. 19. It makes me happy to think that I learn English.

Appendix 2: The plan of a learner training session

Topic(s)	1.Learning styles 2.Language learning strategies
Aims	1. Developing an awareness of learning styles and language learning strategies 2. Allowing students to get an understanding of their own styles and strategies
Date	20th December 2005
Duration	50 minutes
Place	ZF. 05
No. of studs.	30
Materials	1. Learning style inventory (comprised of items related to perceptual modalities: Auditory Visual, Kinesthetic/Tactile, adapted from Boydak, 2001) 2. Turkish version of SILL (Strategy Inventory for Language Learning, Oxford, 1990) (translated by Dursun, 2007) 3. A form to be used for reflection at the end of the session
Activities	**I. Learning styles** 1. Completion of 'Learning style inventory' 2. Whole-class discussion on different learning styles **II. Language learning strategies** 1. Brainstorming about different strategies that students use in learning English 2. Completion of SILL (Strategy Inventory for Language Learning) 3. Whole-class discussion about different strategies employed by students
Procedures	*1. Greeting and establishing rapport:* The teacher and students greet each other. **I. Learning styles** *2. Lead-in:* The teacher gives an example of two individuals' learning styles and draws attention to the difference between them. Then she asks several students how they prefer to learn in general. *3. Main-activity:* a) A copy of 'Learning style inventory' is distributed to each student. Students complete the inventory and find out their own learning style. b) Whole-class discussion takes place about the students' preferred learning styles and explanations are made by the teacher when necessary. **II. Language learning strategies** *2. Lead-in:* The students brainstorm about different strategies they employ when learning vocabulary (here given as an example language area to be studied). Some of the responses are written on the board. *3. Main-activity:* a) The students are instructed about how to complete the form of SILL and calculate the points in order to find out their preferences in using different categories of strategies. b) The teacher and the students extend the discussion on various strategies and how effective they are found by particular students. *4. Reflection:* The students write answers to these questions: • What have I learned during this training session? • How and where can I use what I have learned? *5. Wrap-up*: The teacher and the students review what they covered in the session.

ASIAN EFL JOURNAL
The EFL Professional's Written Forum

What Item Response Theory (IRT) Can Reveal to Us:

An Analysis of a Twenty-Item Vocabulary and Structure Test

Zhang Jianmin
Chen Zhiteng
Xiao Xi

Bio Data:

Zhang Jianmin is an associate professor at the School of International Studies, Zhejiang University, P.R. China.
Chen Zhiteng is a senior teacher of English at Longshan High School, Ruian, Zhejiang Province, P.R.China.
Xiao Xi is a primary teacher of English at Lingxi No.2 Middle School, Cangnan, Zhejiang Province, P.R.China.

Abstract

A language test serves two basic functions: 1) it tries to measure the true language ability of a student; 2) it aims to evaluate classroom teaching. Based on the results of an English test given to one class at a high school, this paper aims to answer two questions: 1) are the scores of some students from the test compatible with their regular performance in English? 2) is the test good enough to give us useful and reliable information about the test and the test takers? By comparing the test scores with those from other tests and by using the Rasch model, the authors find that IRT shows a very strong capacity in interpreting the test scores and predicting the language ability of the students in question.

Keywords: language test, scores, item response theory, interpret, revelation

I. Introduction

The most fundamental objective of administering a language test to students and other learners is to measure their language ability though technically and theoretically this has often proved difficult. In analyzing and interpreting test results, two testing theories are normally used. One is Classical Test Theory (CTT) and the other is Item Response Theory (IRT). CTT is based on the true score theory in that it is assumed that the observed score is composed of the true score and the error score. The observed score is usually seen as an estimate of the true scores of that test-taker

plus/minus some unobservable measurement error (Crocker & Algina, 1986; Hambleton & Swaminathan, 1985). According to many researchers in language testing, CTT was the leading framework for analyzing and developing standardized tests. Since the beginning of the 1970's, IRT has largely replaced the role CTT had and is now the major theoretical framework used in this scientific field (Crocker & Algina, 1986; Hambleton & Rogers, 1990; Hambleton, Swaminathan, & Rogers, 1991). One of the major weak points of CTT is that it is sample-dependent (sample of test takers here), that is, item parameters are obtained by calculating the mean of the items in question and their correlation coefficients with the whole test scores. In other words, since the sampled test takers are different in levels of ability, then such indices as facility values and discriminatory powers of test items will be different. As a consequence, it is almost impossible to determine the norm with which to measure the language ability of test takers unless we can get a fairly accurate scale or yard stick for such use. If for instance, the test-takers with different ability levels take a test, scores will be wildly different. Neither an easy test nor a difficult one can distinguish or discriminate the test takers. Therefore, it is difficult to compare test-takers' results between different tests. Item Response Theory is based on an entirely different philosophy of psychometrics in that

> in practical test development work, we need to be able to predict the statistical and psychometric properties of any test that we may build when administered to any target group of examinees. We need to describe the items by item parameters and examinees by examinee parameters in such a way that we can predict probabilistically the response of any examinee to any item.(Lord, 1980, p.11).

In other words, from the response patterns we can observe from those test takers to the items, we can judge fairly accurately and independently both item performance and the ability of the test taker. Certainly, we have to make sure, as Wright and Stone (1979) put it, each person's response pattern must be assessed to determine whether the person was responding in an acceptably predictable way given the expected hierarchy of responses (i.e. the items are arranged in order to form a hierarchy).Unlike CTT, therefore, we do not need to stress on reliability or validity of a test. We concentrate rather on the response pattern of individual test takers. According to Bond

and Fox (2001, p.8), the Rasch model provides us with useful approximations of measures that help us understand the processes underlying the reason why people (test takers included) and items behave in a particular way.

IRT does not need large numbers of test items to obtain scores of test takers to determine their language abilities. With ten calibrated test items, the abilities of some test takers can be revealed by referring to the response patterns they show on these items. (By some, we mean we only chose 20 test takers out of 45 in our research. By doing so, we hoped to avoid complicated calculations of the scores from all the 45 test takers.) In this sense, the latent ability of a test-taker is independent of the content of a test. The relationship between the probability of answering an item correctly and the ability of a test-taker can be modeled in different ways depending on the nature of the test (Hambleton et al., 1991). According to IRT, a test-taker with high ability should have a high probability of answering an item correctly. Theoretically, if the items that are used to test the ability of a test taker are arranged according to their facility values, then we can expect to see an orderly response pattern of that test taker. That is, the test taker is not likely to answer correctly the items beyond his/her ability. Even if he or she gets the items correct, the response pattern is disorderly. From this disorderliness, we can judge that guess work is involved in doing the test.

IRT seeks to reveal latent psychological constructs in terms of observable item responses. This information is useful in developing and evaluating tests, as well as interpreting examinees' scores on the latent characteristics in question. If with only a small number of test items we could know a lot about the test takers, the advantages of IRT will show themselves as are revealed by Henning:

> If only those items are used that approximate in difficulty the known ability region of the examinees, then fewer items will be required. … If items are pre-calibrated, banked and randomly summoned for any given measurement task, then there is less risk of a security breakdown that would disqualify large numbers of items for future use. All these advantages app up to greater economy of items over time and use (Henning, 2001, p.111).

And this is what we want to reveal from the results of a test by using the one

parameter model or the Rasch Model. The reason for our choice of the Rasch One-Parameter Model is that for one thing, we are more concerned with the parameter: scale of person ability and item difficulty; for another, this model is not constrained by sample size (size of test takers). For the two parameter model, a sample of more than 200 test takers is required and the three parameter model would need a sample size of more than 1000 (Henning, 1987, p.116). The objective of our research here is very simple: we intend to find out if the scores from the test are compatible with: 1) the regular performance of these test takers; 2)with the assessment of their language teachers. At the same time we want to test if IRT is powerful in making sound judgment of test takers' abilities or of the quality of a test. Moreover, we also attempt to see if IRT could make better interpretations of test results as well as those or any test takers.

II. Methods and Materials

1. Subjects and Test Items

A total of 45 students from Longshan High School in Zhejiang Province, China took an English test. They were 17 to 18 years of age at the beginning of the school year in September 2006. The given test, officially known as Final English Test Paper for 10 Schools in Wenzhou, Zhejiang Province for the first term of 2006 School Year, contains four sections: listening comprehension, reading comprehension, vocabulary & structure and essay writing. The participants were aware that the test was given just as a quiz. It was neither a mid-term test, nor a final test. They were told that the results of the test would not be counted to form the final grade for the subject of English. The English of these students for our study was below the average of their peers, but they showed willingness and cooperation in taking the test. For easy calculation as is usually the case, we selected 20 students from the 45 ones that took the test, 10 in the upper group and 10 in the lower group according to their scores in the test in question. Furthermore we only used the scores of the vocabulary and structure part for our analysis and discussion. We chose this part of the test for two reasons: first, the items are context-independent of one another; secondly, they cover

a wider range of language points to be tested. That part contains 20 multiple-choice items and each item counts one point. So the full score of this section is 20.

Table 1 Scoring Matrix for a 20-Item Vocabulary and Structure Test

Item / person	17	9	3	15	20	4	12	16	18	1	7	8	10	11	5	2	6	14	13	19	score	
A	1	1	1	1	1	1	1	1	0	1	0	1	1	1	0	0	1	0	0	0	13	Upper Group
B	1	1	1	1	1	0	1	0	1	1	1	1	1	1	0	1	0	0	0	0	13	
C	1	1	0	1	1	1	0	1	1	0	1	1	0	1	0	0	1	1	0	1	13	
D	1	1	0	1	1	1	0	1	1	1	0	1	1	0	0	0	0	1	1	0	12	
E	1	0	1	1	1	1	1	1	0	1	0	0	1	0	0	1	0	0	0	0	10	
F	1	1	0	1	0	1	1	1	1	0	0	1	1	1	0	0	0	0	0	0	10	
G	1	1	1	1	0	0	0	1	1	1	1	0	1	0	1	0	0	0	00	0	10	
H	1	0	1	1	1	0	0	1	0	1	1	1	0	0	0	0	0	1	0	0	9	
I	0	0	1	1	0	1	1	0	1	1	1	0	0	0	1	1	0	0	0	0	9	
J	1	1	1	0	1	0	1	0	0	0	0	0	0	1	0	0	1	0	1	0	8	
K	1	1	1	0	0	1	0	1	0	1	0	0	1	1	1	0	0	0	0	0	8	Lower Group
L	1	0	0	1	0	1	1	0	0	0	1	0	0	0	0	1	1	0	0	0	7	
M	0	1	1	0	1	1	0	0	1	0	0	0	0	0	1	0	0	0	0	0	7	
N	0	1	1	1	0	0	0	0	0	0	1	0	1	0	1	0	1	0	0	0	7	
O	0	1	1	0	1	0	0	0	1	0	0	1	0	1	0	0	0	0	0	0	6	
P	1	1	0	0	1	0	0	0	0	0	0	1	0	0	1	1	0	0	0	0	6	
Q	1	1	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1	0	0	5	
R	1	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	4	
S	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	3	
T	1	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	3	
total	16	15	12	11	10	9	9	9	9	8	8	8	8	7	6	5	5	4	2	2		

2. Data Analysis

The selection of 20 students as subjects and their scores of the 20-item test as data for analysis of our study were made by referring to their regular performances in similar English tests over the past school year. These students were then arranged in order according to their scores of the 20-item test as shown in Table 1 above. Here 1 stands for correct choice and 0 for wrong choice. The persons or items for which all responses are correct or incorrect are usually eliminated, because they do not have any discrimination power, but none of the examinees or items in our case belonged to this category.

With this matrix ready, we can now calculate the logit incorrect value for each

possible number correct. We only need to follow the procedures proposed by Henning (1987, p.119) in doing the calculation. The logit incorrect value for each item was computed as the natural logarithm of the ratio of the proportion incorrect to the proportion correct. In fact, we do not need to do such tedious calculation. By consulting Table F in Appendix A (Henning,1987, p.171), we can easily get the logit correct and logit incorrect values. This procedure places item difficulties on an interval scale and eliminates the boundaries inherent in the zero-to-unity range of classical p-values. The zero point, or origin, of the item difficulty calibrations is arbitrarily set at the mean of the logit incorrect values for all analyzable items. This is done by subtracting the mean adjustment value, computed as the sum of the frequency times logit values divided by the sum of the frequencies, from the logit incorrect value for each item (Henning 1987, p.119). We do so in order to calculate the difficulty level of the items as shown in Table 2 below.

Table 2 provides us with information about the items. For example, we note in the second column1 6 students out of 20 got Item 17 correct, so the proportion correct is 0.8 (16/20), while the proportion incorrect is 0.2. But Item Frequency is 1, which means only one item (Item17) was done correctly by 16 students. Logit Incorrect can be obtained by consulting Table F in Appendix A (Henning,1987, p.171). Finally the Initial Item Difficulty can be calculated easily.

Table 2 only reveals information about item difficulty. To know more about the 20 students' performance on this test, we need to calculate the initial person ability by using logit correct values instead of logit incorrect values. This is because we are effectively subtracting item difficulties from person abilities in order to place both these estimates on the same single ability-difficulty continuum (Henning,2001, p.120). Table 3 shows the result of this calculation.

Table 2 Calculation of Item Difficulty Calibrations

item name	number correct	item freq.	prop. corr.	prop. incor.	logit incor.	freq.x logit	freq.x $logit^2$	item diff.
17	16	1	16/20=0.80	0.20	-1.39	-1.39	1.93	-1.83
9	15	1	15/20=0.75	0.25	-1.10	-1.10	1.21	-1.54
3	12	1	12/20=0.60	0.40	-0.41	-0.41	0.17	-0.85

15	11	1	11/20=0.55	0.45	-0.20	-0.20	0.04	-0.64
20	10	1	10/20=0.50	0.50	0.00	0.00	0.00	-0.44
4,12,16,18	9	4	9/20=0.45	0.55	0.20	0.80	0.16	-0.24
1,7,8,10	8	4	8/20=0.40	0.60	0.41	1.64	0.67	-0.03
11	7	1	7/20=0.35	0.65	0.62	0.62	0.38	0.18
5	6	1	6/20=0.30	0.70	0.85	0.85	0.72	0.41
2,6	5	2	5/20=0.25	0.75	1.10	2.20	2.42	0.66
14	4	1	4/20=0.20	0.80	1.39	1.39	1.93	0.95
13,19	2	2	2/20=0.10	0.90	2.20	4.40	9.68	1.76
Total		20				8.79	19.31	

Logit Incorrect = Ln□Proportion Incorrect/Proportion Correct)

Mean Adjustment =□Sum of Frequency x Logit)□/□Sum of Frequency)

Initial Item Difficulty = Logit Incorrect - Mean Adjustment, e.g. mean adjustment = 8.79/20 = 0.44

Table 3 Calculation of Initial Person Ability Measures

persons	person score	person freq.	prop. corr.	logit corr.	freq.x logit	freq.x $logit^2$	person measure
T,S	3	2	3/20=0.15	-1.74	-3.48	6.06	-2.16
R	4	1	4/20=0.20	-1.39	-1.39	1.93	-1.81
Q	5	1	5/20=0.25	-1.1	-1.10	1.21	-1.52
P,O	6	2	6/20=0.30	-0.85	-1.70	1.45	-1.27
N,M,L	7	3	7/20=0.35	-0.62	-1.86	1.15	-1.04
K,J	8	2	8/20=0.40	-0.41	-0.82	0.34	-0.83
I,H	9	2	9/20=0.45	-0.20	-0.40	0.08	-0.62
G,F,E	10	3	10/20=0.50	0.00	0.00	0.00	0.42
D	12	1	12/20=0.60	0.41	0.41	0.17	0.83
C,B,A	13	3	13/20=0.65	0.62	1.86	1.15	1.04
Total		20			-8.48	13.54	

Logit correct = Ln□Proportion Correct/Proportion Incorrect)

Mean Adjustment =□Sum of (Frequency x Logit)□/□Sum of Frequency□

Initial Person Measures = Logit Correct Values - Mean Adjustment

For example: Mean Adjustment = -8.48/20 =-0.42

III. Findings and Discussion

The scoring matrix for a 20-item multiple choices test depicts (Table 1) that no

students obtained a score of less than 3 or more than 13. We could say therefore that the actual range of scores was 3-13, or 10, the possible range being 0-20, or 20. Since this was a multiple-choice type of test with four options, one would expect the examinees to get a sore of at least 25 percent by mere guessing. Thus, a score of 5 or below would actually be meaningless for discriminating among the participants in the ability being tested. Hence, we conclude that the effective range of this test was 5--13, or 8. Generally speaking, the broader the range, the more effective in discriminating among examinees on the ability under consideration will be. From Table 1, we know that some items are too difficult for these students (Items 13, 14, 19), and the level of several students is too low. We need to determine whether there is some correlation between the two.

The response patterns of some persons as revealed in Table 1 are informative. The response pattern of Person R, for example, is typical of a person who was making wild guesses, for he exhibited a highly unlikely response pattern. This is especially so since with an estimated ability -1.81, this person missed Item 3 and got Item 7 correct. He was one of the two who got Item 19 correct which has a difficulty level of 1.76. He failed in Item 3 with only a difficulty level of -0.85. By referring to his disorderly responses to the items, it is natural for us to come to this conclusion that guess work was involved in making the choices. This was later verified after we interviewed Person R. Consider Person Q. He only got five choices correct and one of them was Item 14 which has a difficulty level of 0.95. Referring to his ability measure, we know it is -1.52, meaning it is far below his ability to get Item 14 correct. The only possible conclusion is that he got the correct choice by mere guessing because he failed to correctly answer much easier items such as Items 3, 15, 20, 4, 12. Here what we need to point out is that CTT makes no provision for the identification of such persons and, therefore, increases the possibility that invalid scores may be reported for certain persons in the sample of examinees. IRT overcomes this weakness and can reveal this information to us.

To verify if the ability of these 20 test takers matches their regular performance, we have checked with their regular record and ranks among 45 students in the whole

class. We also checked with the assessment of these 20 test takers by their English teacher, but we did not tell the teacher that we were doing this kind of research for fear that the facts he gave us might be contaminated. That is, to avoid the facts he gave us were in our favor. The regular record of their English performance and their assessment by the teacher match very well with the scores they obtained from the 20-item test as is shown in Table 4 below.

This 20-item test is of good quality, except that some items are too difficult for these students. For example, for items 13 and 19, the average final calibration is 0.06, which means its calibration is appropriate. It is equal to 0.5-0.6 according to the classical test theory (CTT). However, the level of the participants of the test is comparatively low. The average of the final measure is – 0.45, which is far below 1.

Table 4 Matrix of Scores from Four Tests and the Testees' Ranks

test / persons	20-item test		mid term exam		final exam		monthly exam		
	rank	score	rank	score	rank	score	rank	assessment	
A	1	13	1	106	3	98	4	102	Upper Group
B	1	13	2	104	2	102	3	105	
C	1	13	3	103	4	97	6	96	
D	2	12	4	101	1	106	2	107	
E	3	10	5	99	5	95	7	95	
F	3	10	6	99	1	106	5	97	
G	3	10	7	96	6	83	10	81	
H	4	9	8	94	8	83	8	91	
I	4	9	8	94	6	92	12	75	
J	5	8	9	92	7	85	9	87	
K	5	8	10	59	9	70	14	65	Lower Group
L	6	7	11	58	11	62	17	40	
M	6	7	12	56	10	69	18	39	
N	6	7	13	54	13	52	1	112	
O	7	6	14	48	15	48	11	80	
P	7	6	15	47	11	68	15	61	
Q	8	5	16	46	15	48	16	45	
R	9	4	17	45	14	50	18	39	
S	10	3	18	45	12	54	13	69	
T	10	3	19	38	16	29	19	37	

Conclusions

From the above discussion, we may now come to these conclusions:

1) IRT is more powerful in revealing the latent traits of the test takers;
2) IRT is more economical so far as language testing is concerned because it does not need too many items to measure the true language ability of test takers;
3) From the data of other tests given to the same group of students, we can judge that classroom teaching is going on smoothly as the participants of our study show fairly compatible results of their English studies.

Notwithstanding, this kind of study needs to be widely replicated if we are to safely say that IRT is a powerful tool to measure and assess language teaching and testing.

References

Baker, F. B. (1992). *Item response theory parameter estimation techniques*. New York: Marcel Derker, Inc.

Baker, F. B. (2001). *The basics of item response theory* (second edition). ERIC Clearinghouse on Assessment and Evaluation

Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F.M. Lord and M.R. Novick (Eds.), *Statistical theories of mental test scores* (pp. 397-472). Reading, MA: Addison-Wesley,

Bond, T. G. & Fox C. M. (2001). *Applying the Rasch model*. New Jersey: Lawrence Erlbaum. Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Holt, Rinehart and Winston.

Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: Principles and applications*. Boston: Kluwer-Nijhoff.

Hambleton, R. K., & Rogers, J. H. (1990). Using item response models in educational assessments. In W. Schreiber & K. Ingenkamp (Eds.), *International developments in large-scale assessment* (pp. 155-184). England: NFER-Nelson.

Hambleton, R. K., H. Swarmnathan and H. J. Rogers. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage.

Harvey R. J. & Hammer A. L. (1999). Item response theory. *The Counseling Psychologist*, *27*(3), 353-383

Henning, G. (2001). *A guide to language testing: Development, evaluation and research*. Foreign Language Teaching and Research Press & Heinle /Thomson Learning Asia

Jaeger, R. M. (1991). *Series Editor's Foreword*. In Hambleton, R. K., H. Swarmnathan and H. J. Rogers. (1991). *Fundamentals of item response theory.*

Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. NJ: Lawrence Erlbaum.

Wright, B. & Stone, M. (1979). *Best test design*. Chicago: MESA Press.