Assessing Specific Language Ability: A Theoretical Framework

Mohammed Naoua

Department of Arts and English Language, Faculty of Arts and Languages
Chahid Hamma Lakhdhar University, Eloued, Algeria

ABSTRACT

Assessing specific language ability can be defined as the process of administering tests to specific types of test takers in order to gather information about their levels of language competence, and about their capacity of using this competence in situations similar to their fields of interest whether academic or job-related. This paper attempts to provide a theoretical 'Doglassan' framework, which can be used as a checklist for describing and isolating the specific constructs to be tested, and for telling how these constructs can be tested. The research concludes with a set of recommendations relevant to testers, test takers, and test tasks. In this context, the paper recommends the spread of assessment and measurement literacies amongst teachers. It also calls for the identification of test takers' specific communicative needs, and their levels of language and background knowledge before test design. Regarding tasks, and for reinforcing the concept of interactiveness between test takers' language ability and the test input, the research recommends the enrichment of test content with subject-specialist cues.

Key Words: Assessment, Communicative competence, Languages ability, Measurement, Tests.

INTRODUCTION

Assessment in languages for specific purposes (LSP) refers to the process of making inferences about test takers' sectors of language knowledge and the extent to which they can use this knowledge in specific target contexts related to their academic specialties, or to their present time occupational domains (Bachman, 1990; Bachman and Palmer, 1996; Douglas, 2000, 2001, 2005). For obtaining valid and reliable scores reflecting the abilities to be measured and reducing the factors, which may undermine the credibility of the score interpretations, two main criteria need to be provided. First, the testing instruments no matter how narrow their scope, should be informed with detailed models that describe the abilities intended to be evaluated. Second, providing frameworks that explain how these abilities can validly be tested (Alderson, 2000; Purpura, 2004; Douglas, 2010, 2013).

In an attempt to explain how this process can be conducted, the researcher organizes his paper into seven sections. Section one reviews some of the literature relevant to communicative language ability from the perspectives of sociolinguistics. Section two suggests some definitions of LSP, and delimits its participants and contexts. Section three focuses on describing the process of LSP assessment, and on highlighting the major features of its tests. Section four introduces Douglas's (2000) three-componential model of specific language ability (SLA). Section five delineates the types and approaches of the constructs be tested; and section six proposes the facets which explain how these constructs can be measured. Finally, in section seven we propose a set of recommendations aimed at reinforcing the concept of validity in ESP tests.

1. The Conceptualization of Communicative Language Ability (CLA)

Since Hymes' seminal article 'On Communicative Competence' (1972), the notion of communicative language ability (CLA) has been theorized with respect to the knowledge of the rules of grammar, and the capacity of implementing those rules in real contexts (Widdowson, 1983). According to Hymes (1972), conducting a communicative purpose requires competence for grammar and competence for use. In the same way, Canale and swain (1980: 6) distinguish two types of competence: communicative competence and communicative performance. The former encompasses grammatical and sociolinguistic competencies. However, the latter refers to

"the realization of these competencies and their interaction in the actual production and comprehension of utterances". On his part, Widdowson (1983: 8) characterizes (CLA) as the capacity "to produce and understand *utterances* by using the resources of the grammar in association with features of context to make meaning". We conclude with Bachman (1990: 84) who sees this ability as the implementation one's knowledge of the rules of grammar in "appropriateness with contextualized communicative language use".

2. Definition of LSP

Teaching languages for specific purposes can be defined as the process of training or instruction that is designed for homogeneous groups of learners or trainees on the basis of an extensive analysis of their specific academic or professional communicative needs (Basturkmen and Elder, 2004; Basturkmen, 2006, 2010; Naoua, 2016b). The latter refer to the real-world situations, which learners or trainees may encounter whether in the work place, or during their further studies. It is worth mentioning here that this type of instruction can be provided during pre-service, in-service, or post-service courses (Hutchinson and Waters, 1987; Dudley-Evans and St. John, 1998; Basturkmen and Elder, 2004; Hyland, 2006, 2009; Douglas, 2010b; Hyland and Shaw, 2016).

Applied linguists organize LSP, or more specifically ESP into two main divisions: English for academic purposes (EAP), and English for occupational purposes (EOP) (Mumby, 1978; Basturkmen, 2010; Douglas, 2013). The former concerns the persons who use the language as a medium of instruction for acquiring subject-specialist academic education. However, the latter concerns those who use language in areas relevant to the specificity of their working place.

Maintaining the same division but with a slight modification, Hutchison and Waters (1987) propose a three-fold arrangement which includes English for science and technology (EST), English for business and economics (EBE), and English for social sciences (ESS) (see fig 1.). Each of these sub-divisions is further organized into English for academic purposes, and English for occupational purpose.

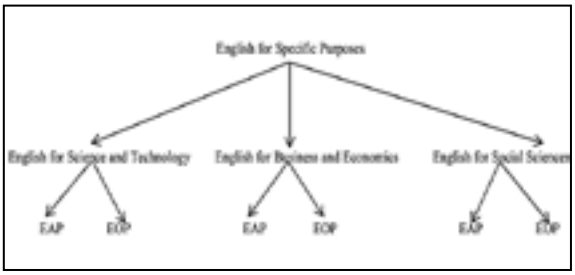

Fig: 1: Naoua , 2016b: 4

2.1. Key Features of LSP

Two key features need to be identified in LSP teaching and testing: describing language use in the specific target domains, and needs analysis (Bachman and Palmer, 1996; Basturkmen and Elder, 2004; Basturkmen, 2006, 2010). Language use can be defined as the "creation or interpretation of intended meanings in discourse by an individual, or as the dynamic and interactive negotiation of intended meanings between two or more individuals in a particular situation" (Bachman and Palmer, 1996: 62). However, target language use domains refer to the sets of specific tasks that ESP learners or test takers can encounter outside the classroom environment, or the testing situation; and to which test designers and users are supposed to generalize the score interpretations. Concerning needs analysis, it can be conceived as the methodical gathering of "information about the communicative demands faced by those in the target situation" (Basturkmen and Elder, 2004: 674). The link between learners' needs and course or test design is of great importance because when the test content is highly field-specific, it will engage test takers' background knowledge to be involved in the testing situation (Douglas, 2000; Bloor and Bloor, 2004; Hyland, 2006, 2009; Basturkmen, 2006, 2010; Hyland and Shaw, 2016).

3. LSP Testing

LSP tests refer to the measurement instruments, which are administered to homogeneous groups of test takers in order to quantify their specific mental language traits and/ or to examine the extent to which they can use language in specific target contexts similar to their fields of study, or relevant to their occupational domains. In order to ensure a high level of validity, these measurement instruments should derive their content and tasks from a wide-range analysis of the examinees' specific target needs (Bachman, 1990; Bachman and Palmer, 1996; Basturkmen and Elder, 2004; Douglas, 2000, 2013; Hyland, 2006, 2009; Basturkmen, 2006, 2010; Hyland and Shaw, 2016).

3.1. Aspects of LSP Tests

Language testers delineate three features that are specific to communicative specific tests. These include specification of content, authenticity of task, and interaction between background knowledge and language knowledge (Basturkmen and Elder, 2004; Douglas, 2000, 2010a, 2013). Concerning the first feature, Douglas (2000: 90) reasons that measuring specific language ability requires the interaction between test takers' language knowledge on the one hand, and the specific content of the test on the other. Consequently, he recommends that ESP "test input must be rich in specific purpose cues to help insure that test takers will engage an appropriate discourse domain [because] the more highly specialized test content becomes, the greater the influence of specific purpose background knowledge". Specificity of content in ESP tests can be achieved with the guidance of subject-specialist informants.

As for the second feature 'authenticity of task', this can be delineated with reference to two factors: the extent of correspondence between test tasks and real-life tasks in specific target domains, as well as the extent of interaction between test takers' internal language traits and the test input (Douglas, 2000, 2001, 2005; Bachman and Palmer, 2010). In this context, Bachman (1991) proposes two types of authenticity: situational and interactional. The former measures the extent of relevance of test content and tasks to the real target context tasks. Bachman explains, "If test takers were specialists in engineering, it is likely that inclusion of technical terms and topics from engineering would tend to increase the situational authenticity of the test" (690). The author distinguishes this type from the real-life approach, which achieves authenticity by sampling real-world tasks and incorporating them in language tests. Concerning interactional authenticity, it can be seen as "a function of appropriate response and is realized when sender and receiver engage in interaction mediated by the language" (Widdowson, 1979: 162). According to Bachman (1991: 691), situational authenticity focuses "on the relationship between the test task and nontest language use [while] interactional authenticity resides in interaction between the test taker and the test task".

The third characteristic of ESP tests is background knowledge. The latter can be defined as the long-term specific-subject knowledge that learners or workers have internalized as a result of their study of their topics of specialty, or of their occupation in specific domains (Bachman, 1990; Bachman and Palmer, 1996; Douglas, 2010a). In 'general communicative competence tests', background knowledge can be considered as construct irrelevant difficulty, which, "leads to construct scores that are invalidly low for those individuals adversely affected" (Messick, 1995: 742). Conversely, when the test input is highly familiar, whether 'deliberately or inadvertently', to some test takers rather than others, we can speak of construct irrelevant easiness. The latter can lead "to scores that are invalidly high for the affected individuals as reflections of the construct under scrutiny" (Messick, 1995: 743). However, in LSP testing, and due to the homogeneity of tests, background knowledge is considered as a part of the construct to be measured.

4. Describing Specific Language Ability (SLA)

There is almost a unanimous point of view on the part of Language testers that every test has a model of language ability behind it (Bachman, 1990; Alderson *et al*.,1995; Bachman and Palmer, 1996; Alderson, 2000; Fulcher and Davidson, 2007, 2009, 2012). This model or theory refers to some "abstract theoretical descriptions of what it means to be able to communicate in a second language" (Fulcher and Davidson, 2007:36). In the same way, these testers see that "every test is an operationalization of some beliefs about language, whether the constructor refers to an explicit model or merely relies upon 'intuition'" (Alderson *et al*., 1995: 16-17). However, their divergence is on whether LSP teaching and testing would stand on a theoretical description of the ability to be taught, or tested. The first trend, which includes linguists such as Widdowson (1979, 1983, 1984, 2003), Hutchinson and Waters (1987), Dudley-Evans and St. John (1998), and Basturkmen and Elder (2004) sees LSP teaching as atheoretical, and represents only a special case of communicative language teaching. Additionally, in their point of view, "all uses of English, as any other language, are specific [and] all uses of language serve a particular purpose" (Widdowson, 2003: 61). Conversely, the other trend represented by Bachman (1990), Bachman and Palmer (1996), Alderson and Bachman (2000-2006), and Douglas (2000, 2010a, 2010b, 2013) strongly argues that:

These assertions are not true, that there is a theoretical justification for ESP, that ESP is different from general purpose language, that language knowledge and specific purpose background knowledge are both part of the ESP construct, and that specific purpose language testing is not only possible but necessary (Douglas, 2010b: 3).

4.1. Douglas's Model of Specific Language Ability

Specific language ability (SLA) in Douglas's (2000) model builds largely upon Bachman (1990, 1991), and Bachman and Palmer (1996). As it is included in Fig 2, Douglas (2000) proposes a three-constituent model comprising language knowledge (LK), strategic competence (SC), and background knowledge (BK). Language knowledge includes the long-term linguistic traits internalized by language learners during their academic study, or language acquisition (Purpura, 2004). Background knowledge refers to the subject-specialist information that learners or trainees have acquired during their study, or work in a given specific field (Douglas, 2010a). Concerning strategic competence, it is conceived as a set of mental and communication strategies, which on the one hand enable language knowledge to interact with background knowledge, and allow these mental traits to interact with the external context, or the specific purpose input to create and interpret discourse on the other.

Fig 2: Naoua, 2016a: 62

4.1.1. Language Knowledge

Language knowledge can be defined as the "domain of information in memory that is available for use by the metacognitive strategies in creating and interpreting discourse" (Bachman and palmer, 1996: 67). Language knowledge is further subdivided into organizational and pragmatic knowledge. The former enables language users or test takers to produce or understand grammatically acceptable utterances, and to combine sentences to form texts or to move from sentence-level to discourse construction. However, the latter allows us to create and comprehend discourse by relating it to the intentions of participants and to the

characteristics of the context it occurs in (Alderson, 2000).

4.1.1.1 Organizational Knowledge

On its turn, organizational knowledge is organized into two types: grammatical and textual. The first type, which helps us in producing or understanding correct and accurate sentences or utterances, encompasses awareness of knowledge of vocabulary, syntax, phonology, and graphology. The second type, which allows us to combine sentences or utterances to form texts or parts of discourse, comprises knowledge of cohesion and conversational organization (Luoma, 2004; Mackay, 2006).

4.1.1.2 Pragmatic Knowledge

Concerning pragmatic knowledge, it enables language users to produce and comprehend "utterances or sentences and texts to their meanings, to the intentions of language users, and to relevant characteristics of the language use setting" (Bachman and palmer, 1996: 69). We can speak of two large areas of pragmatic knowledge: functional and sociolinguistic. The former determines how utterances/sentences or responses can be related to language users or examinees' intentions. The latter delineates how the meaning of sentences or utterance can be interpreted with reference to the social context.

4.1.1.3.1 Functional Knowledge

Functional knowledge covers ideational, manipulative, heuristic**,** and imaginative functions (Bachman, 1991; Halliday, 1973, 2002, 2004; Purpura, 2004). Ideational functions enable people to express and understand the meaning of utterances based on their experience of the world around them. Manipulative knowledge, which enables participants to use language to affect the world around them, includes three functions: instrumental, regulatory, and interpersonal. The first type is performed to get other people do things for us in response to our commands, suggestions, requests, or warnings. The second type is used to control people's behavior as a result of law or regulations implementation. The third type is used "to establish, maintain, and change interpersonal relationships (examples include greetings and leave-takings, complements, insults, and apologies)" (Bachman and Palmer 1996: 70). Concerning heuristic functions, these enable language users to extend their knowledge of the world around them. These functions include the use of language for teaching and learning, or problem solving. We conclude with imaginative functions, which enable us "to use language to create an imaginary world or extend the world around us for humorous or esthetic purposes" (Bachman and Palmer, 1996:70).

4.1.1.3.2 Sociolinguistic Knowledge

This area of language knowledge enables people to create and interpret discourse with reference to the social context. Sociolinguistic knowledge covers the awareness of the conventions, which govern the use of registers, dialects, standard languages, as well as grasping from the linguistic meaning the cultural meaning encoded in it (Lado, 1961; Bachman, 1990; Douglas, 2000).

4.1.2. Background Knowledge

We have mentioned previously that LSP tests have three main features: specificity of content, authenticity of task and interaction between language knowledge and background knowledge. We have also indicated that BK in general language tests forms a source of bias because it leads to scores that are invalidly higher or lower than what they should normally be. Contrariwise, in ESP contexts; BK which refers to test takers' prior knowledge, or familiarity to test content, is considered as a part of the construct to be measured. To illustrate this point, LSP tests designed for electrical engineering specialties are supposed to include some knowledge relevant to engineers in these branches. Consequently, in SPLA testing, we are concerned with measuring the degree to which the test content can involve learners'

linguistic traits and their BK to interact with the test input. In this context, Bachman (1990: 274) argues convincingly that if we design a test for the purpose of measuring learners' SLA, "we are, in effect, defining specialized knowledge as part of the language ability to be tested, and this test should properly be used only with individuals whom we believe to have learned those specific abilities to some degree".

4.1.3. Strategic Competence

Strategic competence (SC) has been revisited several times in the literature of communicative competence. Canale and Swain (1980), for instance, who limit its role to compensating for deficiencies in competence, or in performance, define it as the "verbal and nonverbal communication strategies that may be called into action to compensate for breakdowns in communication due to performance variables or to insufficient competence" (p. 30). Bachman (1990:84) conceives this competence as the mental processes responsible for "implementing the components of language competence in contextualized communicative language use". Since language use involves interactiveness amongst different components such as individuals' LK, BK, and the external context, these mental processes can be seen as a "set of metacognitive components, or strategies, which can be thought of as higher order executive processes that provide a cognitive management function in language use, as well in other cognitive activities (Bachman and Palmer, 1996: 70). In the same way, Douglas (2000) considers it as the processes, which control the interaction between language users' internal traits and the external context. The author identifies two sets of strategies: metacognitive and communication strategies. The first type is engaged when participants perceive that the situation does not require language, such as in the case of performing instructions, or in carrying out a laboratory experiment. The second type is engaged when the situation is perceived to be communicative. According to Douglas (2000: 77), these strategies are "hierarchically arranged so the higher-level metacognitive processes can engage communication at lower level".

4.1.3.1. Phases of Strategic Competence

In LSP learning and testing, strategic competence operates at four levels: assessment, goal setting, planning, and control of execution. At the first phase, test takers assess the features of the specific situation (test input) and attempt to engage an appropriate discourse domain. At the second phase, the examinees determine the communicative purpose that the situation requires them to achieve. At the third phase, they draw a plan that specifies the components of LK and BK to be incorporated for achieving this goal. At the final stage, the plan will be executed by providing responses to the test tasks (Douglas, 2000, 2005, 2010a, 2010b, 2013).

5. Defining the Constructs to be measured

The main purpose of language tests is to generate scores. To provide evidence that these scores are real indicators of test takers' performance on the tests, the latter need to measure the constructs that we intend to measure, and nothing else (Bachman and Palmer 1996). So, before highlighting the relationship between scores and constructs, let us first review some of the definitions provided to constructs in the literature. Cronbach and Meehl (1955: 383), for instance, conceive a construct as "some postulated attribute of people, assumed to be reelected in test performance, in test validation the attribute about which we make statements in interpreting a test is a construct". In the same way, Anastasi (1986: 2-3) thinks of these terms as "theoretical concepts of varying degrees of abstraction and generalizability which facilitate the understanding of empirical data". As far as language testing is concerned, constructs can be defined as the psychological concepts or traits that underlie our linguistic behavior. To be measured, these traits need to be

operationalized (manifested) in language performance in the form of tests (Fulcher, 2010). For example, if we want to measure learners' reading ability, we can administer tests that elicit information about their capacity of skimming or scanning (Alderson, 2000). Similarly, some mental abstractions or constructs of grammatical knowledge can, for instance, be operationalized and inferred by means of tests, which measure active/ passive transformations (Purpura, 2004).

5.1. Approaches to Construct Definition

Three approaches to construct definition have been described in the literature: ability-based, performance-based, and interaction-based approaches (Chappelle, 1998). Ability or trait-based definitions focus on what test takers, or language users have in terms of stored language knowledge. Task or context-based definitions relate the constructs to the observed behavior. In other words, they concentrate on what participants can do with the stored knowledge. However, the interactionalist approach sees constructs as production of the interaction between language users and the external context by means strategic competence (Chappelle, 1998; Weigle, 2002; Purpura, 2004; Bachman, 2007).

5.2. Relationship between Constructs and Topical Knowledge

Bachman and Palmer (1996) identify three options for defining constructs with respect to topical knowledge (Bachman, 2007; Luoma, 2004, Purpura, 2004). The first option limits the construct definition solely to language users' mental traits, excluding topical knowledge. We usually design these types of tests when we want to measure test takers' levels of language competencies. The second option combines between linguistic traits and topical knowledge. In this situation, we assume that test takers have homogenous topical knowledge, which usually occurs in LSP testing. In assessing specialized discourse, Douglas (2005: 860) emphasizes that we need to "interpret test performance in terms of a composite construct of specific purpose language ability that includes both specific purpose language knowledge and field specific content knowledge". However, the third option considers language ability and topical knowledge as separate constructs. We can see this in exams that attempt to measure "theme-based language programs, where [the] topic serves as a context for language learning" (Purpura, 2004: 159).

6. Test Method Characteristics

In their definition of LSP testing, language testers remind us that it refers to the process of making inferences about test takers' specific language competencies and the degree to which they can use these competencies in specific target domains relevant to their fields of study (Bachman and Purpura, 2008; Shohamy, 2008). According to these testers, this process stands upon two constituents: the 'what' and the 'how'. The former delimits the constructs that we plan to test, and the latter describes the methods, which inform us how these constructs can be measured. It also specifies and describes the characteristics of test takers and test tasks to minimize their negative impact on the obtained scores.

6.1. Test Takers' Characteristics

Test takers' characteristics refer to the factors which do not form part of the ability to be measured, but which do have their impact on the obtained scores. These include test takers' personal attributes, their topical knowledge, affective schemata, and different levels of language ability. Age, gender, social stratification, and native language, for instance, constitute the components of personal traits. Topical knowledge refers to test takers' familiarity to the test content. As for the 'affective schemata', these "comprise the user's emotional responses to the situation"(Luoma, 2004: 98). Promoting these variables can improve the participants' performance on the test (McKay, 2006; Fulcher and Davidson, 2012).

6.2. Test Task Characteristics

Language testers distinguish five characteristics for test tasks, which include the rubric, the input, the expected response, the interaction between the input and the expected response, and assessment (Alderson, 2000; Bachman, 1990, 1991; Bachman and Palmer, 1996; Douglas, 2000). Fig 3 illustrates how these components can generate task design. The first aspect explains the participants' progression in completing the test items. The second one specifies the specific material that the examinees are supposed to process. The third feature refers to the type of information that test designers attempt to collect by administering the test. The fourth feature examines the extent of interaction between the input and the expected response. As for assessment, it describes the features of the construct to be measured, and how its tasks will be measured. The identification of these characteristics helps us design interactive items, which maximize test takers' background knowledge to be involved by the input.

Fig 3: The Role of Task Characteristics in Task Design

6.2.1. Characteristics of the Rubric

The test rubric refers to the features, which "provide the structure for particular test tasks and that indicate how test takers are to proceed in accomplishing the tasks" (Bachman and Palmer, 1996: 50). Douglas (2000) identifies five components for the rubric, which include the specification of objectives, procedures for responding, structure of the communicative events, time allotment, and evaluation (how the tasks are to be scored).

6.2.2. Characteristics of the Input

The input in LSP contexts can be conceived as "the specific purpose material in the TLU (target language use) situation that language users process or respond to. In test situation, the input is the means by which, the features of context are established and controlled (Douglas, 2000: 55, [our explanation]). Characteristics of the input describe the prompt (contextual stimuli for eliciting appropriate responses), input data, and the extent of task interactional or situational authenticities. LSP testers signal the importance of specific input data in engaging test takers' language and background knowledge. Based on Hymes (1972), Brown and Yule (1993), Kramsch (1993), and Bachman and Palmer (1996), Douglas proposes a framework for enriching input data specifying eight components: the setting, participants, purpose, form and content, tone, language, norms of interaction, and genre.

6.2.3. Characteristics of the Expected Response

Characteristics of the expected response refer to the type of performance that test developers seek to gather or elicit from test takers (Bachman, 1990; Bachman and Palmer, 1996; Purpura, 2004). In this context, test developers distinguish between actual responses and expected responses. The former refer to the situations where test takers decide to respond in a way that is not anticipated by test designers (Douglas, 2000). Consequently, test developers recommend that we should be clear in delineating the format, the type of responses (selected or constructed), response content (language or background knowledge), and the extent of authenticity (interactional or situational) (Bachman and Palmer, 1996; Douglas, 2000).

6.2.4. Relationship between the Input and Response

This feature describes the extent of interaction between the characteristics of the input and test takers' expected response. As we have implied above, the more the input is field- specific the more it engages test takers' linguistic mental traits and background knowledge to interact with the test input. This interaction is characterized in terms of reactivity (reciprocal/ non-reciprocal), scope, and directness (Bachman and Palmer, 1996). The first feature can be seen as the extent to which test takers' responses can be altered, or adjusted as a result of the feedback provided in the input. Of course, reactivity can be highly reciprocal such as in interviewing when both testers and testees alter questions or responses depending on the feedback they receive from the other interlocutors. However, in non-reciprocal tests such as reading comprehension or composition tasks, the input and responses do not change as a result of the interaction. The second feature of the interaction between input and response concerns its scope. The latter, which varies between narrow or broad continuums, can "pertain to the amount or variety of input that the participant must process before responding" (Douglas, 2000: 65).

6.2.5. Assessment

The final set of characteristics is made up of three components: construct definition, criteria for correctness, and procedures for rating test takers' performance on the test. According to Douglas (2000) ESP constructs result from the interaction between learners' or test takers' language knowledge, background knowledge, and the test specific input by means of strategic competence. For more explanation of the constructs to be measured, see sections (6. 6.1. 6.2.). However, the criteria for correctness specifies "how the correctness of the response is determined: by means of an objective scoring key, multiple value rating scales, judgment of correct/ incorrect, etc." (Bachman and Palmer, 1996: 52). Concerning the procedures taken for scoring the response, the authors see them as "the steps involved in scoring the test: scored in a particular sequence, all rated by the same raters, etc." (52).

7. Recommendations

The literature that we have reviewed concerning the assessment of language ability in specific contexts suggests that this process is scientific, logical, and methodological. In other words, if this procedure is conducted on a haphazard basis, it brings about unexpected results. Therefore, for building valid and interactive specific measurement instruments, this paper proposes a set of recommendations relevant to test designers/users, test takers, and test content. Concerning the first type, we recommend the reinforcement of assessment literacy amongst teachers and measurement officials so that they will be aware of what they test and how they test it. Concerning test takers, a prior identification of their specific communicative needs should be identified, and their levels of language ability and background knowledge need to be determined. As for the test itself, and for validity reasons, this should be designed to measure the constructs that test developers intended to test, and nothing else; otherwise, the resulting scores cannot be interpreted as real indicators of the abilities being measured. Finally yet importantly, and for reinforcing the interactiveness between test takers' strategic competence and test tasks, the latter need to demonstrate similarity to tasks in real contexts, where test scores are supposed to be generalized.

CONCLUSION

In conclusion, this paper has reviewed the literature relevant to the assessment of languages for specific purposes. It has defined this process as the administration of tests in order to collect information about the participants' language competencies and about using these competencies in real target domains, similar to their academic

specialties, or corresponding to their work-place contexts. The paper has also signaled that designing valid and dependable tests requires the description of the specific language abilities to be tested, and the specification of the methods, which tell how these abilities can be measured. Language testers theorize that the former is made up of three components: language knowledge, background knowledge, and strategic competence. Concerning the test method, it specifies two types of characteristics: one type is related to the participants, and the other is relevant to the tasks they are supposed to work on. The description of the participants' characteristics enables us to conduct an extensive needs analysis for the purpose of providing them with the most valid tests and the most appropriate testing environment. This is because the more the input is highly field-specific, the more it engages the participants' competencies to be involved by its content. The paper has concluded with a list of recommendations intended to improve language testing in LSP contexts.

REFERENCES

Alderson, J, C. 2000. Assessing Reading. Cambridge University Press, Cambridge

Alderson, J, C., and Bachman, L. F. Eds. 2000–2006. The Cambridge Language Assessment Series. Cambridge University Press, Cambridge.

Alderson, J. C., Clapham, C., and Wall, D. 1995. Language Testing Construction and Evaluation. Cambridge University Press, Cambridge.

Anastasi, A. 1986. Evolving concepts of test validation. Annual Review of Psychology. 37 (1):15. Accessed on 13-04-2015 from http://www.annualreviews.org/doi/pdf/10.1146/annurev. ps.37. 020186.000245

Anastasi, A., and Urbina, S. 1997. Psychological Testing (7thed). Prentice Hall, Upper Saddle River, NJ.

Bachman, L. F. 1990. Fundamental Considerations in Language Testing. Oxford University Press, Oxford.

Bachman, L. F. 1991. What does language testing have to offer? TESOL Quarterly. 25 (4): 671-704. Accessed on 05-12-2015 from http://bit.ly/2yWt059.

Bachman, L. F. 2007. What is the construct? The dialectic of abilities and contexts in defining constructs in language assessment. pp.41-71. *In*: Fox, J. Wesche, M., Bayliss, D., Cheng, L., Turner, C. E., and Doe, C. (Eds.). Language Testing Reconsidered. University of Ottawa Press, Ottawa, Canada.

Bachman, L. F., and Palmer, A. S. 1996. Language Testing in Practice. Oxford University Press, Oxford.

Bachman, L. F., and Purpura, J. E. 2008. Language assessments: gatekeepers or door openers? pp.456-468. *In*: Spolsky B., and Francis, M. H. (Eds.). The Handbook of Educational Linguistics. Blackwell Publishing Ltd, Oxford, UK.

Basturkmen, H. 2006. Ideas and Options in English for Specific Purposes. Erlbaum Associates, Mahwah, NJ.

Basturkmen, H. 2010. Developing Courses in English for Specific Purposes. Palgrave Macmillan, Basingstoke, UK.

Basturkmen, H., and Elder, C. 2004.The practice of LSP. pp. 672–694 *In*: Davies A., and Elder, C. (Eds.). The Handbook of Applied Linguistics.Blackwell, Oxford.

Bloor, T., and Bloor, M. 2004. The Functional Analysis of English: A Hallidayan Approach. (2nded). Arnold, London.

Brown, G., and Yule, G. 1983.Discourse Analysis. Cambridge University Press, Cambridge.

Canale, M., and Swain, M. 1980. Theoretical bases of communicative approaches to second language teaching and testing. Applied Linguistics. 1(1): 1–47. Accessed on 07-03-2016 from http://applij.oxfordjournals.org/content/I/1/1.full.pdf+html.

Chapelle, C. A.1998. Construct definition and validity inquiry in SLA research. pp. 32–70. *In*: Bachman L. F., and Cohen, A. D. (Eds.) Second Language Acquisition and Language Testing Interfaces. Cambridge University Press, Cambridge.

Cronbach, L. J., and Meehl, P. E. 1955. Construct validity in psychological tests. Psychological Bulletin. (52): 281-302. Accessed on 02-01-2014 from <http://bit.ly/2kSSSsP>.

Douglas, D. 2000. Assessing Languages for Specific Purposes. Cambridge University Press, Cambridge.

Douglas, D. 2001. Three problems in testing language for specific purposes: Authenticity, specificity and inseparability.pp. 45-51. *In:* Elder, C, Brown, A., Grove, E., Hill, K., Iwashits, N., Lumley, T., McNamara, T., and O'Loughlin, K. (Eds.) Experimenting with Uncertainty: Essays in Honour of Alan Davies. Cambridge University Press, Cambridge.

Douglas, D. 2005. Testing language for specific purposes. pp.857-868. *In*: Hinkel, H. (Ed.). The Handbook of Research in Second Language Teaching and Learning. Lawrence Erlbaum Associates Publishers, Mahwah, NJ.

Douglas, D. 2010a. Understanding Language Testing. Lawrence Erlbaum Associates Publishers, London, UK:

Douglas, D. 2010b. This won't hurt a bit: Assessing English for nursing. Taiwan International ESP Journal. 2(2):1-16. Accessed on 08-02-2015 from http://tiespj.tespa.org.tw/this-wont-hurt-a-bit-assessing-english-for-nursing.

Douglas, D. 2013. ESP and assessment. pp. 367-383. *In*: Paltridge, B., and Starfield, S. (Eds.). The Handbook of English for Specific Purposes. John Wiley and Sons, Inc, Chichester, UK.

Dudley-Evans, T., and St. John, M.1998. Developments in English for Specific Purposes. Cambridge University Press, Cambridge, UK.

Fulcher, G. 2010. Practical Language Testing. Hodder Education, London, UK.

Fulcher, G., and Davidson, F. 2007. Language Testing and Assessment: An Advanced Resource Book. Routledge, London.

Fulcher, G., and Davidson, F. 2009. Test architecture, test retrofit. Language Testing. 26 (1): 123–144 Accessed on 13-04-2015from <http://bit.ly/2gcpZTh>.

Fulcher, G., and Davidson, F. (Eds). 2012. The Routledge Handbook of Language Testing. Routledge, Abingdon, Oxon.

Halliday, M. A. K. 1973. Explorations in the Functions of Language. Edward Arnold, London.

Halliday, M. A. K. 2002. On Grammar. Continuum, London.

Halliday, M. A. K. 2004. An Introduction to Functional Grammar (3rd ed). Oxford University Press, Oxford.

Hutchinson, T., and Waters, A. 1987. English for Specific Purposes: A Learning-Centered Approach. Cambridge University Press, Cambridge.

Hyland, K. 2006. English for Academic Purposes: An Advanced Resource Book. Routledge, London.

Hyland, K. 2009. Academic Discourse: English in a Global Context. Continuum International Publishing Group, London, UK.

Hyland, K., and Shaw, P. Eds. 2016. The Routledge Handbook of English for Academic Purposes. Routledge, Abingdon, Oxen.

Hymes, D. 1972. On communicative competence. pp. 269–293. *In:* Pride, J., and Holmes, J. (Eds.) Sociolinguistics. Penguin, Harmandsworth, UK.

Kramsch, C. 1986. From language proficiency to interactional competence.The Modern Language Journal. 70(4): 366-372. Accessed on 17-08-2015 from <http://bit.ly/2yqPgDe>.

Lado, R. 1961. Language Testing: The Construction and Use of Foreign Language Tests: A Teacher's Book. McGraw Hill, New York.

Luoma, S. 2004. Assessing Speaking. Cambridge University Press, Cambridge.

McKay, P. 2006. Assessing Young Language Learners. Cambridge University Press Cambridge, UK.

Messick, S.1995. Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. American Psychologist. 50 (9): 741-749.

Mumby, J. 1978. Communicative Syllabus Design. Cambridge University Press, Cambridge, UK.

Naoua, M. 2016a. An evaluation of English language testing in the baccalaureate examination. PhD Thesis. University of Constantine, Algeria.

Naoua, M. 2016b. Specialized discourse in language tests. pp. 1-12. Colloque International les Discours Spécialisés: Enjeux, Descriptions et Pratiques. October19-21, 2016, Gabes, Tunisia.

Purpura, J. A. 2004. Assessing Grammar. Cambridge University Press, Cambridge

Savignon, S, J. 1972. Communicative Competence: An Experiment in Foreign Language Teaching. Center for Curriculum Development. Philadelphia, PA.

Savignon, S, J. 2002. Interpreting Communicative Language Teaching: Contexts and Concerns in Teacher Education. Yale University Press, Yale.

Shohamy, E. 2008. Introduction to volume 7: Language testing and assessment. pp. xiii-xxii *In*: Shohamy, E., and Hornberger, N. H. (Eds.) Encyclopedia of Language and Education: Volume 7.Language testing and assessment (2nded). Springer Science and Business Media, New York.

Weigle, S, C. 2000. Assessing Writing. Cambridge University Press, Cambridge.

Widdowson, H. G. 1973. An applied linguistic approach to discourse analysis. Ph.D. Thesis. University of Edinburgh, Edinburgh.

Widdowson, H. G. 1979. Explorations in Applied Linguistics 1. Oxford University Press, Oxford.

Widdowson, H. G. 1983. Learning Purpose and Language Use. Oxford University Press, Oxford.

Widdowson, H. G. 1984. Explorations in Applied linguistics 2. Oxford University Press, Oxford.

Widdowson, H. G. 2003. Defining Issues in English Language Teaching. Oxford University, Oxford.

نموذج نظري لقياس المقدرة اللغوية المتخصصة

محمد نوة

قسم الآداب واللغة الإنجليزية، كلية الآداب واللغات
جامعة الشهيد حمه لخضر، الوادي، الجزائر

الملخص

يعرّف تقييم المقدرة اللغوية المتخصصة بتلك العملية التي تهدف إلى توزيع اختبارات على نوع خاص من الممتحنين لجمع بيانات حول مستوى كفاياتهم اللغوية، وكذا قياس مدى تمكنهم من استعمال تلك الكفايات في سياقات مماثلة لمجالات تخصصاتهم الأكاديمية والمهنية. تبرز هذه الورقة أهمية استعمال نموذج دوغلاس كمرجع نظري لتعريف وحصر المكونات اللغوية المتخصصة المستهدفة بعملية التقييم، ومن ثم تقديم الطرق المناسبة لكيفية قياسها. تختتم هذه الورقة بمجموعة من التوصيات تخصّ صانعي الاختبارات والممتحنين وكذا محتويات تلك الاختبارات؛ إذ توصي الدراسة بنشر ثقافة التقييم والقياس في أوساط المدرّسين، والحرص على أن تكون مواضيع الاختبارات متجانسة مع الاحتياجات التواصلية المتخصصة للممتحنين، وبناء اختبارات ذات صبغة تفاعلية عالية محفزة لكفايات الممتحنين اللغوية والمعرفية للتجاوب السريع مع نص الاختبار وأسئلته.

الكلمات المفتاحية: الاختبارات، التقييم، القياس، الكفاية التواصلية، المقدرة اللغوية