

Impact of Metacognitive Awareness on Self-Assessment Skills of Lower-Intermediate Level EFL Students

Fatemeh Hemmati¹ & Amirali Mohammadkhani²

¹Payame Noor University of Tehran, fatemehhemmati2002@yahoo.co.uk

²Corresponding author, Payame Noor University of Tehran, aa.m.khani@gmail.com

Received: 29/09/2014

Accepted: 07/06/2015

Abstract

This study focuses on the relationship between lower-intermediate level EFL learners' metacognitive awareness and accuracy in self-assessment of a speaking test. We tried to raise the learners' metacognitive awareness through practice in goal-setting and planning. To do so, 103 lower-intermediate level students took a pretest of speaking, completed a metacognitive awareness questionnaire, and had an immediate recall interview. Next, they were divided into 2 groups and attended a course in which the students in 1 group received goal-setting and planning treatments. After administration of the posttests, the findings revealed that the goal-setting and planning treatment led the learners to gain higher levels of metacognitive awareness, $t(101) = 2.45$, $p = .019$, and this increased awareness could increase their accuracy in self-assessment. Such findings indicate that presenting the objectives of different tasks to learners and goal-setting can be applied in language learning classes to enhance their awareness and improve their self-assessment skills.

Keywords: Self-Assessment; Metacognitive Awareness; Planning; Goal-Setting

1. Introduction

It has been argued that tests, by nature, are artificial contexts for language use (Spence-Brown, 2001), and no matter how reliable and valid they are, tests make test-takers follow rules which are different from nontest interaction (Spolsky, 1985). In accordance with these opinions, Cohen (1998) believes that as long as the task is part of a test, students are likely to use strategies that they would not use under nontest conditions. Based on these ideas, it can be suggested that traditional tests lack the merit of measuring the knowledge and skills of language learners with high degrees of validity and reliability. This deficiency must be addressed in some ways, and it indeed has received some attention in the field of language testing under the heading of dynamic and alternative assessment.

Dynamic assessment is a continuing diagnostic, prescriptive approach to instruction (Schneider & Ganschow, 2000). It causes a continuous interaction between teacher and learner. Through this interaction learners can discover solutions

to learning problems. The concept of dynamic assessment originated with Reuven Feuerstein (1980), who believed in the facilitative and mediating role of the instructor (Schneider & Ganschow, 2000).

Close to the concept of dynamic assessment is alternative assessment. Huerta-Macías (2002) describes alternative assessment as an alternative to standardized testing and all of the problems found with it. She believes that there is no single definition for describing this concept and many labels can distinguish it from traditional testing. García and Pearson (1994) categorized the following terms under the heading of alternative assessment: performance assessment, authentic assessment, portfolio assessment, informal assessment, situated assessment, and assessment by exhibition.

Brown and Hudson (1998), however, argued that to speak of alternative assessment is counterproductive because some may imply that it can act as an excuse to refrain from responsible test construction. Therefore, they proposed the term *alternatives* in assessment. According to this notion, assessment is not only test construction, and other alternatives like the ones mentioned in García and Pearson (1994) can be used for this purpose (Brown, 2004).

Needless to mention, García and Pearson's list (1994) was far from complete. One of the techniques and approaches in alternative assessment which was not included in García and Pearson's (1994) list is the use of self-assessment. Self-assessment is defined as a process by which learners' performance is evaluated by themselves (McNamara, 2000). Heilenman (1990) argued about the importance of self-assessment stating that tests can measure a small portion of learners' language use, and learners themselves can be aware of a wide range of their resources. This advantage of self-assessment over other forms of testing caused researchers to ponder on this alternative in assessment (e.g., Andrade, Du, & Wang, 2008; Pounder, 2000; Ross, 1998; Sullivan & Lindgren, 2002).

2. Literature Review

2.1. Self-Assessment

Learners usually view assessment of language skills from two distinct points. The first one is assessment in the form of examinations and the other one is in the form of self-reports and self-assessment (Oscarson, 1989). Self-assessment is justified in this way that cooperation of learners in the process of evaluation provides some affective support for the learners (Oscarson, 1989). Furthermore, Oscarson (1989) stated the rationale for self-assessment as follows: (1) Promotion of learning which, he believed, would be achieved by learners' engagement in the process of evaluation. (2) Raised level of awareness: This may happen due to systematic questions and answers about the content of the course. (3) Improved

goal-orientation: Self-assessment can cause learners to think about a variety of goals for themselves. (4) Expansion of range of assessment: Learners' involvement in evaluation process may include aspects which could not be assessed by an outsider. (5) Shared assessment burden: This is not empirically proved. (6) Beneficial postcourse effect which is autonomy of learners.

In line with Oscarson's ideas, Harris (1997) considered self-assessment as one of the pillars of learner autonomy. He believed that one of the essential elements of self-directed language learning is the opportunity for learners to assess their own progress. Self-assessment can help learners become more active and realize that they have the ultimate responsibility for learning. He also remarked that self-assessment can help learners know their own strengths and weaknesses and encourage them to think about their needs in order to improve their performance (Harris, 1997).

Harris (1997) continued with this idea that systematic self-assessment enables learners to look back and assess their own progress holistically. Comparison of learners' self-assessment scores and final exam scores can affect learners in their choice of future goals. Finally, systematic self-assessment facilitates learners' development in other activities like organizing and planning learning, thinking about learning styles, and discussion of learning and communication strategies.

These statements about usefulness of this approach to assessment made researchers conduct studies on this topic. For example, Ulicsak (2004) in his study about the development of children's group skills found that by explicit scaffolding, it is possible to reflect on collaboration and providing feedback using the children's self-assessment of these skills. His model incorporated existing training schemes, including procedural prompting, assigning roles, modelling exchanges, and giving feedback. The results showed that reflection on collaboration and self-assessment of skills may improve the recall of the material studied. In another study, Patri (2002) investigated the agreement of teacher-assessment, self-assessment and peer-assessment of ethnic Chinese students in the presence of peer feedback. The study used a self- and peer-assessment questionnaire containing 14 items related to the organization of the presentation content, use of language and interaction with the audience. The findings revealed that with firmly set criteria, peer-assessment is similar to teacher-assessment. However, self-assessment is not as accurate as peer-assessment.

The literature in this area of study contains numerous studies like the abovementioned ones showing conflicting results. Malabonga, Kenyon, and Carpenter (2005) conducted two studies investigating technical aspects of a computer-mediated test, the Computerized Oral Proficiency Instrument (COPI), particularly in contrast to a similar tape-mediated test, the Simulated Oral Proficiency Interview (SOPI). In their first study, 55 university students took the

COPI and SOPI in one of three languages: Arabic, Chinese, or Spanish. The researchers sought to find how participants used self-assessment to choose an appropriate starting level on the COPI. The results showed that 92% of the examinees were able to use the self-assessment instrument to select test tasks at appropriate difficulty levels.

In a study about the promotion of self-assessment skills in writing computers were used to record all the processes during writing and in a peer session the recorded process was elaborated by every individual learner. This process caused different impacts on the learners which were all reported to be influential in learners' writing ability (Sullivan & Lindgren, 2002).

Andrade, Du, and Wang (2008) investigated the effect of assigning reading of a model written assignment and generating a list of criteria for the assignment on self-assessing according to a rubric. The instruments involved a model paper to scaffold the process of generating a list of criteria for an effective story or essay, receiving a written rubric, and using the rubric to self-assess first drafts. The results indicated the effect of treatment and of previous achievement on total writing scores, as well as effects on scores for the individual criteria on the rubric. These results suggest that using a model to generate criteria for an assignment and using a rubric for self-assessment can help elementary school students produce more effective writing.

In the Japanese context, proficiency tests and self-assessment measures were used to investigate whether subgroups of Japanese heritage language learners would demonstrate different language behaviours from those of traditional Japanese as foreign language learners and which domains of language use and skills would specifically identify such differentiation (Kondo-Brown, 2005). The findings indicated that there were striking similarities between the traditional Japanese as a foreign language learners group and 2 Japanese heritage language learners.

Employing the effectiveness criteria contained in the Competing Values Model of Organizational Effectiveness (Quinn & Rohrbaugh, 1981, 1983), Pounder (2000) suggested a set of organizational effectiveness self-rating scales for Hong Kong higher education institutions. The study underscored the qualitative benefits of the scale development procedure and resulting scales and notes that could be employed, and the qualitative benefits enjoyed, in other higher education systems.

Closely related to the current study, AlFallay's (2004) investigated the role of some psychological and personality traits of learners in the accuracy of self- and peer-assessments. The selected traits were motivation types, self-esteem, anxiety, motivational intensity, and achievement. Based on the results of the study, he concluded that assessment is a multifaceted process, which is affected by various

psychological and personality traits of the raters. He reported that learners with high classroom anxiety were more accurate in their assessment. The study also showed that long-term practice and sufficient feedback had a positive effect on the accuracy of assessment. Finally, the study showed that students with low self-esteem were the most accurate in assessing their performance, whereas learners with instrumental motivation were the least accurate.

Proponents of self-assessment believe that the notion that learners' own assessment of acquired skills is inherently unreliable is not supported by the literature (Oscarson, 1989). However, some studies showed a tendency among learners, especially less proficient learners, to overestimate their abilities. But there is something to add here about the argument. LeBlanc and Painchaud (1985) argued that learners evaluate their performance and think about *what* and *how much* they have learned, but the answers to these kinds of questions are not useful in terms of a graded instrument. This is so regardless of learners' level of proficiency, because if learners want to assess their linguistic ability based on an established model they must learn about all the components on that model like the one proposed by Bachman (1990) and the more proficient and less proficient learners are equally prone to this lack of knowledge. LeBlanc and Painchaud (1985) also argued that if the statements used for self-assessment are not abstract sentences, their scores highly correlate with proficiency tests. Harris (1997) citing Chamot and O'Malley (1994) stated that self-rating requires the student to exercise a variety of learning strategies and higher order thinking skills that not only provide feedback to the student but also provide direction for future learning.

With all that has been said, there are many factors that can affect the process of self-assessment. One of the factors not dealt with thoroughly is metacognitive awareness. Rivers (2001) distinguished between two types of metacognition: metacognitive self-assessment and metacognitive self-management. By the former he means having the ability to assess one's own cognition, and by the latter he means having the ability to manage one's cognitive development. He further added that self-assessment is a more critical skill than self-management.

With River's (2001) categorization of self-assessment under the heading of metacognition, and for the sake of proving its importance in the process of self-assessment, the next section of this paper is allocated to this concept.

2.2. Metacognition

Research has suggested that learners of different ages and varying proficiencies acquire some knowledge about learning, which influences their approach to learning and the expectations they hold about the outcome of their efforts (Wenden, 1998). Wenden defined this knowledge as *metacognitive*

knowledge. Citing Flavell and Wellman (1977), Wenden mentioned that metacognitive knowledge is the relatively stable information human thinkers have about their own cognitive processes and those of others. This knowledge can be explicitly stated and learners can become conscious of it and articulate what they know about it.

Wenden (1998) also cited different studies in the L2 literature and talked about other terms used to refer to metacognitive knowledge: learner beliefs (Horwitz 1987), learners' naive psychology of learning (Wenden 1987), and learner representations (Gremmo & Riley 1995; Holec 1987). Metacognitive knowledge has been classified as person knowledge, task knowledge, and strategic knowledge. According to Goh (1997):

Person knowledge consists of general knowledge learners have about how learning takes place and how different factors like age, aptitude, and learning styles can influence language learning. Person knowledge also includes what learners know about themselves as learners, and the beliefs they have about what leads to their success or failure in learning a language. Task knowledge refers to what learners know about the purpose, demands, and nature of learning tasks. It also includes their knowledge of the procedures that constitute these tasks. Strategic knowledge is what learners know about strategies. More specifically, it is knowing about which strategies are likely to be effective in achieving learning goals. It includes understanding how best to approach language learning. (p. 362)

Wenden noted that there might sometimes be overlaps between task and strategic knowledge because learners often needed to know what strategies to apply in order to accomplish a task (Goh, 1997).

Metacognition consists of two separate and distinct components of metacognitive knowledge and metacognitive strategies (Brown, Bransford, Ferrara, & Campione, 1983, as cited in Wenden, 1998). Metacognitive knowledge is defined as information learners have about their learning, whereas metacognitive strategies are general skills through which learners manage, direct, regulate and guide their learning. Examples are planning, monitoring and evaluating. According to Flavell (1979, as cited in Wenden, 1998), metacognitive knowledge is the key player in many cognitive activities like oral communication of information, oral persuasion, oral comprehension, reading comprehension, and writing which are language-related.

Wenden (1998) described a model of the affective learning process and believed when a task is assigned, the learners' perception of the task's demands causes an assessment process that is based on two categories of metacognitive knowledge and what they know about themselves as learners (person knowledge) and their perception of the task's purpose (task knowledge). The model suggests that all three kinds of metacognitive knowledge can be useful at the right time. Task knowledge allows learners to analyze a task in order to realize what is necessary to be done about it. Person knowledge enables learners to understand their abilities and preferences. Finally, strategic knowledge helps them in selecting strategies to deal with anticipated difficulties (Wenden, 1998).

On the effect of metacognitive awareness on L2 listening, Goh (1997) suggested that we can ask learners to describe the way they listen and what they know about being a second/foreign language listener. One way learners can report is diary writing where they record their observations, reactions, and perceptions. This 'listening diary' is different from more general diaries in which learners write about different aspects of language learning. In this way listening class should therefore be a place with lots of opportunities to share ideas. When students share information on their reflections and are fully aware of the various aspects of L2 listening, they will have a better chance to become autonomous listeners.

Vandergrift (1999), citing recent studies on the differences in strategy use between effective and less effective listeners, pointed to the potential role that metacognitive strategies can play in enhancing successful L2 listening. He stated that students should evaluate the results of their decision-making during a listening task and the teacher can encourage self-evaluation and reflection. This can be done by asking them to reflect on the strategies they use. Group or class discussions on different ways learners used to achieve success can also stimulate reflection and meaningful evaluation.

Mori, Sato, and Shimizu (2007), reflecting on studies in those years, claimed that individuals' choice of strategies for a challenging learning task reflects their views on language learning, in general, and the nature of a given task, in particular (Benson & Lor, 1999; Cotterall, 1995; Sakui & Gaies, 1999; Wenden, 1998; Yang, 1999).

Goh and Taib (2006), stating how to help learners to understand mental and emotional processes in their learning, believed that teachers can guide learners in discovering important aspects of the listening process and refer to awareness-raising achieved in this way as metacognitive instruction. They argue that metacognitive instruction includes both training learners directly to employ relevant strategies as well as helping them increase their metacognitive knowledge. Wenden (1998), in her work on learner autonomy, strongly advocated helping language learners

develop these areas of metacognitive knowledge in order to self-appraise and self-regulate their learning. She also added that teachers can encourage reporting and discussion of the thought processes that learners engage in during listening tasks in order to have greater metacognitive awareness.

One interesting point in Goh and Taib's (2006) study is their point on young learners. They claimed that young learners need more guidance from their teachers in mediating their perceptions about strategies and task demands. As children are not cognitively mature, they cannot assume multiple perspectives on many things, including the uses of specific strategies and how their application might change according to task demands.

In another study by Vandergrift and Tafaghodtari (2010), it was argued that application of metacognitive knowledge is a shared mental characteristic of all successful learners. They continue that metacognition accounts for a relatively high percentage of variance in learning performance. They also refer to an extensive body of evidence that learners' metacognition can directly affect the process and the outcome of their learning (Goh, 2008; Victori & Lockhart, 1995; Wenden, 1998), and that it is positively linked to motivation and self-efficacy (Dörnyei & Skehan, 2003; Vandergrift, 2005). They also point that learners with a high degree of metacognitive knowledge have the facility to apply knowledge and are better at processing and storing new information, finding the best ways to practice, and reinforcing what they have learned (Vandergrift & Tafaghodtari, 2010).

As it was said, Wenden (1998) classified metacognitive awareness as knowledge about the person, the task, and the strategies. This classification had its own value; however, Alexander, Schallert, and Hare (1991) classified the components of metacognitive awareness as self-knowledge, task knowledge, strategic knowledge and plans and goals. They insisted that this fourth dimension can drastically affect the other three components.

To our interest, Coutinho (2007) conducted a study and investigated two variables of achievement goals and metacognition. Coutinho believed that achievement goals are the types of outcomes students pursue in learning environments. Achievement goals are divided to two main types: mastery goals and performance goals. Mastery goals are related to learning and mastery of content and have been linked to adaptive outcomes such as strong self-efficacy, good metacognition, and good performance. Coutinho (2007) concluded that mastery goals and performance goals are related to academic success but only through metacognition. Thereby, students with mastery goals are predicted to have good metacognition, and this leads to academic success. Similarly, students with performance goals are expected to have poor metacognition, which translates to poor academic results.

Finally, in a comprehensive study on the effect of goal-setting and self-assessment, Kato (2009) provided these two techniques to groups of Japanese language learners in elementary, intermediate, and advanced levels. Elementary and intermediate students' comments at the end of the project showed that they preferred self-assessment activities and did not find goal-setting activity particularly beneficial. Based on his findings, he postulated that these learners appreciated the self-assessment activity because it helped them to overcome the challenges they faced in language learning. He also called for more research with more in-depth analysis investigating the effect of goal-setting and self-assessment on learners' language skills.

As it was mentioned by Rivers (2001), self-assessment is a metacognitive function and as Alexander et al. (1991) claimed, this function could be affected by goal-setting. However, Kato's (2009) study, as one of the few studies on the issues of goal-setting and self-assessment, showed that goal-setting could not prove to be effective for language learners. On the contrary, Coutinho (2007) links goals to academic success through metacognition. In order to shed more light on the effect of goal-setting on metacognitive awareness and its subsequent effect on self-assessment skills of learners, it is believed that if we could raise the awareness of learners through goal-setting and planning as one of the components of awareness and study its impact on self-assessment we can pave the way for the enhancement of this alternative in assessment. The current study addresses the following questions:

1. Is there any relationship between lower-intermediate level EFL learners' metacognitive awareness and their self-assessment on a speaking test?
2. Is there any difference between learners who are metacognitively aware and those who are not on their accuracy in self-assessment in a speaking test?
3. Does raising awareness through planning and goal-setting have any impact on other components of metacognitive awareness?

3. Methodology

3.1. Participants

One-hundred and three lower-intermediate level learners in four branches of Safir Language Academy in different parts of Tehran were chosen as the participants. Because the classes in this language school are segregated based on gender, the classes were held with only male or only female participants. The number of the male students was 49 and they formed four classes. The rest of the participants were female and they formed five classes. The experimental group, which was randomly chosen, included 52 students (28 males in two classes, 24 females in two classes) and the control group consisted of 51 students (21 males in

two classes, 30 females in three classes). The age range for students was 28 to 38. All the participants had academic education, and a few of them were postgraduate students. All the participants shared their L1 (i.e., Persian). The groups were homogenous in both measures of metacognitive awareness and accuracy in self-assessment. The difference between the scores obtained from the participants in measure of metacognitive awareness measured by Metacognitive Awareness Inventory (Schraw & Dennison, 1994, see Appendix) was nonsignificant as depicted in Table 2. The accuracy of self-assessment compared to teacher's assessment was also calculated through Pearson correlation which was not significant as displayed in Table 1.

3.2. Instruments

In order to measure the learners' speaking ability, two speaking sections from two sets of Cambridge Key English Test (KET) were chosen. The scores given to the students were considered to be from 1 to 25 based on the premise that the score of KET speaking section is 25% of the total score of the standardized exam.

To measure the learners' metacognitive awareness, Metacognitive Awareness Inventory (Schraw & Dennison, 1994) was used because it is the most useful instrument for measuring metacognitive awareness and has been used in other studies as well (Coutinho, 2007; Hamman & Steven, 1998). This questionnaire consists of 52 items and is categorized under two headings: knowledge about cognition and regulation of cognition. The section under the heading of "knowledge about cognition" has three subcategories of declarative knowledge, procedural knowledge, and conditional knowledge. The other section under the heading of "regulation of cognition" has five subcategories, which are planning, information management strategies, comprehension monitoring, debugging strategies, and evaluation. Due to the fact that the treatment duration was short, we used a 5-item Likert type in order to be able to capture the minor improvements in the learners' metacognitive awareness. The inventory was presented to the participants in English because we believed that the learners in this level were capable of comprehending the sentences used in the questionnaire. Moreover, to avoid any misapprehension, all the new words used in the inventory were pretaught to the learners before they took the inventory.

Also, for the sake of triangulation, an immediate recall interview was conducted. In these interviews, the learners were asked to reflect on the way they scored themselves, what reasons or criteria they had, and how fair they were in their scoring. Their answers were recorded and analyzed as self-reports for further insights.

3.3. Procedure

First of all, a pretest (i.e., KET speaking test) was administered to both groups, and the students were asked to score their performance holistically from 1-25. Then, in the immediate recall interview, we asked the students to talk about how fair they were in scoring themselves and what were their criteria in scoring. They were also asked to complete the metacognitive awareness questionnaire. The teacher, who was one of the researchers, also scored the performance of the students based on the scale provided for the test. The scores obtained from the self-assessment and teacher-assessment of speaking test and the learners' answers to metacognitive awareness questionnaire were compared. As Tables 1 and 2 show, no statistically significant difference was found in the two groups.

Following this phase, the students were put in two groups of experimental and control. In the experimental group, the learners were asked to set a goal for every session in class and estimate whether they achieved the goal or not at the end of the session. The teachers also asked about whether they knew what the objective of each task they did was. The enabling and terminal objectives of each activity in the class were also explained to them so they could compare their goals with these objectives. The duration of this treatment in the course was 12 90-min sessions that happened over one month. The learners in the control group had the same course with the same activities and tasks; however, no goal-setting activity was done in their classes, and the objectives of tasks were not explained to them. At the end of the course, both groups took a speaking section of KET and, similar to their pretest, were asked to score themselves holistically. The teacher also scored them based on the evaluation criteria for the exam. Immediately after each exam, the learners had a recall interview in which they explained the criteria with which they assessed themselves. In the end, Metacognitive Awareness Inventory (Schraw & Dennison, 1994) was presented to the participants and they completed it. The data collected from each instrument were analyzed, and the results are reported in the next section.

3.4. Data Analysis

The strength of association between oral interview self-assessment scores from both the control and experimental groups with teacher scores was measured using Pearson correlation. The reports from recall interviews were also analyzed for further insight. And, the scores of Metacognitive Awareness Inventory (Schraw & Dennison, 1994) were compared through *t* test on the pretest to the posttest. It must be mentioned that the alpha for achieving statistical significance in all statistical procedures was set at .05.

4. Results

The scores obtained from the learners' self-assessment and teacher-assessment of oral interviews on the pretest and the posttest are provided in Table 1. On the pretest, both experimental and control groups highly overestimated themselves compared to their teacher. The correlation between self-assessment and teacher-assessment for both the experimental ($r = -0.265$, $p = 0.223$) and control groups ($r = -0.224$, $p = 0.343$) on the pretest was negative. This shows that both groups were equally incompetent in assessment of their own performance. However, on the posttest, the correlation between teacher-assessment and self-assessment in the experimental group was statistically significant ($r = 0.664$, $p < 0.001$). This is compared to the correlation obtained from the two variables in the control group ($r = 0.332$, $p = 0.153$) which was not statistically significant. This finding supports the idea that the goal-setting treatment could be effective in increasing learners' accuracy in self-assessment.

Table 1. *Pearson Correlation Between Teacher-Assessment and Self-Assessment of Oral Interview on the Pretest and the Posttest*

| | Mean Exp ¹ TA ³ | Mean Exp SA ⁴ | Pearson | Sig. | Mean Control ² TA | Mean Control SA | Pearson | Sig. |
|----------|---|--------------------------------|---------|------|------------------------------------|-----------------------|---------|------|
| Pretest | 20.21 | 23.91 | -.265 | .223 | 20.00 | 23.25 | -.224 | .343 |
| Posttest | 20.04 | 21.52 | .664* | .001 | 19.30 | 22.90 | .332 | .153 |

1: experimental group, 2: control group, 3: teacher-assessment, 4: self-assessment

Another instrument used to collect the data was the Metacognitive Awareness Inventory (Schraw & Dennison, 1994) which consisted of 52 items, measuring different aspects of metacognitive awareness of the learners. Each aspect of this awareness in the questionnaire was compared between the groups on both the pretest and posttest. As Table 2 shows, none of the items was different in the groups.

However, the difference between the learners on the posttest items showed that metacognitive awareness in the participants of the experimental group rose because scores of five aspects out of seven showed statistically significant differences. The learners' declarative knowledge in both groups increased; however, the participants in the experimental group showed much higher increase in this aspect of metacognitive awareness, $t(101) = 3.54$, $p = .001$. The results also revealed that the learners in the experimental group showed higher levels of procedural knowledge on the posttest, $t(101) = 2.69$, $p = 0.010$. The scores of the planning aspect also showed significant differences in the experimental group compared to the control group $t(101) = 2.20$, $p = .033$. The same goes for comprehension monitoring $t(101) = 6.07$, $p = .000$, and evaluation, $t(101) = 6.26$, $p = .000$. The overall score of the participants in the questionnaire also showed improved the

results for the experimental group, $t(101) = 2.45$, $p = .019$. Table 2 shows Metacognitive Awareness Inventory (Schraw & Dennison, 1994) on the pretest and the posttest:

Table 2. *Metacognitive Awareness Inventory Results on the Pretest and the Posttest*

| Metacognitive Awareness Inventory | Mean Exp (Pretest) | Mean Control (Pretest) | t | Sig. | Mean Exp (Posttest) | Mean Control (Posttest) | t | Sig. |
|-----------------------------------|--------------------|------------------------|-------|------|---------------------|-------------------------|------|-------|
| Declarative Knowledge | 16.34 | 16.80 | -0.19 | .850 | 31.30 | 19.60 | 3.54 | .001* |
| Procedural Knowledge | 8.00 | 8.40 | -0.35 | .728 | 11.30 | 7.80 | 2.69 | .010* |
| Conditional Knowledge | 15.21 | 15.00 | 0.10 | .914 | 11.95 | 11.75 | 0.11 | .910 |
| Planning | 17.58 | 19.60 | -1.61 | .114 | 21.30 | 15.40 | 2.20 | .033* |
| Information Management Strategies | 46.95 | 41.50 | 1.02 | .310 | 47.82 | 42.00 | 1.38 | .172 |
| Comprehension Monitoring | 13.69 | 17.50 | -1.47 | .149 | 29.82 | 13.65 | 6.07 | .000* |
| Debugging Strategies | 13.04 | 15.25 | -1.07 | .290 | 15.21 | 12.75 | 1.29 | .203 |
| Evaluation | 13.03 | 17.70 | -1.91 | .063 | 25.82 | 11.40 | 6.26 | .000* |
| Overall Score | 143.86 | 151.75 | -0.95 | .345 | 194.54 | 134.35 | 2.45 | .019* |

The results from the analysis of immediate recall interview on the pretest showed that the learners did not have a well-founded criterion for scoring their performance. Most of them who scored themselves with high grades claimed that their performance was completely acceptable and when they were asked to instantiate their claim, they could hardly provide any response. From these interviews, it was revealed that the learners lacked a reasonable and efficient monitoring ability.

However, on the posttest, the recall interviews showed that they gained the ability to monitor themselves. They also reported that they gained some insights into suitable criteria for scoring their performance. These reports make the discussion of some points necessary. These points are introduced in some detail in the next section of this paper.

5. Discussion

This study sought to find answers to three questions which were posed earlier in this paper. The reported results of the study about self-assessment accuracy and metacognitive awareness are used here to answer the proposed questions.

In answering the first question about the relationship between the learners' metacognitive awareness and their accuracy in self-assessment of a speaking test, the findings revealed a positive correlation between these two variables. About the second and the third questions, the results showed that the learners who were metacognitively aware were more accurate in self-assessing their performance in speaking tests and the goal-setting activity improved other components of their metacognitive awareness.

Drawing on the attained results from the self-assessment of the learners in both groups on the pretest, we can be sure that the two groups were similar in their self-assessment ability. The low grades in almost all the items of Metacognitive Awareness Inventory (Schraw & Dennison, 1994) and the learners' claims in the recall interview can be used as evidence in favour of the claim that the learners were not capable of handling self-assessment tasks because they were metacognitively unaware.

On the other hand, the learners in the experimental group were shown to have higher metacognitive awareness because their difference with the control group in five aspects was statistically significant. The experimental group also showed a higher and statistically significant correlation with teacher assessment as well. Such a finding corroborates Oscarson's (1989) view on self-assessment. However, it should be mentioned that, contrary to our findings, Oscarson believes that raised awareness and improved goal-orientation are the result of learners' involvement in the process of self-assessment. These claims can be the indication of a reciprocal relationship among these variables. Moreover, we believe that if self-assessment is intended to be used to raise learners' awareness or improve their goal-setting, it should happen over a long period of time. This way, learners will have enough time to reflect on their internal processes and they can improve their goal-setting skills and awareness, which will lead to more accurate self-assessment on their part (Harris, 1997).

The fact that the participants in the experimental group gained higher metacognitive awareness and showed higher correlation with teacher assessment can be used to support the claim in favour of the relationship between metacognitive awareness and accuracy in self-assessment. But we must find what connects our treatment to raised level of awareness and improved self-assessment procedure. This link is created through an analysis of different aspects of metacognitive awareness

inventory and learners immediate recall sessions. Researchers believe that different items of the metacognitive awareness questionnaire measure different subsets of the concept of metacognitive awareness, namely declarative knowledge, procedural knowledge, conditional knowledge, planning, information management strategies, comprehension monitoring, debugging strategies, and evaluation. From the five aspects which were statistically significant two aspects were related to planning and evaluation. Researchers believe that because this increase in these two items is related to each other, it can be postulated that increase in learners' planning and goal-setting skill can directly affect other aspects of learners' metacognitive awareness. This idea supports Alexander et al.'s (1991) idea that planning component of metacognitive awareness can affect other aspects.

Contrary to our findings, Kato (2009) claimed that his participants in elementary and intermediate levels found goal-setting activity ineffective. Because he used goal-setting and self-assessment activities concurrently, his participants were unable to perceive the effect of goal-setting on their self-assessment skills. Moreover, Kato (2009) did not try to measure the degree to which the learners' self-assessment was accurate. In another study measuring the accuracy of self-assessment, Patri's (2002) findings showed that self-assessment, even with a set criteria, is not as precise as peer-assessment and teacher assessment. Although Patri's findings may not confirm our findings, it indicates that awareness of objectives is needed for accurate assessment. Andrade, Du, and Wang (2008), in line with our findings, claimed that criteria should be set for learners in order for self-assessment to be effective and accurate. However, for this desired effect long-term practice and sufficient feedback is essential (AlFalla, 2004).

With this short discussion on the results of this study, we can conclude that the planning and goal-setting can lead to higher metacognitive awareness, and this awareness in turn can lead to more accurate self-assessment in learners. With this interpretation of the results, we move to the next section of the paper to sum up the study.

6. Conclusion

This study drew on the relationship between metacognitive awareness and accuracy of self-assessment in lower-intermediate EFL learners. We believed that planning was one of the components of metacognitive awareness that could play a crucial role in raising awareness in this regard. Therefore, the study aimed to raise this awareness through goal-setting among the learners and presenting objectives of the tasks to them. Based on the findings and interpretation of the results, it can be concluded that planning and awareness of the goals may lead to improved metacognitive awareness and, as a result, this will help learners to be more successful and accurate in the self-assessment process.

The findings of this study can be applied to language learning classes by presenting the objectives of different tasks to learners and by asking them to set goals for own their language learning. Such activities, at least in lower levels, could enhance learners' awareness and improve their language skills. In order to shed more light on the issue, researchers can conduct similar studies on learners in different proficiency levels to observe the effect of these treatments on their awareness and self-assessment accuracy. Similar studies could be conducted studying the effect of planning and goal-setting on learners' metacognitive awareness and self-assessment skills of adolescents and children. Such studies could be especially insightful because learners of these ages represent different stages of cognitive development, and similar treatments used in this study could result in different effects on their development. Finally, researchers can conduct experiments by providing treatments that affect other aspects of metacognitive awareness to see whether planning will also be affected and whether learners' self-assessment skills will be developed.

References

- Alexander, P. A., Schallert, D. L., & Hare, V. C. (1991). Coming to terms: How researchers in learning and literacy talk about knowledge. *Review of Educational Research, 61*, 315-343.
- AlFallay, I. (2004). The role of some selected psychological and personality traits of the rater in the accuracy of self- and peer-assessment. *System, 32*, 407-425.
- Andrade, H., Du, Y., & Wang, X. (2008). Putting rubrics to the test: The effect of a model, criteria generation, and rubric-referenced self-assessment on elementary school students' writing. *Educational Measurement: Issues and Practice, 27*, 3-13.
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Benson, P., & Lor, W. (1999). Conceptions of language and language learning. *System, 27*, 459-472.
- Brown, H. D. (2004). *Language assessment: Principles and classroom practices*. White Plains, New York: Pearson Education.
- Brown, J. D., & Hudson, T. (1998). The alternatives in language assessment. *TESOL Quarterly, 32*, 653-675.
- Cohen, A. D. (1998). Strategies and processes in test taking and SLA. In L.F. Bachman & A. D. Cohen (Eds.), *Interface between second language acquisition and language testing research* (pp. 90-111). Mahwah, NJ: Lawrence Erlbaum Associates.

- Cotterall, S. (1995). Readiness for autonomy: Investigating learner beliefs. *System*, 23, 195-205.
- Coutinho, S. A. (2007). The relationship between goals, metacognition, and academic success. *Educate*, 7, 39-47.
- Dörnyei, Z., & Skehan, P. (2003). Individual differences in second language learning. In C. Doughty & M. Long (Eds.), *Handbook of second language acquisition* (pp. 589-630). Malden, MA: Blackwell.
- Feuerstein, R. (1980). *Instrumental enrichment: An intervention program for cognitive modifiability*. Baltimore, MD: University Park Press.
- Goh, C. (1997). Metacognitive awareness and second language listeners. *ELT Journal*, 51, 361-369.
- Goh, C. (2008). Metacognitive instruction for second language listening development: Theory, practice and research implications. *RELC Journal*, 39, 188-213.
- Goh, C., & Taib, Y. (2006). Metacognitive instruction in listening for young learners. *ELT Journal*, 60, 222-232.
- Hamman, L. A., & Steven, R. J. (1998). *Metacognitive awareness assessment in self-regulated learning and performance measures in an introductory educational psychology course*. Paper presented at the Annual Meeting of the American Educational Research Association.
- Harris, M. (1997). Self-assessment of language learning in formal settings. *ELT Journal*, 51, 12-20.
- Heilenman, L. K. (1990). Self-assessment of second language ability: The role of response effects. *Language Testing*, 7, 174-201.
- Huerta-Macías, A. (2002). Alternative assessment: Responses to commonly asked questions. In J. C. Richards & W. A. Renandya (Eds.), *Methodology in language teaching: Anthology of current practice* (pp. 338-343). Cambridge: Cambridge University Press.
- Kato, F. (2009). Student preferences: Goal-setting and self-assessment activities in a tertiary education environment. *Language Teaching Research*, 13, 177-199.
- Kondo-Brown, K. (2005). Differences in language skills: Heritage language learner subgroups and foreign language learners. *The Modern Language Journal*, 89, 563-581.
- LeBlanc, R., & Painchaud, G. (1985). Self-assessment as an L2 placement instrument. *TESOL Quarterly*, 19, 26-40.

- Malabonga, V., Kenyon, D. M., & Carpenter, H. (2005). Self-assessment, preparation and response time on a computerized oral proficiency test. *Language Testing, 22*, 59-92.
- McNamara, T. (2000). *Language testing*. Oxford: Oxford University Press.
- Mori, Y., Sato, K., & Shimizu, H. (2007). Japanese language students' perceptions on Kanji learning and their relationship to novel Kanji word learning ability. *Language Learning, 57*, 57-85.
- Oscarson, M. (1989). Self-assessment of language proficiency: Rationale and applications. *Language Testing, 6*, 1-13.
- Patri, M. (2002). The influence of peer feedback on self and peer-assessment of oral skills. *Language Testing, 19*, 109-131.
- Pounder, J. S. (2000). A behaviourally anchored rating scales approach to institutional self-assessment in higher education. *Assessment and Evaluation in Higher Education, 25*, 171-182.
- Rivers, W. P. (2001). Autonomy at all costs: An ethnography of metacognitive self-assessment and self-management among experienced language learners. *The Modern Language Journal, 85*, 279-290.
- Ross, S. (1998). Self-assessment in second language testing: a meta-analysis and analysis of experiential factors. *Language Testing, 15*, 1-20.
- Sakui, K., & Gaies, S. J. (1999). Investigating Japanese learners' beliefs about language learning. *System, 27*, 473-492.
- Schneider, E. & Ganschow, L. (2000). Dynamic assessment and instructional strategies for learners who struggle to learn a foreign language. *Dyslexia, 6*, 72-82.
- Schraw, G., & Dennison, R. S. (1994) Assessing metacognitive awareness. *Contemporary Educational Psychology, 19*, 460-475.
- Spence-Brown, R. (2001). The eye of the beholder: Authenticity in an embedded assessment task. *Language Testing, 18*, 463-481.
- Spolsky, B. (1985). The limits of authenticity in language testing. *Language Testing, 2*, 31-40.
- Sullivan, K., & Hall, C. (1997). Introducing students to self-assessment. *Assessment and Evaluation in Higher Education, 22*, 289-306.
- Sullivan, K., & Lindgren, E. (2002). Self-assessment in autonomous computer-aided second language writing. *ELT Journal, 56*, 258-266.

- Ulicsak, M. H. (2004). "How did it know we weren't talking?": An investigation into the impact of self-assessments and feedback in a group activity. *Journal of Computer Assisted Learning*, 20, 205-211.
- Vandergrift, L. (1999). Facilitating second language listening comprehension: Acquiring successful strategies. *ELT Journal*, 53, 168-176.
- Vandergrift, L. (2005). Relationships among motivation orientations, Metacognitive awareness and proficiency in L2 listening. *Applied Linguistics*, 26, 70-89.
- Vandergrift, L., & Tafaghodtari, M. (2010). Teaching L2 learners how to listen does make a difference: An empirical study. *Language Learning*, 60, 470-497.
- Victori, M., & Lockhart, W. (1995). Enhancing metacognition in self-directed language learning. *System*, 23, 223-234.
- Wenden, A. (1998). Metacognitive knowledge and language learning. *Applied Linguistics*, 19, 515-37.
- Yang, N. (1999). The relationship between EFL learners' beliefs and learning strategy use. *System*, 27, 515-535.

Appendix
Metacognitive Awareness Inventory
 (Schraw & Dennison, 1994)

| | True | False |
|---|------|-------|
| 1. I ask myself periodically if I am meeting my goals. | | |
| 2. I consider several alternatives to a problem before I answer. | | |
| 3. I try to use strategies that have worked in the past. | | |
| 4. I pace myself while learning in order to have enough time. | | |
| 5. I understand my intellectual strengths and weaknesses. | | |
| 6. I think about what I really need to learn before I begin a task | | |
| 7. I know how well I did once I finish a test. | | |
| 8. I set specific goals before I begin a task. | | |
| 9. I slow down when I encounter important information. | | |
| 10. I know what kind of information is most important to learn. | | |
| 11. I ask myself if I have considered all options when solving a problem. | | |
| 12. I am good at organizing information. | | |
| 13. I consciously focus my attention on important information. | | |
| 14. I have a specific purpose for each strategy I use. | | |
| 15. I learn best when I know something about the topic. | | |
| 16. I know what the teacher expects me to learn. | | |
| 17. I am good at remembering information. | | |
| 18. I use different learning strategies depending on the situation. | | |
| 19. I ask myself if there was an easier way to do things after I finish a task. | | |
| 20. I have control over how well I learn. | | |
| 21. I periodically review to help me understand important relationships. | | |

| | | |
|---|--|--|
| 22. I ask myself questions about the material before I begin. | | |
| 23. I think of several ways to solve a problem and choose the best one. | | |
| 24. I summarize what I've learned after I finish. | | |
| 25. I ask others for help when I don't understand something. | | |
| 26. I can motivate myself to learn when I need to | | |
| 27. I am aware of what strategies I use when I study. | | |
| 28. I find myself analyzing the usefulness of strategies while I study. | | |
| 29. I use my intellectual strengths to compensate for my weaknesses. | | |
| 30. I focus on the meaning and significance of new information. | | |
| 31. I create my own examples to make information more meaningful. | | |
| 32. I am a good judge of how well I understand something. | | |
| 33. I find myself using helpful learning strategies automatically. | | |
| 34. I find myself pausing regularly to check my comprehension. | | |
| 35. I know when each strategy I use will be most effective. | | |
| 36. I ask myself how well I accomplish my goals once I'm finished. | | |
| 37. I draw pictures or diagrams to help me understand while learning. | | |
| 38. I ask myself if I have considered all options after I solve a problem. | | |
| 39. I try to translate new information into my own words. | | |
| 40. I change strategies when I fail to understand. | | |
| 41. I use the organizational structure of the text to help me learn. | | |
| 42. I read instructions carefully before I begin a task. | | |
| 43. I ask myself if what I'm reading is related to what I already know. | | |
| 44. I reevaluate my assumptions when I get confused. | | |
| 45. I organize my time to best accomplish my goals. | | |
| 46. I learn more when I am interested in the topic. | | |
| 47. I try to break studying down into smaller steps. | | |
| 48. I focus on overall meaning rather than specifics. | | |
| 49. I ask myself questions about how well I am doing while I am learning something new. | | |
| 50. I ask myself if I learned as much as I could have once I finish a task. | | |
| 51. I stop and go back over new information that is not clear. | | |
| 52. I stop and reread when I get confused. | | |