EFL Teachers' Epistemological Beliefs and Their Assessment Orientations

Ammar Abdullah Mahmoud Ismail

Faculty of Education., Sohag University, Egypt

E-mail: abdallai2000@yahoo.com

Abstract

Epistemological beliefs—beliefs about the nature of knowledge, where it resides, and how knowledge is constructed and evaluated—have been the target of increased research interest lately. Heretofore, emphasis has been directed to language teaching/learning aspects and strategies. Language assessment practices have not yet received due attention in epistemic research literature. The current study examined the relationship between pre-service EFL teachers' epistemological beliefs and their assessment orientations. Dimensions of epistemological beliefs were assessed via a questionnaire designed and validated by the researcher based on Schommer's work. Two assessment orientations were examined including: (a) transmissive surface- processing orientation and (b) constructive deep-processing orientation. The study involved 114 preservice EFL teachers enrolled in the Professional Diploma in Teaching Program in the Abu Dhabi University, the United Arab Emirates. Results of the study showed that EFL teachers' epistemological beliefs have a direct bearing on their assessment orientations and practices. EFL teachers with naive epistemological beliefs tended more to adopt surface-level assessment orientations whereas those with sophisticated epistemological beliefs showed more tendency to adopt deeper level approaches to assessment in language settings. Results are discussed in terms of backwash effects on foreign language instruction, curriculum development, and teacher education. Suggestions for further research are also discussed.

Keywords: Epistemological beliefs, assessment orientations, surface processing, deep processing

1. Introduction

Foreign language learning settings at different schooling levels have increasingly been witnessing a paradigm shift towards constructivist, self-regulated leaning, and connectivist approaches. Learner autonomy, self-construction of knowledge and understandings, problem solving on own initiative, and self-regulation of learning behaviors and contexts have all been the target of education in different settings. Students are encouraged to actively engage in the language learning striving, test their own hypotheses, draw their own conclusions, and construct their own knowledge and understandings.

Contemporary language education theory and practice thereof, as Hillocks (2002) states, see that "effective learning must be constructive in nature and that students learn best and perhaps only when they construct knowledge for themselves within the frameworks of their existing knowledge"(P. 22). As communicative competence is the ultimate target in foreign language learning settings, as Liu and Zhang (2014), and Schell and Janicki (2013) argue, language learning should by no means be a process in which teachers transmit knowledge to students but one in which students construct their communicative competence on their own initiative. This, as Wang and Zhang (2012) state, requires that teachers' roles be changed from the traditional knowledge transmitters into guides and counselors helping students construct and assimilate new information.

This paradigm shift in foreign language education resulted in a shift in assessment practices from tests designed to assess learners' mastery of discrete areas of the linguistic system to tests designed to tap communicative proficiency and deeper manipulation of language input in authentic language learning contexts (Brindley, 2001; Richards, 2001; Yoshida, 2006). Assessment procedures nowadays, as Gibbs and McDonald (2003) state, are seen not only as significant tools for crediting students, but also as invaluable instruments for monitoring and supporting students' progress. Most importantly, as Segers (2003) states, assessment of students' performance has been changed towards more accent on cognitive competencies such as problem solving and assessment in authentic contexts. On a larger scale, exams are nowadays viewed as a positive driving force that could infuse positive changes in the entire educational system (Andrews, 2004; Braun & Kanjee, 2006).

Yet, evidence from language assessment research literature indicates that a disparity still exists between advocated constructivist approaches to language instruction and actual assessment practices. Structural approaches accenting discrete point testing and rote learning techniques are still haunting assessment practices in different settings (Masters, 2013). This, in view of consequential validity research literature, is likely to have its detrimental consequences on all

aspects of the language curriculum and instructional practices thereof (Greenwald, Nosek, & Sriram, 2006, Zhao, 2013). Assessment practices, as Knight (2002) describes them, are in a disarray and risky to treat them as reliable.

This is of particular significance in exam-driven learning environments, the current one is an example, where what is tested determines to a great extent what is taught. Conclusions of the researcher of the current study come in line with the ones mentioned above. Exams, worksheets, tests, and other formative as well as summative assessment tools developed by EFL teachers in the current context accented shallow processing of language input. More often than not, exam items were repetitive in nature and primarily accented testing different aspects of the language system in an isolated fashion. Emphasis on the formal features of the language system was paramount at the expense of functional aspects such as communication in authentic contexts, self-expression of ideas, and meaning-making endeavors. This impacted not only their instructional procedures but as well the learning practices of their students. Students focused on types of exam items stressed by the teacher and targeted language tasks evaluated in teacher-developed assessment tools. The kind of knowledge and language input deemed necessary for them was that covered by the teacher since this will be the kind of knowledge valued and credited by those teachers.

Identifying reasons behind this recoil in assessment practices towards structural approaches and discrete point techniques might help improve language assessment and avoid the detrimental effects this might have on various aspects of the language curriculum and instructional practices thereof. The current study focuses on personal epistemologies of language teachers - their beliefs regarding the nature of linguistic knowledge, where it resides, and how it should be taught and tested - as a potential factor behind teachers' resistance to the new assessment paradigm and their tendencies to adopt surface-level orientations in language assessment at the expense of deeper-level processing orientations, an area that has not heretofore been explored in assessment research literature.

1.1 Statement of the Problem

It is obvious from the brief introduction mentioned above that foreign language assessment practices prevailing in the Emirati context of pre-tertiary foreign language education targets surface-level processing mechanisms at the expense of deeper manipulation of language input. Discrete point testing and structural practices are still dominating the assessment culture despite repeated calls for adopting communicative, constructivist, and connectivist models of language teaching and testing. More often than not foreign language teachers target lover levels of information processing and cognitive manipulation of language content at the expense of higher level processes. It seems that the problem is rooted in issues deeper than the given interpretations thus far. Little is known about the reasons behind this steady tendency of foreign language teachers to stick to such dysfunctional assessment practices.

The current study fills in a gap in existing epistemic research and language assessment literature via investigating the potential link between the assessment orientations that language teachers are likely to adopt and the personal epistemologies they tend to hold in foreign language learning settings. Three research questions have been tackled, including:

a) What epistemological beliefs do pre-service EFL teachers hold?

b) What assessment orientations do they tend to adopt?

c) Is there a relationship between the epistemological beliefs they hold and their assessment orientations?

1.2 Significance of the Study

Significance of the study stems from a number of considerations, as follows.

a) First, the study focuses on an area that has not received due attention in language assessment literature. This relates to teachers' epistemological beliefs as a potential factor behind their resistance to adopting the new assessment paradigm and their tendencies to embrace rote learning and surface processing techniques. Research on personal epistemologies, as will be discussed in the literature review below, indicates that these have influential impact on various aspects of foreign language instruction. No published research heretofore -- as far as the researcher knows -- has explored the likelihood that their impact extends to the assessment practices of foreign language teachers in this context.

b) Second, the study deals with foreign language teachers in a context that is heavily driven by exams, where what is tested determines to a great extent what is taught. Language skills coverage, curricular focus, and methodological decisions are all impacted by the kind of assessment practices prevailing herein. Identifying teachers' personal theories behind these assessment practices is a key to any successful assessment reform.

c) Third, the study focuses on both formative and summative assessment practices of foreign language teachers and thus taps the assessment process in its entirety rather than focusing on discrete skills and isolated tasks.

d) Finally, the study covers assessment practices of language teachers at different levels of schooling, including primary, intermediate and high school levels, and thus gives a more comprehensive picture about assessment across different educational levels.

2. Review of Literature

A paradigm shift has been witnessed in language assessment over the past two decades leading to rigorous changes in the assessment culture and techniques. This paradigm shift in assessment followed a corresponding shift in instructional practices, where rote learning approaches to language instruction have been increasingly abandoned for more constructivist ones. The pendulum, as Edwards (2004) and Hillocks (2002) argue, has shifted away from educating learners by transmitting a body of knowledge by teachers based on pre-selected textbooks to establishing classroom

environments rich in opportunities for self-construction and self-discovery of knowledge. The discrete point tradition of teaching and testing, as McNamara (2001) states, was seen as "focusing too exclusively on knowledge of the formal linguistic system for its own sake rather than the way this knowledge is used to achieve communication" (P.14).

Consequently, assessment of students' performance shifted towards more accent on their ability to construct knowledge and monitor their own learning endeavors rather than their ability to reproduce the information they were previously fed by teachers and more knowledgeable others. As Segers (2003) and Segers et al. (2003) state, assessment has been changed towards more accent on cognitive competencies such as problem solving and assessment in authentic contexts. It has shifted away from tests designed to assess learners' mastery of discrete areas of the linguistic system to tests designed to tap communicative proficiency and deeper manipulation of language input in different language use situations (Brindley, 2001; Richards, 2001; Yoshida, 2006).

Evidence from empirical research indicates that this shift in assessment practices has resulted in better gains in language learners' performance and achievement. For example, Gibbs and McDonald (2003) stated that changes in assessment design affected student learning by producing more productive learning activity. Similarly, in a series of studies by Dori (2003), it was concluded that instruction and curriculum decisions at the school and classroom levels as well as students' learning outcomes were improved by changing the assessment culture. Gielen et al. (2003) concluded that constructive forms of assessment were key tools in generating appropriate student learning activity, capturing student time and attention, providing timely help and feedback, helping students internalize discipline's standards, and fostering cooperative learning.

Yet, although communicative approaches and constructivist models of language instruction have been in existence for quite a while and have impacted various aspects of language teaching and learning, they have not heretofore had their full potential effect on language assessment practices. A number of researchers and educators have expressed concerns that more efforts are still needed to fully put this new assessment paradigm into practice. As stated by Lokon (2005) and Edwards (2006), there is a growing perception among educators and researchers that a separation is occurring between the goals of actual learning and testing in contemporary education. A similar stance was echoed in Cunningham's (2002) words that "school practices are still dominated by structural methodologies accenting discrete point testing and rote learning techniques that take language structures rather than functions as their primary focus" (P. 22).

These concerns about current assessment practices have been fueled by studies on test washback, referring to the impact of exams on language teaching and learning (Hayes & Read, 2003; Saville & Hawkey, 2003; Hawkey, 2004; Wall & Horak, 2006; Edwards, 2006; Apichatrojanakul, 2011; Estaji & Tageddin, 2012; Aftab et al, 2013; Akpinar & Cakildereb, 2013; Green, 2013). Conclusions of washback research show that current assessment practices have negative effects on curriculum innovation, content, and methodology (Cheng & Watanable, 2004; Wall, 2005; Wall & Horak, 2006). These negative washback effects occur when teachers assess communicative performance objectives with multiple choice structure tests, which as Flinders (2005) states, are likely to "limit the learning opportunities and narrow the educational experience of students" (P.8). In thus doing, as Seherer (2005) argues, educators might fail in making deep learning the primary function of language instruction.

Negative test washback effects extend to various aspects of language instruction including scope, sequence, depth of teaching as well as attitudes towards the content, methods of teaching, and learning. For example, Aftab et al. (2014) and Abedi and Linquanti (2012) reported a strong negative washback from the examinations on teaching methodology, context, and learning. Similarly, O'Sullivan (2006) concluded that "the skills coverage is inadequate because language knowledge rather than language use is being tested" (P.19). Popham and Snyder (2000) see that "exams have been perceived to be harmful and teachers' tendencies to teach to the test are often cited as an impediment to introducing new instructional practices" (P. 460). Hayes and Wyward (2002) see that test content validity has been compromised. Such assessment practices, as Popham (2001) and (1999) argue, often lead to a narrowing of curricular focus and promote unsound skill-and-drill instructional activities, which would in turn undermine educational quality.

Results of washback studies are consistent regarding both high-stakes language exams and teacher-designed classroom-based ones. For example, in an empirical study investigating the washback effects of a high-stakes English test in China (the National Matriculation English Test), Luxia (2007) concluded that "the communicative features of the NMET writing test were not taught the way intended by the test constructors. Both teachers and learners neglected the communicative context of writing while emphasizing the test situation and the assumed preferences of the markers" (P.5). This concurs with Marshall's view (2005) that standardized graduation exams are likely to cause students to become rote learners of facts that lack problem solving skills. Similarly, Amrein and Berliner (2002) stated that standardized testing has a narrowing effect on the curriculum as too much focus on testing skills may reduce opportunities for deeper learning. Similar results were reported by Qi (2004), Wall (2005), Wall and Horak (2006) and Yıldırım, (2010).

Conclusions regarding the negative washback effects of teacher-developed exams are not different from those of high-stakes exams. For example, Lokon (2005) conducted a survey on 129 senior high school English teachers in Japan, in which 59% of the teachers believed that their oral communication classes were ineffective, and 16% of the teachers stated that they had changed their oral communication classes to teach test-taking strategies. Similarly, in a survey of language assessment practices, Hillocks (2002) concluded that current practices are a sort of "educational McDonaldization" as most writing tests" amount to little more than superficial prewriting exercises and what is purportedly in-depth analysis often amounts to little more than a blithe regurgitation of previously stated facts" (P. 14). As Hillocks argues, the narrow focus of such assessment practices fosters writing instruction on surface mechanics

rather than deep content. Similarly, Schuster (2004) described most writing tests as tests of drafting rather than tests of writing, resulting in superficializing writing instruction.

These negative backwash effects of assessment on language education led researchers and language educator to call for a reform in the assessment culture, albeit their disagreement on the core problem. For example, Popham (2001) sees that the problem lies in the "educators' understanding of educational measurement that does not extend much beyond the care and feeding of teacher-made classroom tests" (P. 27). Hillocks (2002) blames the "No Child Left Behind" policies for promoting the superficial performance of explicit test features at the expense of in-depth learning. Datnow, Hubbart, and Mehan (2002) argue that for lasting educational reform to work, fundamental shifts in the attitudes towards testing are needed by all stakeholders in the system. Most teachers, as Sacks (1999) states, are often adverse to adopting deep approaches since their teaching philosophies are so deeply rooted in practice-drill modes of instruction, leading to a "profound disconnection between knowledge and practice" when it comes to testing (P.XI).

Although these interpretations of the problem sound logical regarding assessment practices in general, the problem in classroom-based teacher-designed tests seems a bit different. In this context, classroom teachers have more freedom to construct their own tests based on a plethora of options available to them. A number of questions arise here. First, what makes language teachers in the same language context opt for surface discrete point testing approaches rather than deeper level ones (or vice versa)? Second, what makes language teachers adverse to adopting deeper approaches to teaching and assessment in language settings? Third, what causes the profound disconnection between knowledge and practice in assessment settings?

One of the under-researched areas that might provide answers to these questions is personal epistemologies of language teachers -- their beliefs about the nature of the linguistic knowledge, how it is constructed, and how it should be evaluated. This line of research has topped the research agenda in different domains of knowledge. Yet, it has not heretofore received due attention in foreign language assessment settings.

2.1 Personal Epistemologies

Personal epistemology, according to Hofer (2004), is "a field that examines what individuals believe about how knowing occurs, what counts as knowledge and where it resides, and how knowledge is constructed and evaluated" (P. 1). Work on personal epistemologies was initiated by Perry in 1970s. Perry (cited in Schommer-Aikins, 2004:19) described personal epistemological beliefs as developing along a continuum of four stages, including: dualism, multiplicity, relativism, and commitment. Individual epistemological beliefs change from one stage to another.

Dualism refers to beliefs in knowledge as comprised of absolute truths (right/wrong) that can be transmitted from authorities. In the second stage, Multiplicity, individuals start viewing knowledge as multiple rather than dualistic, comprising both absolute truths and personal opinions (their inner voice) though they still believe in knowledge as certain. The third stage, according to Perry, marks a shift in epistemic development when individuals start believing in knowledge as constructed and reasoned. Truth is considered tentative and relative to individuals' personal interpretations and experiences as validated and supported with evidence. The final stage, Commitment, marks the highest level of epistemic development when individuals become relativist thinkers showing more commitment to particular knowledge beliefs in a flexible manner.

In (1992), Baxter-Magolda proposed an epistemic reflection (ER) model similar to that of Perry. According to this model, epistemological beliefs develop along four stages in which individuals' knowledge beliefs change from beliefs in absolute knowledge in initial stages of epistemic development to beliefs in relativist knowledge, constructed knowledge, and contextual knowledge, respectively. Absolute knowers view knowledge as certain and believe that authorities have all the answers. Transitional knowers discover that authorities are not all-knowing and begin to accept the uncertainty of knowledge. Those who are independent knowers question authority as the only source of knowledge and begin to hold their own opinions as equally valid. Contextual knowers are capable of constructing an individual perspective by judging evidence in context. Each of these stages, according to Baxter-Magolda (2004), leads to "particular expectations of the learner, peers, and instructors in learning settings, as well as to an understanding of how learning should be evaluated and how educational decisions are made" (P. 29).

Schommer-Aikins'work (1994) marked a shift in understanding epistemological beliefs. While Perry and Baxter-Magolda viewed epistemic development as a sequential process of epistemic sophistication, Schommer-Aikins conceptualized epistemological beliefs as a system of five more or less independent dimensions in a frequency distribution ranging from epistemic naiveté to epistemic sophistication. The proposed dimensions include stability of knowledge, structure of knowledge, source of knowledge, control of knowledge, and speed of knowledge acquisition. The "stability of knowledge" dimension ranges from beliefs in knowledge as fixed and unchanging to beliefs that knowledge is tentative and ever-changing. The "structure of knowledge" ranges from beliefs in knowledge as bits and pieces to beliefs in knowledge as integrated. The "source of knowledge" dimension includes beliefs in knowledge as handed down by omniscient authorities to knowledge as derived by reason and supported with evidence. The "speed of knowledge acquisition" ranges from beliefs in knowledge acquisition as quick or not at all to beliefs in acquisition as gradual and time consuming. The "control of knowledge" dimension refers to beliefs in individuals' ability to learn as fixed at birth to ability to learn being improvable through sustained effort. According to Schommer-Aikins (1994, 2004), belief dimensions are more or less independent, i.e. individuals might have sophisticated beliefs on a given dimension but naïve beliefs on other dimensions.

Empirical research on personal epistemologies revealed that students' epistemic positions overshadow various aspects of their learning. For example, in a study by Ali and Ammar (2005), it was concluded that EFL tertiary students' epistemological beliefs have significant impact on their language learning strategies and foreign language classroom anxiety. In a subsequent study on a similar sample of tertiary EFL students, Ammar and Abdel-Majeed (2006) found a positive correlation between students' epistemological beliefs and their goal orientations, cognitive engagement, perceived linguistic competence, and achievement. Similar studies correlated learners' epistemological beliefs with their cognitive processing (Kardash & Howell, 2000; Qian & Alvermann, 2000; Hofer, 2001), self-regulated learning strategies (Donald, 2002; Schreiber, Shinn & Weems, 2003), interpretation of instructional practices (Hammer &Elby, 2002; Hammer, 2002; Many et al., 2002), and metacognitive awareness (Hofer, 2004). Research also indicated that students' epistemological beliefs strongly impact their motivation to learn and utilize strategies (Chan, 2002).

Similarly, empirical evidence indicates that epistemological beliefs of language teachers overshadow various aspects of their teaching. For example, teachers' epistemic stances impact their teaching orientations (Hofer, 2000; Beuhl & Alexander, 2001) as well as their conceptions about teaching/learning (Chan, 2003). Likewise, research indicated that teachers' epistemological beliefs impact how they react to different classroom problems (Donmoyer, 2001) as well as how they take decisions in language settings (Tsai, 2000; Brownlee et al. 2001).

Yet, despite the remarkably growing evidence regarding the impact of personal epistemologies on various aspects of language instruction in different settings, language assessment practices have not heretofore received due attention in the epistemic research literature. This becomes really critical in exam-driven educational settings, where classroom practices are often tailored to exam purposes and where, as Flinders (2005) argues, "what is tested determines what is taught" (P. 8). The current study attempts at shedding light on this area through investigating the nature of epistemological beliefs that EFL teachers are likely to hold, the kind of assessment orientations they tend to embrace, and any potential relationship between their epistemological beliefs and assessment orientations.

3. Method

3.1 Participants

The study involved 114 EFL pre-service teachers doing their full-time practicum at Al-ain Educational District in Abu Dhabi (ADU), the United Arab Emirates, in the academic years 2012, 2013, 2014. Their practicum experience was part of the Professional-diploma-in-teaching Program (a post-graduate program). This program admits university graduates with first degrees in language studies (TEFL, TESL, communication, literature, applied linguistics, ... etc.) who opt for teaching as their future career after finishing their BA degrees. As a job requirement, they had to have a professional degree in language teaching before joining schools. The study sample involved female students only as the English majors admitted those years were all female students (save two). Those two did not submit the required materials and were thus excluded from the final study sample.

3.2 Instrumentation & Data Collection Procedures

A set of procedures were following to collect data regarding preservice EFL teachers' epistemological beliefs and their assessment orientations in language settings. Their epistemological beliefs were tapped using a questionnaire while their assessment orientations were tapped through a bi-methodic approach as detailed below.

3.2.1 Epistemological Beliefs

The epistemological beliefs of the study sample were tapped using an Epistemological Beliefs Questionnaire. The questionnaire is a further-developed version of the one used in an earlier study (Ali & Ammar, 2005) based on Schommer's Epistemological Beliefs Questionnaire (EBQ). The earlier version of the questionnaire contained 65 items dealing with EFL students' beliefs about the nature of the linguistic knowledge, how it is constructed and evaluated and how knowledge acquisition occurs.

The earlier version was submitted to a jury of TEFL specialists to decide on its validity for assessing the epistemological beliefs of the study sample. Necessary changes by deletion, addition, or modification were made based on the jurors' recommendations. Then, the questionnaire was administered to 234 tertiary students in the Emirati EFL context. A confirmatory factor analysis using principal component factoring with Varimax rotation and Eigen values of 1.0 or greater was then conducted to decide on its factorial structure and item loadings. Items that loaded on more than one factor or those with loadings less than (0.5) were eliminated from the final analysis. An examination of the resulting Scree plot suggested five epistemic dimensions rather the six dimensions in the earlier version of the questionnaire. Items of the two dimensions of the "structure of knowledge" and the "integration of knowledge" in the earlier version loaded together on one factor.

The final form of the questionnaire included (40) items distributed along five dimensions of epistemological beliefs including, source of knowledge, control of knowledge, certainty of knowledge, structure of knowledge, and speed of knowledge acquisition. The first dimension, source of knowledge, ranges from students' believing in "knowledge as handed down by experts and authority" to "knowledge as obtained by reason and evidence". The second dimension, Control of Knowledge, ranges from "students' beliefs in linguistic ability as inborn and fixed at birth" to "beliefs in linguistic ability as resulting from expenditure of effort over time. The third dimension, Certainty of Knowledge, ranges from beliefs in "knowledge as certain and absolute" to "knowledge as tentative and ever-changing". The fourth dimension, the structure of knowledge, ranges from "beliefs in linguistic knowledge as simple and fragmentary" to "beliefs in linguistic knowledge as complex, integrated, and in constant interplay with the context of language use

situations". The fifth dimension, the speed of knowledge acquisition, ranges from "beliefs in language acquisition as quick or not at all" to "beliefs in acquisition as gradual and time-consuming".

The questionnaire used Likert type ratings. Items were assigned scores on a 5-point rating ranging from Totally Agree '5' to "Totally Disagree"1". Items with negative statement were reverse-coded so that high scores on the questionnaire indicated sophisticated epistemological beliefs; the opposite holds true.

3.2.2 Assessment Orientations

The researcher tried to tap both the process and the product components of assessment by focusing on EFL teachers' ongoing (formative) as well as summative assessment practices. While the latter is more product- oriented, the former provides useful information about the process of learning (Carbery, 1999). In so doing, a bi-methodic approach was utilized including self-report protocols and teacher-designed exam papers/assessment sheets. In the former, teachers were asked to write a weekly report on their assessment practices summarizing their answers on three basic themes including: (1) the new language content areas studied each week, the learning outcomes they tried to achieve, and the assessment techniques/tasks/practices they utilized. The overall objective was to tap preservice teachers' assessment in context. In the latter, they were asked to submit copies of the assessment exercises, tasks, observations, or any other tools targeting assessment they have designed and utilized throughout their practicum experience.

Three raters (inter-rater reliability = .91) were involved in analyzing data about the assessment orientations of teachers based on the weekly self-report protocols and the assessment tools they designed. A preservice EFL teacher's final score on assessment practices was obtained through a 3-step process. First, weekly reports were analyzed and scored on a 5-point rating scale ranging from "1" referring to "absolutely surface" assessment orientation to "5" referring to "absolutely constructive" orientation. A mean score for each teacher was obtained by averaging the scores given by the three raters on all reports. Second, exam papers/tasks/sheets were similarly analyzed and a mean rating given to each teacher based on whether assessment practices targeted surface-level processing or deeper level features. (3) Finally, individual teachers' scores from the previous two steps were summed up and averaged to obtain mean scores on their assessment orientations.

4. Findings

Pre-service EFL teachers' scores on the Epistemological Beliefs Questionnaire as well as on their assessment practices were calculated and tabulated to provide answers to the research questions as detailed below.

4.1 EFL Teachers' Epistemological Beliefs
To answer the first question regarding the epistemological beliefs that preservice EFL teachers tend to adopt, their responses to the Epistemological Beliefs Questionnaire were scored and tabulated. Descriptive procedures (Mean scores and standard deviations) were used as basis for analysis as outlined in Table (1).

Table 1. Mean Scores and Standard deviations Obtained by pre-service EFL Teachers' on their Epistemological Beliefs

Epistemic Dimension	Maximum Score	Mean Scores Obtained	Standard Deviations
Source of knowledge	5	2.65	0.78
Certainty of knowledge	5	2.82	0.57
Structure of knowledge	5	2.48	0.83
Control of knowledge	5	2.74	0.69
Speed of knowledge acquisition	5	3.21	0.84

As shown in the table above, EFL teachers seem to hold mixed beliefs about the nature of linguistic knowledge and how this knowledge is constructed and should be evaluated, though their epistemological beliefs tend to be generally naive. Their mean score on the "Source of Knowledge" dimension is (2.65) on a 5-point scale, indicating that they tend to strongly believe in knowledge transmitted from teachers to be true and valid. and should be taken for granted without questioning its credibility. Similarly, they have a strong belief that the textbooks assigned to students should be considered as the only source of knowledge. How much an EFL student gets out of school, according to their responses, depends heavily on the quality of the instructors.

Similarly, they showed naïve epistemological beliefs on the "Certainty of knowledge" dimension of the questionnaire. Their mean score is (2.82 on a 5-point scale), indicating that they tend to believe in linguistic knowledge as fixed and non-changing. Even when handling dynamic features of the linguistic system such as communication, they tend to view communicative situations as repetitive and thus ask their learners to memorized expressions in the "studied situations" so that they be able to rehearse them when encountering similar situations outside. They seem unable to distinguish between established fixed features of the linguistic system such as grammar and structures and the dynamic features characteristic of language communication. Likewise, as indicated by their responses to items of the questionnaire, they

believe that there is usually one best way for handling language tasks and that students should be told exactly what they are supposed to learn and what they have to do. Negotiation of ideas or attempts of self-expressions, according to their responses, should be controlled or guided by teachers.

Likewise, they tend to hold naïve epistemological beliefs regarding the "structure of knowledge" dimension. Their mean score is (2.48), referring to their strong conviction that linguistic knowledge is fragmentary and discrete. As language skills and areas are studied separately in the language classrooms, this view is carried over to language learners that linguistic knowledge is fragmentary rather than one whole phenomenon. Separation of language skills and areas for the sake of study is misinterpreted in a way that depicts these skills and other features of the linguistic system as fragmentary rather than integrated. As their responses to items of the questionnaire indicate, they tend to focus on the here and now of language study rather than look at the whole picture of the language phenomenon as a process of communication that requires integration of language skills, aspects, and areas.

Their mean scores on the "Control of knowledge" is (2.74), indicating that they have naïve beliefs that ability to learn a foreign language is innate and fixed at birth. They tend to believe that in case one is not born with innate linguistic ability, s/he will not succeed in language learning regardless of how much they expend in time and effort. As their responses to items of the questionnaire indicate, they believe that low-ability language learners have few chances for achievement and that modest EFL performance is basically due to poor innate ability. They see individual differences among EFL students as basically due to different levels of linguistic aptitude. As well, they see high language achievers as having distinguished innate abilities to handle language tasks and perform well in language learning settings.

The only dimension they showed relatively sophisticated beliefs is the "Speed of Knowledge Acquisition". Their mean score is (3.21), indicating that those teachers have fairly strong convictions that language acquisition/learning is a continuous process that takes time. They responded negatively to items such as "usually the first time you try a new language course, you can tell how well you are going to do at it", "it is useless to spend lengthened periods of time studying difficult language content" and "It's a waste of time to work on language tasks that have no precise answers". They provided positive answers to items such as "Working on language tasks for extended periods of time pays off for really smart students", "Learning English is a process of building up knowledge and competence gradually" and "Becoming a competent EFL learner takes a lot of time and needs concentrated effort".

4.2 EFL Teachers' Assessment Orientations

To answer the second question regarding EFL teachers' assessment orientations, mean scores obtained by individual teachers on their (i) weekly reports and (ii) their assessment tools were averaged and tabulated. Descriptive statistics (mean scores and standard deviations) were used as basis for analysis as shown in Table (2).

Table 2. Mean Scores and Standard Deviations Obtained by EFL teachers on their assessment practices

Number of participants	Maximum score	Mean score obtained	Standard Deviation
114	5	2.59	0.67

Data shown in Table (2) indicate that preservice EFL teachers tended more to adopt surface discrete-point assessment orientations rather than deeper-level constructive ones. Their mean score was (2.59) on a 5-point rating. When testing their students, whether formatively or summatively, they focused on testing their students' ability to recall information previously presented to them in class. Quizzes, exams, and other assessment tools tended to be emulative of those in the workbook. Decontextualized language use was paramount in most assessment tools.

Analysis of their self-reported assessment practices revealed that they were not able to distinguish between the fairly static features of linguistic knowledge and the changing dynamic ones. They tested dynamic features of language as if they were static and non-changing. For example, they handled communicative situations and contexts as if they were repetitive and consequently focused on testing their students' abilities to memorize expressions and structures the way they were studied in the classroom. Creative expressions in different communicative contexts were not encouraged and students' abilities to initiate interaction and to take risks in language classes were rarely assessed. Likewise, when handling reading comprehension passages, they used to focus on explicit questions with little attention to implicit of inferential questions, i.e. those targeting deeper content and deeper cognitive engagement. Listening tasks were handled the same way.

4.3 Epistemological Beliefs and Assessment Orientations

To answer the third research question regarding the relationship between preservice EFL teachers' epistemological beliefs and their assessment orientations, the assessment practices of teachers with sophisticated epistemological beliefs (the highest 27% on the EBQ, N=32) were compared to the assessment practices of those with naïve epistemological beliefs (the lowest 27% on the EBQ, N=32) using ANOVA procedures as shown in Table (3).

Table 3. Univariate analysis of the impact of EFL teachers' epistemological beliefs on their assessment orientations

	Sum of Squares	df	Mean Square	F- value	η2
Between Groups	39.942	1	39.942	71.07**	.534
Within groups	34.844	62	0.562		
Total	74.786	63			

As indicated by data in the above table, there is a statistically significant difference between mean scores of EFL teachers with sophisticated personal epistemologies and those with naïve epistemologies favoring the former group {F (1, 63) = 71.07, P .05}. EFL teachers in the sophisticated epistemologies group (N=32, M= 3.47, SD = 0.86) showed more tendency to adopt deeper level assessment orientations than their peers in the naïve epistemologies group (N=32, M = 1.89, SD= 0.62). Those in the naive group tended more to use surface level assessment practices when evaluating their students' language performance.

An examination of the effect size indicates that the epistemological beliefs of EFL teachers were able to explain about 53% of the variance in their assessment orientations (η2 = .534). It is normal that EFL teachers who believe in the certainty of knowledge to focus on fixed features of language (grammar and structures) rather than the dynamic features characteristic of authentic communication. Similarly, those who believe in the integration of knowledge – language is one whole rather than a sum of bit and pieces -- tend more to assess language in its entirety rather than discrete points and skills.

Again, pre-service teachers with naïve epistemological beliefs could not match their assessment with the level of cognitive processing required by different language tasks/situations. Their assessment often targeted surface level features. The mismatch between deeper level processing needed for functioning effectively in communicative situations and the surface level practice-and-drill assessment would have its negative wash-back effects on teaching and learning alike.

5. Discussion

Results of the current study provide evidence that the paradigm shift in language instruction from structural role learning to communicative and constructivist approaches has not yet had its full impact on language assessment in EFL contexts. The assessment practices of many EFL teachers are still oriented towards surface level processing of language input and the discrete point testing of students' language performance. Assessment of students' deeper processing of language input and assessment of language performance in authentic language use situations have not yet received due attention.

These assessment practices as depicted in the weekly reports of the preservice teachers as well as the assessment tools they developed to evaluate their students' performance rarely go beyond the feed and repeat philosophy of language teaching and testing. The only information that's tested is that embedded in assigned textbooks or class reading materials. Assessment situations and question types are all repetitive of evaluations previously practiced in class settings. Students are expected just to recall information previously fed to them the same way and format. This empirical evidence concurs with what Sacks (1999) and Popham (2001) stated regarding language assessment practices that are still focusing on discrete point testing at the expense of deeper processing.

These results and the empirical evidence hereof provide some logical answers to the many issues raised in the language assessment literature. On the onset of this study, a basic question was raised regarding possible reasons behind the disconnection between knowledge and practice in language assessment (Sacks, 1999), and the separation between the goals of actual learning and testing in contemporary education (Lokon, 2005; Edwards, 2006). As well, questions were raised regarding reasons behind teachers' resistance to give up their practice-and-drill approaches of teaching and testing for deeper, more constructive ones (Sacks, 1999). More importantly, a question was also raised regarding reasons behind differences in assessment washback effects among language learners (Estaji&Tageddin, 2012).

It seems that answers to these questions reside, to some extent, in the teachers' personal epistemologies about the nature of linguistic knowledge and how it should be taught and tested. This conclusion is in line with words of James Back regarding the relationship between assessment and epistemology. In an extended dialog via Dorantes' weblog on professional testing (May, 22nd, 2006), James Bach defended the idea that testing is applied epistemology. As Bach stated "epistemology reframes testing expertise. With an epistemic focus, being a tester means something more than writing test cases. It makes one better able to slice and dice specifications, manage evolving test designs, and sound credible in one's analysis".

These personal epistemological beliefs of teachers seem to determine to a great extent their readiness to adopt deep approaches to assessment in language settings. It's quite clear that EFL teachers under consideration have naïve epistemic stances regarding the nature of knowledge and how linguistic knowledge should be taught and tested. This epistemic naiveté is strongly correlated with and have a significant impact on their assessment orientations that seem to

focus on surface level discrete point testing approaches. It is natural for teachers who view linguistic knowledge as fixed, fragmentary, and handed down by omniscient figures such as teachers and more knowledgeable others to test this knowledge using discrete point surface level techniques. Judging the dynamic features of communication as static and fragmentary leads to more accent on memorizing expressions and rehearsing utterances in simulated situations rather than constructing self-expressions in authentic contexts; the opposite holds true. EFL teachers' testing orientations will echo this very perspective.

The critical point here is that the long-advocated teachings of communicative, constructive, and self-regulated language learning approaches when assimilated in such naïve epistemological frameworks are more likely to lead to poor outcomes in both language teaching and testing. Speaking about the issue of disconnection between knowledge and practice in language testing raised by Sacks (1999), it seems that even with new teachings of communicative, constructive, and self-regulatory approaches, naïve epistemological beliefs undercut the effects of these approaches and lead teachers to recoil to their "comfort zone" of long-practiced surface-level assessment approaches. Any lasting reform in language assessment culture should render due attention to changing the epistemological beliefs of all stake-holders in the system. Otherwise, the effects of training in new approaches would be assimilated in maladaptive epistemic frameworks leading to less desired outcomes in language teaching and assessment alike.

Results of the study did not support the "ceiling effect hypothesis" in epistemic research. According to this hypothesis (Barnard, 2007:91) the epistemic development of individuals is more likely to reach its peak by the end of their formal education experience. Being at the top of the formal educational ladder, prospective EFL teachers are supposed to have developed sophisticated epistemological beliefs regarding the nature of the linguistic knowledge and how it is constructed and evaluated. Yet, it seems, based on the results of the current study, that mere progression up the educational ladder is not indicative of epistemic sophistication. Developing sophisticated epistemological beliefs relies heavily on the quality of educational experiences made available to students. When students are educated in contexts that accent rote learning and discrete point teaching and testing, they come to the conclusion that knowledge is fixed, non-dynamic, and disconnected. This view is likely to carry over to their future students in a cycle of beliefs and practices that seems to remain and reinforce itself.

6. Conclusion

On the onset of the current study a basic question was raised regarding the reasons behind EFL teachers' tendencies to adopt surface-level processing orientations in language assessment at the expense of deeper-level constructivist ones. Personal epistemologies of language teachers - their beliefs about the nature of linguistic knowledge, where it resides, and how it should be taught and tested-- were investigated as potential factors behind those teachers' assessment orientations, an area that has not heretofore received due attention in language assessment literature. Findings of the study carry evidence that these epistemological stances are core determinants of assessment practices in the foreign language settings. They determine to a great extent the type of assessment tools and the requisite information processing levels thereof. Naive epistemological beliefs are more likely to induce surface-level processing orientations; whereas sophisticated epistemological beliefs are more tuned to deeper-level assessment orientations. Any successful assessment reform needs to start with reframing these epistemological beliefs of language teachers. Otherwise, contemporary education principles and practice thereof will be assimilated in maladaptive epistemic frameworks leading to a cycle of dysfunction that reinforces itself and carries over to future generations of foreign language teachers and learner.

6.1 Implications

Results of the current study carry a set of implications that may inform various aspects of language instruction. The first of these implications concerns the need to balance surface and deep approaches to language teaching and testing. Results indicate that EFL teachers in the current study are more oriented towards surface-level processing of language learning input. Their assessment practices focus on students' ability to recall content previously presented to students or practiced in assigned textbooks. Higher level thinking or deeper processing of language input where students are asked to construct their own knowledge and think divergently in learning settings is often neglected in language assessment. Students are encouraged to emulate language teachers simply because the knowledge that will be valued is that conveyed by the teacher or included in assigned textbooks.

These shallow practices are not likely to enhance students' language learning and thus their communicative abilities. Successful learning, as Ravidran et al (2005) and Zimmerman (2002) argue, necessitates that teachers inculcate in their students a vast repertoire of strategies, both surface and deep level ones. As well, it necessitates that teachers train students in how to activate and orchestrate these strategies according to the requirements of the language tasks at hand.

Another implication relates to the need to adopt whole language approaches in language study situations. The prevailing practice herein is to study language skills and areas separately. Even if this separation is done for study purposes, it is more likely to bring about misconceptions about the nature and structure of language on the part of teachers and students alike – their epistemological beliefs. It is normal for language learners receiving decontextualized study of grammar and lexis to be oblivious of the relationship between these and other language areas and skills. They can't view language in its entirety. The same applies to the study of the different components of the language system such as syntax, vocabulary, semantics, morphology, phonology…etc.

Again, as the ultimate target of foreign language study is to develop students' communicative competence, knowledge of the language system--systemic knowledge-- alone is not enough to lead to successful communication in language use situations. According to Swain and Canale (cited in Johnson, 2001) not only should integration occur among different

subcomponents of the language system (syntax, vocabulary, phonetics and phonology, semantics, language skills…etc.), but also between these and the other two components of communicative competence, including sociolinguistic competence and strategic competence. Adopting whole language approaches in teaching and testing will help develop students' ability to communicate better, on the one hand, and promote more sophisticated epistemological beliefs about the nature of linguistic knowledge, on the other hand.

Matching types of test items to the nature of the language areas being assessed is another area that can be informed by the results of the current study. Assessing dynamic features of language such as communication and negotiation of ideas as if they were static and non-changing is more likely to bring about rote learning practices in teaching and testing and develop convergent rather than divergent thinking. In designing test items, test designers should take into account matching the type of test items or assessment tasks to the kind of requisite information processing. Convergent test items are more suitable for fixed formal features of the language system; whereas divergent assessment is more congruent with the dynamic aspects. Testing communication with multiple choice questions superficializes communication and testing alike. Likewise, testing fixed features of the language system such as grammar as if they were fluid would negatively affect students' language accuracy.

Finally, reframing the epistemological beliefs of teachers is a key to success in assessment reform. Reframing beliefs here refers to changing naïve beliefs about the nature of language and language learning into more sophisticated ones. Based on epistemic literature (Schommer-Aikins,2004; Hofer, 2004) sophisticated epistemology refers to believing in knowledge as incremental rather than bits and pieces, constructed by individuals based on evidence rather than handed down by omniscient others, and obtained from a plethora of sources rather than studied in assigned textbooks as the only source of knowledge. When teachers start seeing knowledge as relative and contextual rather than absolute or discrete, constructed by learners rather than transmitted by teachers, and negotiated rather than imposed on students, their assessment practices will change accordingly. This is particularly important in exam-driven language learning settings, where what is tested determines what is taught.

6.2 Suggestions for Further Research

As indicated by the results of the current study, personal epistemologies of foreign language teachers determine to a great extent their assessment orientations in language teaching/learning settings. It seems that research in personal epistemology would provide answers to the many pending questions in language assessment research literature. For example, test washback effects have been as issue in assessment research for their potential impact on content, scope, curricular decisions, instructional practices, and the overall educational policies. Little is known about what determines the type and gravity washback effects on these variables. Personal epistemologies of teachers and learners might be a determinant factor behind this, an area still open for further research. Other issues that research on personal epistemology and assessment needs to handle include score generalizability issues, test-taking strategies, test anxiety, and attributions of failure or success in language assessment contexts; to mention just a few.

In addition to the above suggested general areas of research on personal epistemology and assessment, there are some assessment issues that are specific to the nature of epistemological beliefs and would thus require due accent in future research. One of these issues is the domain specificity-generality issue. Research on epistemological beliefs indicate that individuals' beliefs are different across domains of knowledge or disciplines of study (Hofer, 2000; Alexander, 2001; Muis et al, 2006; Bromme et al., 2008). These domain differences in epistemological beliefs, as Hofer (2000), and Alexander (2001) argue, seem to be related to differences between the so-called "hard" fields of study such as mathematics and science and the "soft" fields of study such as the social sciences and humanities. Research findings indicate that students hold more naïve epistemological beliefs in the former than the latter. More recently, researchers found differences in epistemological beliefs even between closely related fields of study such as psychology and biological development (Estes et al., 2003) or physiology and human sciences (Knight &Mattick (2006).

The issue that arises here relates to language assessment in ESP contexts, where language (as a soft domain of knowledge) is used in teaching "hard disciplines" such as math, science, medicine …etc. Future research needs to shed light on the nature of epistemological beliefs of language teachers and learners within these contexts compared to those of English-for-academic purposes (EAP), English-for-Specific Purpose (ESP) or General-English (GE) contexts. In particular, research is needed to explore the impact that ESP teachers' epistemologies would have on their assessment practices in ESP contexts, compared to contexts of general English or English for academic purposes.

Another area that might have significant impact on language assessment is the issue of cultural specificity of personal epistemologies. As stated in the epistemic literature, differences in individuals' beliefs exist between cultures. Collectivist cultures (the current one is an example) activate epistemic stances different from those in the individualistic Western cultures (Youn, 2000; Chan &Elliott, 2002; Chan, 2003; Karabenick & Moosa, 2005; Sulimma 2009). A common conviction among researchers nowadays is that unique cultural characteristics might activate particular epistemological beliefs atypical of such environments (Clarebout, 2001) and that language teachers' beliefs and conceptualizations are formulated from experiences that they have within a socio-cultural context (Flores, 2001; Arkoudis, 2003; Ali & Ammar, 2005). Here, a number of issues might arise regarding language assessment including, for example, the generalizability of assessment principles and perspectives across language settings in different cultural contexts, as well as the congruence between the internationally advocated assessment knowledge and principles and the locally oriented policies, philosophies, and practices. Further research is needed on these areas.

Finally, some issues have not been addressed here due to the nature of delimitations in the current study and should be taken into account when generalizing conclusions to other contexts. For example, the current study involved female prospective teachers only. With reported differences in the literature between males and females regarding epistemic positions and ways of knowing (Hofer, 2000; Baxter-Magolda, 2004), future research is needed to investigate the assessment orientations of male language teachers and how different they might be compared to the ones reported here.

Moreover, the current study involved novice teachers whose assessment orientations might have not fully developed due to lack of experience and knowledge of the "actual classroom culture". Further research on in-service teachers is needed to see if teacher experience has any potential impact on personal epistemologies, on the one hand, and on their assessment orientations, on the other hand.

Likewise, the current study was concerned with personal epistemologies and assessment orientations in foreign language contexts. With the documented differences between L1 and FL instructional settings (Wolter, 2001; MacWhinney, 2008; Clark, 2009) further research is needed to investigate if these differences would interfere with teachers' epistemologies regarding linguistic knowledge and consequently with their assessment orientations in these contexts.

References

Abedi, J. & Linquanti, R. (2012). Issues and opportunities in strengthening large scale assessment systems for ELLs. Paper presented at the Understanding Language Conference, Stanford University, CA.

Aftab, A. Qureshi, S. & William, I. (2013). Investigating the washback effects of the Pakistani intermediate English examination. *International Journal of English and Literature, 5*(7), 149-154.

Akpinar, K. D. & Cakildereb, B. (2013). Washback effects of high-stakes language tests of Turkey (KPDS and ÜDS) on productive and receptive skills of academic personnel. *Journal of Language and Linguistic Studies, 9*(2), 81-94.

Alexander, P. A. (2001). In the year 2020: Envisioning the possibilities for educational psychology. *Journal of Educational Psychology, 84,* 261-271.

Ali, M. F, A. & Ammar, A. M. I. (January 2005). An investigation of the relationships between EFL pre-service teachers' epistemological beliefs and their learning strategies, teaching practices and foreign language classroom anxiety. Journal of Scientific Research in Education and Psychology (El-Minia Faculty of Education, Egypt) 18(3) 1-33.

Ammar, A. M. I. & Abdel-Majeed, U. M. (2006). Predicting gifted EFL students' goal orientation, cognitive engagement, perceived linguistic competence, and achievement with epistemological beliefs. Paper presented at the Regional Scientific Conference on Giftedness and Creativity, Jeddah, Saudi Arabia, August, 28-30.

Amrein, A.L. & Berliner, D.C. (2002, March 28). High-stakes testing, uncertainty, and student learning. Educational Policy Analysis Archives, 10 (18). Retrieved January 10, 2006, from http://epaa.asu.edu/epaa/v10n18/.

Andrews, S. (2004). Washback and curriculum innovation. In L. Cheng, Y. Watanabe, & A. Curtis (eds.), Washback in language testing: Research contexts and methods (pp. 37-50). Mahwah, NJ: Erlbaum.

Apichatrojanakul, P. (2011). The Washback Effects of the TOEIC Examination on the teachers and students of a Thai Business School. *Language Testing in Asia, 1*(1), 62-75.

Arkoudis, S. (2003). Teaching English in science classes: Incommensurate epistemologies? *Language and Education,17*(5), 161-173.

Bachman, L. F. (2006, September). Issues in assessing the English proficiency and academic   achievement of English language learners. Paper presented at the Mid-West Association of Language Testers Conference, Urbana-Champaign, IL.

Bailey A. L., Butler, F. A., Stevens, R., & Lord, C. (2007). Further specifying the language demands of school. In A. L. Bailey (Ed.), The language demands of school: Putting academic English to the test (pp. 103-156). CT: Yale University Press.

Bailey, A. L. & Huang, B. H. (2011). Do current English language development/ proficiency standards reflect the English needed for success in school*? Language Testing,* 28, 343-365.

Barnard, L. (2007). The expert ceiling in epistemological beliefs. *Essays in Education*, 19, 85-95.

Baxter-Magolda, M. B. (1992). Knowing and reasoning in college: Gender-related patterns in students' intellectual development. San Francisco: Jossey-Bass.

Baxter-Magolda, M. B. (2004). Evolution of a constructivist conceptualization of epistemological reflection. *Educational Psychologist, 39*(1), 31–42.

Birenbaum, M. (2003). New insights into learning and teaching and their implications for assessment. In M. Segers, F. Dochy & E. Cascallar (Eds.), Optimizing new modes of assessment: In search of qualities and standards (PP. 37-54). Boston: Kluwer Academic Publishers.

Black, P. J., Wilson, M., & Yao, S. (2011). Road maps for learning: A guide to the navigation of learning progressions. *Measurement, 9*(2-3), 71-123.

Boekaerts, M. & Minnaerts, A. (2003). Assessment of students' feelings of autonomy, competence, and social readiness: A new approach to measuring the quality of the learning process through self-and peer-assessment. In M. Segers, F. Dochy & E. Cascallar (Eds.), Optimizing new modes of assessment: In search of qualities and standards (PP. 170-225). Boston: Kluwer Academic Publishers.

Braten, I. & Stromso, E. (2009). Personal epistemology across cultures: Exploring Norwegian and Spanish university students' epistemic beliefs about climate change. *Social Psychology of Education, 12*(4), 529-560.

Braun, H. (2003). Assessment and technology. In M. Segers, F. Dochy & E. Cascallar (Eds.), Optimizing new modes of assessment: In search of qualities and standards. Boston: Kluwer Academic Publishers.

Braun, H. & Kanjee, A. (2006). Using assessment to improve education in developing nations. In H. Braun, A. Kanjee, E. Bettinger & M. Kremer (Eds.), Improving education through assessment, innovation, and evaluation (pp.1-46). Cambridge: American Academy of Arts & Sciences.

Brindley, G. (2001). Assessment. In R. Carter & D. Nunan (Eds.), Teaching English to speakers of other languages (PP. 137-143). Cambridge: CUP.

Brown, H. D. (2003). Language assessment: Principles and classroom practices. New York: Longman.

Brown, H. D. (2005). Testing in language programs: A comprehensive guide to English language assessment (New ed.). New York: McGraw-Hill College.

Brownlee, J. M. (2001). Epistemological beliefs in pre-service teacher education students. *Higher Education Research & Development, 20*(3), 281-291.

Brownlee, J. M. (2001). Knowing and learning in teacher education: A theoretical framework of core and peripheral beliefs in application. *Asia-Pacific Journal of Teacher Education & Development 4*(1), 131-155.

Brownlee, J. P. & Boulton L. G. (2001) Changing epistemological beliefs in pre-service teacher education students. *Teaching in Higher Education, 6*(2), 247-268.

Buehl, M. M. & Alexander, P. A. (2001). Beliefs about academic knowledge. *Educational Psychology Review, 13*(4), 385 – 418.

Burrows, C. (2004). Washback in classroom-based assessment: A study of the washback in Australian adult migrant English program. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), Washback in language testing: Research contexts and methods (PP.113-128). New Jersey: Lawrence Erlbaum Associates.

Campose, M. (2007). Ecology of meanings: A critical constructivist communication model. *Communication Theory, 17*(4), 386-410.

Carbery, S. (1999). Practicalities of ongoing assessment. *JALT Testing & Evaluation, 3*(1), 3-8.

Chan, k. (2002). Students' epistemological beliefs and approaches to learning. Paper presented at the AARE2002 conference held in Brisbane, Australia from 1-5 December.

Chan, K. (2003, November). Pre-service teachers' epistemological beliefs and conceptions about teaching and learning: Cultural implications for research in teacher education. Paper presented at the NZARE AARE Conference, Auckland, New Zealand.

Chan, K. & Elliott, R.G. (2002). Exploratory study of Hong Kong teacher education students 'epistemological beliefs: Cultural perspectives and implications on beliefs research. *Contemporary Educational Psychology,* 27, 392–414.

Chapelle, A. C. & Brindley, G. (2002). Assessment. In N. Schmitt (Ed.), An introduction to applied linguistics, (PP. 267-288). London: Arnold

Chapman, D.W., Snyder, C.W. (2000). Can high stakes national testing improve instruction: Reexamining conventional wisdom. *International Journal of Educational Development 20*, 457–474.

Cheng, L. (2004). The washback effect of a public examination change of teachers' perceptions toward their classroom teaching. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), Washback in language testing: Research contexts and methods (PP.147-170). New Jersey: Lawrence Erlbaum Associates.

Cheng, L. (2005). Changing language teaching through language testing: A washback study. New York: Cambridge University Press.

Cheng, L. (2008). Washback, impact and consequences. In E. Shohamy & N. H. Hornberger (Eds.), Encyclopedia of language and education (Vol. 7): Language testing and assessment (PP. 349-364). Berlin: Springer Press.

Cheng, L. & Watanabe, Y. (2004). Washback in language testing: Research contexts and methods. New Jersey: Laurence Erlbaum & Associates.

Choi, I. (2008). The impact of EFL testing on EFL education in Korea. *Journal of Language Testing, 25*(1), 39-62.

Clarebout, G., Elen, J., Luyten, L. & Bamps, H. (2001). Assessing epistemological beliefs: Schommer's questionnaire revisited. *Educational Research and Evaluation, 7*(1), 53-77.

Clark, E. V. (2009). First language acquisition (2nd Ed.). Cambridge: Cambridge University Press.

Coombe, C. A., & Hubley, N. J. (2003). *Assessment practices.* Alexandria, VA: TESOL.

Cunningham, C. C. (2002). The TOEIC test and communicative competence: Do test score gains correlate with increased competence? Unpublished master's thesis, University of Birmingham, Birmingham, United Kingdom. Retrieved January 20, 2006, from http://www.cels.bham.ac.uk/resources/essays/cunndiss.pdf.

Datnow, A., Hubbard, L. & Mehan, H. (2002). *Extending educational reform.* London: Rutledge Falmer.

Davidson, F., Kim, J. T., Lee, H., Li, J., & Lopez, A. (2007). Making choices in academic English language testing: Evidence from the evolution of test specifications. In A. L. Bailey (Ed.), The language demands of school: Putting academic English to the test (PP. 157-170). New Haven, CT: Yale University Press.

Davies, A. & LeMahieu, P. (2003). Assessment for learning: Reconsidering portfolios and research evidence. In M. Segers, F. Dochy & E. Cascallar (Eds.), Optimizing new modes of assessment: In search of qualities and standards (PP. 141-170). Boston: Kluwer Academic Publishers.

Donald, J. G. (2002). Learning to think: Disciplinary perspectives. San Francisco: Jossey-Bass.

Donmoyer, R. (2001). Paradigm talk reconsidered. In V. Richardson (Ed.), Handbook of research on teaching (PP. 174-197). Washington, DC: American Education Research Association.

Dorantes, M. (2006). Professional testing = applied epistemology. Professional Weblog. Available online at: http://blogs.msdn.com/b/marcod/archive/2006/05.aspx. Retrieved on Wednesday, July 2nd, 2010.

Dori, Y. J. (2003). A framework for project-based assessment in science education. In M. Segers, F. Dochy & E. Cascallar (Eds.), Optimizing new modes of assessment: In search of qualities and standards (PP. 89-118). Boston: Kluwer Academic Publishers.

Edwards, N. (2004). Rediscovering the creative heart of Japanese education: Fostering intrinsic motivation though a love of language. *The Language Teacher, 28*(1), 19-23.

Edwards, N. (2006). The historical and social foundations of standardized testing: In search of a balance between learning and evaluation. JALT Testing & Evaluation SIG Newsletter, 10(1), 8-16.

Estaji, M. & Tageddin, Z. (2012). The learner factor in washback context: An empirical study investigating the washback of the IELTS academic writing test. *Language Testing in Asia, 2*(1), 5-25.

Estes, D., Chandler, M., Horvath, K. J., & Backus, D. W. (2003). American and British college students' epistemological beliefs about research on psychological and biological development. Applied Developmental Psychology, 23, 625-642.

Flinders, D. J. (2005). The failings of NCLB. *Curriculum and Teaching Dialogue, 7*(1), 1-9.

Francis, D. J., Rivera, M., Lesaux, N., Kieffer, M., & Rivera, H. (2006). Practical guidelines for the education of English language learners: Research-based recommendations for the use of accommodations in large-scale assessments. Portsmouth, NH: RMC Research Corporation.

Gallagher, C. J. (2003). Reconciling a tradition of testing with a new learning paradigm. *Educational Psychology Review, 15*(1), 83-99.

Gibbs, G. (2002). Evaluating the impact of formative assessment on student learning behavior. Paper presented at the EARLI/Northumbria Assessment conference, Longhurst, UK, August 29.

Gielen, S., Dochy, F. & Dierick, S. (2003). Evaluating the consequential validity of new modes of assessment: The influence of assessment on learning, including pre-, post- and true assessment effects. In M. Segers, F. Dochy & E. Cascallar (Eds.), Optimizing new modes of assessment: In search of qualities and standards (PP. 37-54). Boston: Kluwer Academic Publishers.

Green, A. (2007). Washback to learning outcomes: A comparative study of IELTS preparation and university pre-sessional language courses. *Assessment in Education, 14*(1), 75-97.

Green, A. (2013). Washback in language assessment. *International Journal of English Studies, 13*(2), 39-52.

Greenwald, A., Bosek, B. & Sriram, N. (2006). Consequential validity of the implicit association test. American Psychologist, 16(1), 56-61.

Hammer, D. (2002, April). Epistemological resources. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans.

Handley, K., & L. Williams. (2011). From copying to learning: using exemplars to engage students with assessment criteria and feedback. Assessment and Evaluation in Higher Education 36, no. 1: 95-108.

Hawkey, R. (2004). An IELTS Impact Study: Implementation and some early findings. Research Notes, 15. Cambridge: Cambridge ESOL.

Hayes, B. & Read, J. (2003). IELTS test preparation in New Zealand. Preparing students for the IELTS academic modules. In L. Cheng, Y. Watanabe & A. Curis (Eds.), Washback in language Testing: Research contexts and methods. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Hayes, D. & Wynyard, R. (2002). The McDonaldization of higher education. CT: Bergin & Garvey.

Herman, J. L. (2004). The effects of testing on instruction. In S. Fuhrman & R. Elmore (Eds.), Redesigning accountability systems for education. New York: Teachers College Press.

Higgins, R., P. Hartley & Skelton, A. (2001). Getting the message across: The problem of communicating assessment feedback. *Teaching in Higher Education 6*(2), 269- 274.

Higgins, R., P. Hartley & Skelton, A. (2002). The conscientious consumer: Reconsidering the role of assessment feedback in student learning. *Studies in Higher Education, 27*(1), 53-64.

Hillocks, J. (2002). The testing trap: How state writing assessments control learning. New York: Teachers College Press.

Hofer, B. K. (2000). Dimensionality and disciplinary differences in personal epistemology. *Contemporary Educational Psychology, 25*, 378-405.

Hofer, B. (2001). Personal epistemology research: Implications for learning and teaching. *Journal of Educational Psychology Review, 13*(4), 353-83.

Hofer, B. (2004): Epistemological understanding as a metacognitive process: Thinking aloud during online searching. *Educational Psychologist, 39*(1), 43-55.

Hughes, A. (2002). Testing for language teachers (revised ed.). Cambridge: Cambridge University Press.

Johnson, K. (2001). An introduction to foreign language learning and teaching. Essex: Pearson Education Limited.

Karabenick, S. & Moosa, S. (2005). Culture and personal epistemology: US and Middle Eastern students' beliefs about scientific knowledge and knowing. *Social Psychology of Education, 8*, 375-93.

Kardash, C. M. & Howell, K. L. (2000). Effects of epistemological beliefs and topic-specific beliefs on undergraduates/cognitive and strategic processing of dual-positional text. Journal of Educational Psychology, 92, 524-535.

Knight, L. V. & Mattick, K. (2006). When I first came here, I thought medicine was black and white: Making sense of medical students' ways of knowing. *Social Science & Medicine, 63*(4), 1084-1096.

Knight, P. T. (2002). Summative assessment in higher education: Practices in disarray. *Studies in Higher Education, 27*(3), 275-287.

Liu, L. & Zhang, Y. (2014). The application of constructivism to the teaching of intercultural communication. *English Language Teaching, 7*(5), 136-141.

Lokon, E. (2005). Will the New Center Test make English language education more communicative in Japanese schools? *The Language Teacher, 29*(11), 7-12.

Loyens, S. Rikers, R. & Schmidt, S. (2009). Students' perceptions of constructivist learning in different programme years and different learning environments. *British Journal of Educational Psychology, 79*(3), 501-514.

Luna, C., & Turner, C. (2001). The impact of the MCAS: Teachers talk about high-stakes testing. English Journal, 91 (1), 79-87.

Luxia, Q. (2007). Is testing an efficient agent for pedagogical change? Examining the intended washback of the writing task in a high-stakes English test in China. Assessment in Education: Principles, Policy & Practice, 14 (1), 51-74.

MacWhinney, B. (2008). A unified model. In N. Ellis & P. Robinson (Eds.),Handbook of Cognitive Linguistics and Second Language Acquisition (341-371). New York: Rutledge.

Many, J. E., Howard, F. & Hoge, P. (2002). Epistemology and pre-service teacher education: How beliefs about knowledge affected our students' experiences. *English Education, 34*(4), 302-322.

Masters, G. (2013). *Reforming educational assessment: Imperatives, principles, and challenges.* Victoria, Australia: ACER Press.

McNamara, T. (2000). *Language testing.* Oxford: Oxford University Press.

Moon, T., Brighton, C., & Callahan, C. (2003). State standardized testing programs: Friend or foe of gifted education? (On gifted students in school). *Roeper Review, 25*(2), 49-60.

Munoz, A. & Alvarez, M. (2010). Washback of an oral assessment system in the EFL classroom. Language Testing, 27 (11), 33-49.

O'Sullivan, B. (2006). *Issues in testing business English. Studies in language testing* (Vol. 17). Cambridge: Cambridge University Press.

Payne, E. & Brown, G. (2011). Communication and practice with examination criteria: Does this influence performance in examinations? *Assessment and Evaluation in Higher Education, 36*(6), 619-626.

Popham, W. (2001). *The truth about testing: An educator's call to action.* VA: Association for Supervision and Curriculum Development.

Qi, L. (2004). Has a high-stakes test produced the intended changes? In L. Cheng & Y. Watanabe (Eds.). *Washback in language testing: Research contexts and methods* (PP. 171-190). New Jersey: Laurence Erlbaum & Associates.

Qian, G. & Alvermann, D. (2000). Relationship between epistemological beliefs and conceptual change learning. *Reading and Writing Quarterly*, 16, 59-74.

Ravindran, B., Greene, B. & DeBacker, A. (2005). Predicting preservice teachers' cognitive engagement with goals and epistemological beliefs. *Journal of Educational Research, 98*(4), 222-232.

Rea-Dickins, P. & Scott, C. (2007). Washback from language testing on teaching, learning, and policy: Evidence from diverse settings. *Assessment in Education: Principles, Policy & Practice, 14*(1), 1-7.

Rivera, C., Collum, E., Shafer-Willner, L., & Sia, J. K. (2006). An analysis of state assessment policies regarding the accommodation of English language learners. In C. Rivera & E. Collum (Eds.), State assessment policy and practice for English language learners: A national perspective (pp. 1-173). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Sacks, P. (1999). The high price of America's testing culture and what we can do to change it. New York: Da Capo Press.

Saville, N. & Hawkey, R. (2003). The IELTS impact study: Investigating washback on teaching materials. In L. Cheng, Y. Watanabe & A. Curis (Eds.), Washback in language Testing: Research contexts and methods. Mahwah, N.J.: Lawrence Erlbaum Associates, Publishers.

Schell, G. & Janicki, T. (2013). Online course pedagogy and the constructivist learning model. Journal of the Southern Association for Information Systems, 1(1). Retrieved online on July 23, 2016 from http://dx.doi.org/10.3998/jsais.11880084.0001.104.

Scherer, M. (2005). Reclaiming testing. *Educational Leadership, 63*(3), 9-15.

Schommer, M. (1994). Synthesizing epistemological belief of research: Tentative understandings and provocative confusions. *Educational Psychology Review, 6*(4), 293-319.

Schommer-Aikins, M, Duell, O. & Barker, S. (2003). Epistemological beliefs across domains using Biglan's classification of academic disciplines. *Research in Higher Education, 44*(3), 347-66.

Schommer-Aikins, M. (2004). Explaining the epistemological belief system: Introducing the embedded systemic model and coordinated research approach. *Educational Psychologist, 39*(1), 19-29.

Schreiber, J. B. & Shinn, D. (2003) Epistemological beliefs of community college students and their learning processes. *Community College Journal of Research and Practice, 27*(8), 699-709.

Segers, M. (2003). Evaluating the overall test. In M. Segers, F. Dochy& E. Cascallar (Eds.), Optimizing new modes of assessment: In search of qualities and standards (PP.119-140). Boston: Kluwer Academic Publishers.

Segers, M., Dochy, F. & Cascallar, E. (2003). The era of assessment engineering: changing perspectives on teaching and learning and the role of new modes of assessment. In M. Segers, F. Dochy & E. Cascallar (Eds.), Optimizing new modes of assessment: In search of qualities and standards (PP. 1-12). Boston: Kluwer Academic Publishers.

Solano-Flores, G., & Trumbull, E. (2003). Examining language in context: The need for new research and practice paradigms in the testing of English-language learners. *Educational Researcher, 32*(2), 3-13.

Solano-Flores, G., & Li, M. (2006). The use of Generalizability (G) Theory in the testing of linguistic minorities. Educational Measurement: *Issues and Practice, 25*(1), 13-22.

Sulimma, M. (2009). Relations between epistemological beliefs and culture classifications. *Multicultural Education & Technology Journal, 3*(1), 74 – 89.

Taylor, L. (2006). The changing landscape of English: Implications for language assessment. *ELT Journal, 60*(1), 51-60.

Topping, K. (2003). Self and peer assessment in school and university: Reliability, validity, and utility. In M. Segers, F. Dochy & E. Cascallar (Eds.), Optimizing new modes of assessment: In search of qualities and standards (PP. 55-88). Boston: Kluwer Academic Publishers.

Tsai, C. (2000). The Effects of STS-oriented instruction on female tenth graders' cognitive structure outcomes and the role of student scientific epistemological beliefs. International Journal of Science Education, 22(10), 1099-1115.

Turner, C. (2006). Professionalism and high-stakes tests: Teachers' perspectives when dealing with educational change introduced through provincial exams. *TESL Canada Journal, 23(*2), 54-75.

Van-Lier, L. & Walqui, A. (2012). How teachers and educators can most usefully and deliberately consider language. Paper presented at the Understanding Language Conference, Stanford University, CA.

Wall, D. (2000). The impact of high-stakes testing on teaching and learning: Can this be predicted or controlled? System, 28, 499-509

Wall, D. (2005). *The impact of high-stakes examination on classroom teaching: A case study using insights from testing and innovation theory*. Cambridge: Cambridge University Press.

Wall, D. & Horak, T. (2006). The impact of changes in the TOEFL examination on teaching and learning in Central and Eastern Europe: Phase 1, the baseline study. TOEFL Research Report (RR-06-18). Princeton: New Jersey.

Wang, J. & Zhang, Y. (2012). Practical application of constructivism and metacognition in computer-aided college English teaching (V.43). In International Conference on Education Technology and Computer. Singapore: IACSIT Press.

Wolf, M. K., Herman, J. L., Bachman, L. F., Bailey, A. L., & Griffin, N. (2008). Recommendations for assessing English language learners: English language proficiency measures and accommodation uses—Recommendations (CRESST Tech. Rep. No. 737). Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Wolter, B. (2001). Comparing the L1 and L2 mental lexicon: A depth of individual word knowledge model. *Studies in Second Language Acquisition, 23*(1), 41-69.

Yıldırım, Ö. (2010). Washback effects of a high-stakes university entrance exam: Effects of the English section of the university entrance exam on future English language teachers in Turkey. *Asian EFL Journal,12*(2), 92-116.

Yoshida, K. (2006). Theoretical frameworks of testing in SLA: Processing perspectives and strategies in testing situations. *JALT Testing & Evaluation, 10*(1), 2-7.

Youn, I. (2000). The culture specificity of epistemological beliefs about learning. Asian Journal of Social Psychology, 3, 87–105.

Youn, I., Yang, K. & Choi, I. (2001). An Analysis of the nature of epistemological beliefs: Investigating factors affecting the epistemological development of South Korean high school students. Asia Pacific Education *Review, 2*(1), 10-21.

Zhao, Z. (2013). An investigation into the consequential validity of a diagnostic college English speaking test. English Language Teaching, 6(2), 171-78.

Zimmerman, B. (2002). Achieving academic excellence: A self-regulatory perspective. In M. Ferrari (Ed.), *The pursuit of excellence through education: The educational psychology series*. (pp. 85-110). New Jersey: Erlbaum.