

CHAPTER ONE

INTRODUCTION

We design language tests so that we can make inferences about a learner's language ability and decisions about the individuals and their performance and language use in specific settings. In order to be able to make these inferences and decisions we must be able to demonstrate a couple of things and take a few influences and characteristics into consideration.

We must be able to identify what language ability is. We have to identify all the factors and characteristics that influence language use and the learner. We consider individual characteristics, personal characteristics, topical knowledge, and affective schemata. We believe that these characteristics can have important influences on both language use and test performance. When tests are designed, it is possible and desirable to design language tests so that these characteristics facilitate rather than impede test takers' performance.

The definition of the construct of communicative language ability in language testing is clearly important because it is only by reference to underlying constructs that test performance can be interpreted successfully. Theoretical models play a crucial role in defining constructs. Douglas (1995: 169) uses the following example, if performance in an interview is interpreted as evidence of a learner's ability to speak in a second language, it is only by sufficiently defining the construct "speaking ability" that the interpretation has any meaning.

According to Douglas (1995: 69), work on construct definition in the 1990s has focused, appropriately, on trait definition (Henning and Cascallar 1992), the nature of context (Douglas and Selinker 1992; 1993), and the role of strategic competence (Chapelle and Douglas 1993, Cohen 1994b, Cohen and Olshtain 1993). In summary, these studies point out the complex nature of the interaction between aspects of external context and language ability. For example, background knowledge, an important learner-internal factor in invoking discourse domains, is relevant to second language performance only when the external context is sufficiently rich, and when the learner's level of acquisition is sufficiently high so as to be able to respond to the

context. In addition, language testers need to be aware of the language processing demands, which are made by test tasks.

Language testing is changing rapidly in South Africa. In schools, performance assessment, where learners have to demonstrate practical command of skills acquired, is rapidly replacing more traditional test formats such as pencil-and-paper tests involving multiple-choice questions. The driving force behind these changes has been government policy, which increasingly is required performance-based assessment in all areas of education, language education being one of them. In adult and vocational education, a new drive for workforce flexibility of credentials, and a concern for accountability in educational expenditure, has led in many societies, to pressure for demonstrable outcomes of learning in terms of concrete, practical and relevant skills. Again, this move has had major implications for language assessment, as language plays a crucial role in the workplace.

We look at two main approaches to second language performance assessment: (1) The work sample approach, which has its origins in general and vocational education and in personnel selection, and has influenced both general purpose and specific purpose assessment in second languages, and (2) a more cognitive and distinctively linguistic approach, in which attention is focused less on the task, which may be relatively unrealistic in real-world terms, but on the qualities of execution in the performance, and/or the evidence it provides about the candidate's control of the underlying linguistic system. We look at general issues in performance assessment – how it is to be identified, what should it look like etc.

Bachman (1990: 18) states that in developing language tests, we must take into account considerations and follow procedures that are characteristic of tests and measurement in the social sciences in general. Likewise, our interpretation and use of the results of language tests are subject to the same general limitations that characterise measurement in the social sciences. The process of measurement is described as a set of steps, which, if followed in test development, will provide the basis for both reliable test scores and valid test use.

According to Davies (1990: 5), taking account of uncertainty within confidence limits, accuracy in measurement is based on estimates of reliability and validity. Reliability

is measured directly in various ways, which can be generalized to a comparison between one set of items and a comparable set in order to estimate consistency of measure. Reliability is also estimated by the amount of interminacy in a test score in terms of the Standard Error of Measurement, on the logical assumption that the larger the spread of scores around a mean the more likely that they are replicable: the negative aspect of this would be for everyone to achieve the same score, a result which in terms of reliability carries no information.

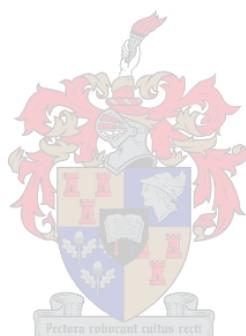
While reliability is concerned with ensuring that a test is a measure (i.e. that whatever it is that it measures it does so accurately), validity has to do with that 'whatever it is' and attempts to provide a theoretical framework, which gives reassurance to the test. It is not surprising that there are several kinds of validity.

Test measurement is also a practical matter. It is important to remember (and to recall) that testing is possible only if it is practicable. A good test (highly valid and reliable) may in practice be unusable because in the situation for which it is intended it would take up too much time, too much skilled manpower, or it might require expensive or elaborate media systems or scoring arrangements, and so on.

McNamara (1996: 193) argues that because test scores are commonly used to assist in making decisions about individuals, the methods used to arrive at these scores are a crucial part of the measurement process. This process, which plays a key role in insuring that the test scores are reliable and that the uses made of them are valid, consists of three steps. The first step is defining the construct theoretically. The second step is defining the construct operationally. The last step of the measurement process, establishing a method for quantifying responses to test tasks, has to do with how scores are derived from test takers' responses to test tasks, and constitutes an essential component of the operational definition of the construct.

We know by now that real-world tasks play a very important role in performance assessment. We look in detail at tasks and different aspects concerned with a task-based approach. We discuss the importance of a needs analysis in task selection, difficulty and sequencing, assessing performance and the steps involved. We discuss the factors affecting reliability, validity and practicality in task-based assessment.

We look at notes on second language learning and teaching which has a huge influence on the design of materials. Following all the theoretical and practical discussions, an example is created – Specific Purpose Sesotho for personnel in Small Business Corporation. From these tasks, several tests are designed.



CHAPTER 2

SECOND LANGUAGE PERFORMANCE ASSESSMENT

2.1 DESCRIBING LANGUAGE ABILITY: LANGUAGE USE IN LANGUAGE TESTS

2.1.1 Language use

According to Bachman and Palmer (1996: 61) 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. In using language to express, interpret, or negotiate intended meanings, language users create discourse. This discourse derives meaning not only from utterances or texts themselves, but, more importantly, from the ways in which utterances and texts relate to the characteristics of a particular language use situation.

Hymes (1972b) describes language use as follows:

The performance of a person is not identical with a behavioural record... It takes into account the interaction between competence (knowledge, ability for use) and the competence of others and the *cybernetic and emergent properties of events themselves*. (Emphasis added) (Hymes 1972b: 283)

Similarly, Savignon (1983) characterises communication as:

Dynamic rather than... static... It depends on the negotiation of meaning between two or more persons... [It] is context specific. Communication takes place in an infinite variety of situations, and success in a particular role depends on one's understanding of the context and on prior experience of a similar kind. (Savignon 1983: 8-9)

Kramsch's (1986) discussion of communicative interaction echoes these notions:

Interaction always entails negotiating intended meanings, i.e., adjusting one's speech to the effect one intends to have on the listener. It entails anticipating the listener's response and possible misunderstandings, clarifying one's own and the other's intentions and arriving at the closest possible match between intended, perceived, and anticipated meanings. (Kramsch 1986: 367)

Bachman and Palmer (1996: 62) states that language use involves complex and multiple interactions among the various individual characteristics of language users, on the one hand, and between these characteristics of the language use or testing situation, on the other. Because of the complexity of these interactions, we believe that language ability must be considered within an interactional framework of language use. The view of language use we present here thus focuses on the interactions among areas of language ability (language knowledge and strategic competence, or metacognitive strategic competence, or metacognitive strategies), topical knowledge, and affective schemata, on the one hand, and how these interact with the characteristics of the language use situation, or test task, on the other.

Candlin (1986) has described communicative competence:

The ability to create meanings by exploring the potential inherent in any language for continual modification in response to change, negotiating the value of convention, rather than conforming to establish principle. In sum... a coming together of organised knowledge to solve new problems of communication that do not have ready-made and tailored solutions. (Candlin 1986: 40)

Bachman and Palmer (1996: 62) observe that we need to keep in mind the involvement of topical knowledge, language knowledge, affect, and strategic competence in carrying out the test task. In describing the construct to be measured, we can choose to include one or more of the following components in the construct definition: language knowledge, possibly topical knowledge, and possibly strategic competence.

2.1.2 Characteristics of individuals

Bachman and Palmer (1996: 64) suggests that language use is affected by a large number of individual characteristics, many of which, such as fatigue or unexpected mood shifts, are largely unpredictable, and there is little we can do to accommodate these in designing language tests. However, there are four sets of individual characteristics whose effects on language test performance are better understood, and that we need to consider in the way we design, develop, and use language tests. These individual characteristics include the following:

- 1) Personal characteristics, such as age, sex, and native language,
- 2) The topical knowledge that test takers bring to the language testing situation,
- 3) Their affective schemata, and
- 4) Their language ability.

2.1.3 Personal characteristics

Bachman and Palmer (1996: 64) proposes that personal characteristics are individual attributes that are not part of test takers' language ability but which may still influence their performance on language tests. In any test development project, the developer will need to develop a specific list of personal characteristics that have to be considered in terms of their potential contribution to the usefulness of the test.

Affective domain:

Affect refers to emotion or feeling. The affective domain is the emotional side of human behaviour. Benjamin Bloom explains the following – first, the development of affectivity begins with receiving. Second, persons must go beyond receiving to responding, committing themselves. Third, affectivity involves valuing, placing worth on a thing, behaviour, or a person. Fourth, is the organization of values into a system of beliefs. Finally, individuals become characterised by and understand themselves in terms of their value system.

The fundamental notions of receiving, responding and valuing are universal. In second language acquisition, learners need to be receptive both to those with whom they are communicating and to the language itself, responsive to persons and to the

context of communication, and to place a certain value on the communicative act of interpersonal exchange. Language is so pervasive a phenomenon in our humanity that it cannot be separated from the larger whole – from the whole persons live and breathe and think and feel.

Self-esteem:

Self-esteem shows self-confidence, knowledge of myself, belief in my capabilities, concepts of myself, acceptance of myself, approval or disapproval of myself, believes in myself to be capable, significant, successful, and worthy. Self-esteem is a personal judgement. The reflection of self is seen in the interaction between self and others. Other people's interaction with me is the mirror I see myself in.

Self-esteem has been categorised into three levels, global, situational or specific, and task self-esteem. Specific self-esteem might refer to second language acquisition in general and task self-esteem might appropriately refer to one's self-evaluation of a particular aspect of the process: speaking, writing, reading, etc. Heyde found that all three levels of self-esteem correlated positively with performance on the oral production measure, with the highest correlation occurring between task self-esteem and performance on oral production measures. Self-esteem appears to be an important variable in second language acquisition. Does high self-esteem cause language success or does language success cause high self-esteem?

Inhibition:

All human beings build sets of defences to protect the ego. Persons with higher self-esteem have lower defences. People with weaker self-esteem have higher defences. Inhibitions protect a fragile ego. Meaningful language acquisition involves some degree of identity conflict as language learners take on a new identity with their newly acquired competence. An adaptive language ego enables learners to lower the inhibitions that may impede success.

The inhibitions, which we place between others and ourselves, are important factors contributing to second language success. Language teaching have been characterised by the creation of context for meaningful classroom communication

such that the interpersonal ego barriers are lowered to pave the way for free, unfettered communication. Anyone learning a foreign language is aware that second language learning actually necessitates the making of mistakes. Adults learning a second language can really only make progress by learning from making mistakes. However, mistakes can be viewed as threats to one's ego. These defences (inhibitions) inhibit learning and their removal can promote language learning.

Risk-taking:

One of the prominent characteristics of good language learners is the ability to make intelligent guesses. Impulsivity could have positive effects on language success. Many factors suggest that risk-taking is an important characteristic of successful learning of a second language. Learners have to be able to 'gamble' a bit, to be willing to try out hunches about the language and take the risk of being wrong.

Success lies in an optimum point where calculated guesses are ventured. The good language learner makes willing and accurate guesses. The silent student in the classroom is one who is unwilling to appear foolish when mistakes are made. Self-esteem seems to be closely connected to a risk-taking factor. Encourage students to guess somewhat more willingly than the usual student is prone to do, and value them as persons for those risks that they take.

Anxiety:

Anxiety is associated with feelings of uneasiness, frustration, self-doubt, and apprehension, to worry. Scovel defined anxiety as "a state of apprehension, a vague fear..." Anxiety can be experienced at various levels. Trait anxiety is more permanent predisposition to be anxious. State anxiety is experienced in relation to some particular event or act.

Foreign language anxiety can be distinguished from other types of anxiety and it can have a negative effect on the language learning process (debilitative anxiety). Facilitative anxiety is some concern – some apprehensions – over a task to be accomplished. Facilitative tension keeps one poised, alert, and just slightly unbalanced to the point that one cannot relax entirely. It is a positive factor. Bailey

shows that facilitative anxiety is one of the keys to success, and closely related to competitiveness. Bailey explained the positive effects of competitiveness by means of the construct of facilitative anxiety. Both too much and too little anxiety may hinder the process of successful second language learning.

Empathy:

Empathy is the process of “putting yourself into someone else’s shoes”, of reaching beyond the self and understanding and feeling what another person is understanding or feeling. Empathy is not synonymous with sympathy. There are two necessary aspects to the development and exercising of empathy: first, an awareness, and knowledge of one’s own feelings, and second identification with another person.

Communication requires a sophisticated degree of empathy. In order to communicate effectively you need to be able to understand the other person’s affective and cognitive states. Oral communication is easier to achieve empathic communication since there is immediate feedback from the hearer. A misunderstood word, phrase, or idea can be questioned by the hearer, and then rephrased by the speaker until a clear message is interpreted. Probably the most interesting implication of the study of empathy is the need to define empathy cross-culturally – to understand how different cultures express empathy.

Extroversion and introversion:

Extroversion is the extent to which a person has a deep-seated need to receive ego enhancement, self-esteem, and a sense of wholeness from other people as opposed to receiving that affirmation within oneself. Extroverts actually need other people in order to feel “good”. Extroverts are not necessarily loudmouthed and talkative. They may be relatively shy but still needs the affirmation of others. Introversion is the extent to which a person derives a sense of wholeness and fulfilment apart from a reflection of this self from other people. Introverts can have an inner strength of character that extroverts do not have.

Stereotypes have influenced teachers’ perceptions of students. Other educators have warned against prejudging students based on perceived extroversion.

Teachers need to consider cultural norms in their assessment of student's presumed "passivity" in the classroom. It is not clear then, that extroversion or introversion helps or hinders the process of second language acquisition. We need to be sensitive to cultural norms, to a student's willingness to speak out in class, and to optimal points between extreme extroversion and introversion that may vary from student to student.

Instrumental, integrative, and assimilative motivation (intrinsic & extrinsic):

Motivation is commonly thought of as an inner drive, impulse, emotion, or desire that moves one to a particular action. Motivation is something that can be global, situational, or task-oriented. Motivation is also typically examined in terms of the intrinsic and extrinsic orientation of the learner. Those who learn for their own self-perceived needs and goals are intrinsically oriented. Those who pursue a goal only to receive an external reward from someone else are extrinsically motivated. Two different clusters of attitudes divided into two basic types of motivation: instrumental and integrative motivation. Instrumental motivation refers to motivation to acquire a language as means for attaining instrumental goals: furthering a career, reading technical material, translation, etc. An integrative motive is employed when learners wish to integrate themselves within the culture of the second language group, to identify themselves with and become a part of that society. Assimilative motivation is the drive to become an indistinguishable member of a speech community, and it usually requires prolonged contact with the second language culture.

Learning a foreign language clearly requires some of all three levels of motivation – global, situational and task-oriented motivation. Motivation in second language acquisition often refer to the distinction between integrative and instrumental orientations of the learner. Integrative motivation may indeed be an important requirement for successful language learning. Second language learning is rarely motivated by attitudes that are exclusively instrumental or integrative. The intrinsic – extrinsic continuum in motivation is applicable to foreign language classrooms around the world. Assimilative motivation is characteristic of persons who perhaps at a very

young age, learn a second language and second culture in order to identify almost exclusively with that second culture.

2.1.4 Topical knowledge

Bachman and Palmer (1996: 64) argue that what they will call topical knowledge (sometimes referred to as knowledge schemata or real-world knowledge) structures in long-term memory. Individuals' topical knowledge needs to be considered in a description of language use because this provides the information base that enables them to use language with reference to the world in which they live, and hence is involved in all language use.

Familiarity of information

Skehan (1998) hypothesises that familiarity of information will lead to greater fluency and accuracy of performance, since the easy access to information should make only limited demands on attention, allowing material to be assembled for speech more easily, and with greater attention to form.

Bygate, Skehan and Swain (2001) states that what they can say, is that familiarity does not have such a strong effect on performance that higher accuracy is guaranteed. In other words, the effect seems weaker than was anticipated. This, however, brings us to the point made immediately above: familiar information does not guarantee more attention being available to achieve a higher level of performance.

2.1.5 Affective schemata

Affective schemata can be thought of as the affective or emotional correlates of topical knowledge Bachman and Palmer (1996: 64) asserts. These affective schemata provide the basis on which language users assess, consciously or unconsciously, the characteristics of the language use task and its setting in terms of past emotional experiences in similar contexts. The affective schemata, in combination with the characteristics of the particular task, determine, to a large extent, the language user's affective responds to the task, and can either facilitate or

limit the flexibility with which he responds in a given context. The affective responses of language users may thus influence whether not only they even attempt to use language in a given situation, but also how flexible they are in adapting their language use to variations in the setting.

A number of language testers have indicated that tests should be designed to elicit test takers' best performance. Bachman & Palmer (1996: 66) suggests that one way to do this is to design the characteristics of the test task to promote feelings of comfort or safety in test takers that will in turn facilitate flexibility of response on their part. However, there needs to be a balance between what the test takers feels comfortable with and what we want to measure. Realizing this, we could try to minimize this threat by building into the interview a warm-up phase, conducted at a level of language with which the test taker feels comfortable, and designed to put the test taker at ease.

2.1.6 Language ability

Bachman (1990: 4) observes that in response to broader views of communicative language ability and communicative language use, much effort is being directed toward developing tests that not only measure a wide range of language abilities, including grammatical, discourse, socio-linguistic, and strategic competencies, but that are also 'authentic', in that they require test takers to interact with and process both the explicit linguistic information and the implicit illocutionary or functional meaning of the test material.

Davies (1990: 52) suggests that there are two main senses of language ability. The first is the more general sense in which we make distinction in terms of knowledge or performance. Even among native speakers we consider some people more fluent than others, better speakers, more effective writers, more efficient readers and listeners and so on. Such distinctions relate to norm-referenced judgements based on which we regard all language ability as normally distributed. The second sense is more often used in relation to second- or foreign-language learners: here we take account of some level of native-speaker ability to which we relate the performance of second language learners. Such an approach may be called criterion referencing.

If we are to make inferences about language ability based on performance on language tests, we need to define this ability in sufficiently precise terms to distinguish it from other individual characteristics that can affect test performance. Bachman and Palmer (1996: 66) maintains. We also need to define language ability in a way that is appropriate for each particular testing situation, that is, for a specific purpose, group of test takers, and TLU domain.

The way Bachman & Palmer defines language ability for a particular testing situation, then, becomes the basis for the kinds of inferences we can make from the test performance. When we define ability this way, for purposes of measurement, we are defining what we called a construct.

Lowe (1988), defines proficiency as follows:

Proficiency equals achievement (ILR functions, content, accuracy) plus functional evidence of internalised strategies for creativity expressed in a single global rating of general language ability expressed over a wide range of functions and topics at any given ILR level. (Emphasis added) (Lowe 1988: 12)

Bachman and Clark (1987) state the advantages of a common metric as follows:

The obvious advantage of such a scale and tests developed from it, is that it would provide a standard for defining and measuring language abilities that would be independent of specific languages, contexts and domains of discourse. Scores from tests based on this scale would thus be comparable across different languages and contexts. (Bachman and Clark 1987:28)

Bachman (1990) defines language ability as involving two components: language competence, or what we call language knowledge, and strategic competence, which we will describe as a set of metacognitive strategies. It is this combination of language knowledge and metacognitive strategies that provide language users with the ability, or capacity, to create and interpret discourse, either in responding to tasks on language tests or in non-test language use.

2.1.6.1 Language knowledge

According to Bachman and Palmer (1996: 67) language knowledge can be thought of as a domain of information in memory that is available for use by the metacognitive strategies in creating and interpreting discourse in language use. Language knowledge includes two broad categories: organisational knowledge and pragmatic knowledge. We recognise that many of the language tests we develop will focus on only one or a few of these areas of language knowledge.

We believe therefore, that the design of every language test, no matter how narrow its focus, should be informed by a broad view of language ability Bachman and Palmer (1996: 67) states.

Organisational knowledge

Bachman and Palmer (1996: 67-8) observes that organisational knowledge is involved in controlling the formal structure of language for producing or comprehending grammatically acceptable utterances or sentences, and for organising these to form texts, both oral and written. There are two areas of organisational knowledge: grammatical knowledge and textual knowledge.

Grammatical knowledge is involved in producing or comprehending formally accurate utterances or sentences. This includes knowledge of vocabulary, syntax, phonology, and graphology.

Textual knowledge is involved in producing or comprehending texts, which are units of language –spoken or written- that consists of two or more utterances or sentences. There are two areas of textual knowledge: knowledge of cohesion and knowledge of rhetorical or conversational organisation.

Knowledge of cohesion is involved in producing or comprehending the explicitly marked relationships among sentences in written texts or among utterances in conversations.

Knowledge of rhetorical or conversational organisation is involved in producing or comprehending organisational development in written texts or in conversations.

Pragmatic knowledge

Pragmatic knowledge enables us to create or interpret discourse by relating utterances or sentences and texts to their meanings, to the language use setting.

Van Dijk (1977) has described pragmatics as follows:

Pragmatics must be assigned an empirical domain consisting of CONVENTIONAL RULES of language and manifestations of these in the production and interpretation of utterances. In particular, it should make an independent contribution to the analysis of the conditions that make utterances ACCEPTABLE *in some situation for speakers of the language*. (Emphasis added) (van Dijk 1977: 189-90)

Bachman and Palmer (1996: 69) suggest that there are two areas of pragmatic knowledge: functional knowledge and socio-linguistic knowledge.

Functional knowledge, or what Bachman (1990) calls "illocutionary competence", enables us to interpret relationships between utterances or sentences and texts and the intentions of language users. Functional knowledge includes knowledge of four categories of language functions: ideational, manipulative, instrumental, and imaginative.

Knowledge of ideational functions enables us to express or interpret meaning in terms of our experience of the real world. These functions include the use of language to express or exchange information about ideas, knowledge, or feelings. Descriptions, classifications, explanations, and expressions of sorrow or anger are examples of utterances that perform ideational functions.

Knowledge of manipulative functions enables us to use language to affect the world around us. This includes knowledge of the following:

- 1) Instrumental functions, which are performed to get other people to do things for us (examples include requests, suggestions, commands, and warnings);
- 2) Regulatory functions, which are used to control what other people do (examples include rules, regulations, and laws); and

- 3) Interpersonal functions, which are used to establish, maintain, and change interpersonal relationships (examples include greetings and leave-takings, compliments, insults, and apologies).

Knowledge of heuristic functions enables us to use language to extend our knowledge of the world around us, such as when we use language for teaching and learning, for problem solving and for the retention of information.

Knowledge of imaginative functions enables us to use language to create an imaginary world or extend the world around us for humorous or aesthetic purposes; examples include jokes and the use of figurative language and poetry.

Socio-linguistic knowledge enables us to create or interpret language that is appropriate to a particular language use setting. This includes knowledge of the conventions that determine the appropriate use of dialects or varieties, registers, natural or idiomatic, expressions, cultural references, and figures of speech.

2.1.6.2 Strategic competence

We conceive of strategic competence 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 a language use, as well as in other cognitive activities Bachman and Palmer (1996: 70) proposes. With respect to language testing, this conceptualisation of strategic competence as metacognitive components provides an essential basis both for designing and developing potentially interactive test tasks and for evaluating the interactiveness of the test tasks we use.

Canale and Swain (1980) describes strategic competence as providing a *compensatory* function when the linguistic competence of the language users is inadequate:

Strategic competence... will be made up of verbal and non-verbal communication strategies that may be called into action to compensate for breakdowns in communication due to performance variables or to insufficient competence. (Canale and Swain 1980: 30)

Canale (1983) has extended this definition of strategic competence to include both the compensatory characteristics of communication strategies and the enhancement characteristic of production strategies:

Strategic competence: mastery of verbal and non-verbal strategies both (a) to compensate for breakdowns in communication due to insufficient competence or to performance limitations and (b) to enhance the rhetorical effect of utterances. (Canale 1983: 339)

We identify three general areas of metacognitive components operate goal setting, assessment, and planning.

A. Goal setting (deciding what one is going to do)

Goal setting involves:

- 1) Identifying the language use tasks or test tasks,
- 2) Choosing, where given a choice, one or more tasks from a set of possible tasks, and
- 3) Deciding whether to attempt to complete the task(s).

Since the purpose of a language test is to elicit a specific sample of language use, tests typically present the test taker with a limited range of tasks, so that the test taker's flexibility in setting goals for performance on test tasks is generally not as great as that enjoyed by language users in non-test language use Bachman & Palmer (1996: 71) observes.

B. Assessment (taking stock of what is needed, what one has to work with, and how well one has done)

Bachman & Palmer (1996: 71) states that assessment provides a means by which the individual relates her topical knowledge and language knowledge to the language use setting and tasks or to the testing situation and tasks. Assessment also takes into consideration the individual's affective responses in the application of assessment strategies.

Corder (1983):

The strategies adopted by speakers, of course, depend upon their interlocutors. What we attempt to communicate and how we set about it are determined by not only our knowledge of the language but also by our current assessment of our interlocutor's linguistic competence and his knowledge of the topic of discourse. (Corder 1983: 15)

Assessing the characteristics of the language use or test task involves identifying the characteristics of the language use task or test task, in order to determine

- 1) The desirability and feasibility of successfully completing the task, and
- 2) What elements of topical knowledge and language knowledge this is likely to require.

Assessing the individuals own topical knowledge and language knowledge involves determining the extent to which relevant topical knowledge and areas of language knowledge are available, which ones might be utilised for successfully completing the task. This aspect of assessment also considers the individual's available affective schemata for coping with the demands of the task.

Assessing the correctness or appropriateness of the response to the test task involves evaluating the individual's response to the task with respect to the perceived criteria for correctness or appropriateness. The relevant criteria pertain to the grammatical, textual, functional, and socio-linguistic characteristics of the response, as well as its topical content Bachman & Palmer (1996: 73) asserts.

C. Planning (deciding how to use what one has)

Bachman & Palmer (1996: 73) maintains planning involves deciding how to utilise language knowledge, topical knowledge, and affective schemata to complete the test task successfully. If the assessment strategies have determined which of these components are available for use, planning involves three aspects:

- 1) Selecting a set of specific elements from topical knowledge and language knowledge (for example, concepts, words, structures, function) that will be used in a plan,
- 2) Formulating one or more plans whose realisation will be a response (interpretation, utterance) to the task, and
- 3) Selecting one plan for implementation as a response to the task.

Formulating a plan may involve an internal prioritisation among the various elements that have been selected, as well as the consideration of how these can be most effectively combined to form a response. The plan thus specifies how the various elements will be combined and ordered when realised as a response. The product of the planning strategy, then, is a plan whose realisation is a response to the task.

Bachman (1990: 102) asserts that this description of the assessment and planning components in communicative language use is similar to Johnson's 1982 characterisation of the processes involved in communication:

There are at least three processes which [a listener] must undertake if he is to fulfil his role as interactant. Firstly, he must "scan" [the speaker's] utterance to extract... its pragmatic information... [which is] that part of the total information conveyed which contributes to the information, which the listener wants to receive... [The listener] approaches the task of listening comprehension prepared to search for certain pieces of information in his interactant's words. Once this information comes, it has to be assessed according to the speaker's aim, and this is the second process which [the listener] must undertake... [The listener] compares, then, what he is told with what he wants to know, identifies any mismatch and then – as a third process – formulates his next utterance. (Johnson 1982: 149)

❖ **Style and strategy:**

Styles, whether related to personality or to cognition characterize the consistent and rather enduring traits, tendencies, or preferences that may differentiate you from another person. Styles are not by any means immutable tendencies.

Strategies on the other hand, are specific methods of approaching a problem or task, modes of operation for achieving a particular end, or planned designs for controlling and manipulating certain information. Strategies are very widely within an individual, while styles are more constant and predictable.

Successful second language learners are usually people who know how to manipulate style (as well as strategy) levels in their day-to-day encounters and cognitive characteristics or tendencies that usually lead to successful acquisition and strive to develop those characteristics.

❖ **Direct (cognitive) strategies and indirect (metacognitive) strategies:**

Direct or cognitive strategies learners apply directly to the language itself. Indirect or metacognitive strategies in which learners manage or control their own learning process. Direct strategies include a number of different ways of remembering more effectively, using all your cognitive processes, and compensating for missing knowledge. Indirect strategies include, organising and evaluating your learning, managing your emotions and learning with others.

2.1.7 Language skills



Language performance is divided in two forms of productive performance, oral and written, and two forms of receptive performance, aural and reading. With all our history of treating the four skills in separate segments of a curriculum, there is nevertheless more recent trend toward skill integration. The class has to model for the students the real-life integration of language skills, get them to perceive the relationship among several skills, and provide the teacher with a great deal of flexibility in creating interesting and motivating lessons.

Communicative interactive framework is a recent trend in second language teaching. Production and reception are quite simply two sides of the same coin. Interactions mean sending and receiving messages. Exceptions are for example, when sending a letter a person is writing the letter, when listening radio or TV a person is only receiving, listening. Written and spoken language often bear a relationship to each other; to ignore that relationship is to ignore the richness of language. The

interrelationship between written and spoken language is an intrinsically motivating reflection of language and culture and society. By attending primarily to what learners can do with language, and only secondarily to the forms of language, we invite any or all of the four skills that are relevant into the classroom arena. Often one skill will reinforce another; we learn to speak, for example, in part by modelling what we hear, and we learn to write by examining what we can read.

Language ability has been considered by language teachers and language testers to consist of four skills: listening, reading, speaking, and writing. Indeed, a model of language proficiency that has been very influential in language testing during the second half of this century describes language ability in terms of the four skills and several components (for example, grammar, vocabulary, and pronunciation).

Bachman and Palmer (1996: 75) argue that these four skills have traditionally been distinguished in terms of channel (audio, visual) and mode (productive, receptive). Thus, listening, and speaking involve the audio channel, and receptive and productive modes, respectively, while reading and writing are in the visual channel, and receptive and productive modes, respectively.

However, is it adequate to distinguish the four skills simply in terms of channel and mode? If it is, then all language use that involves the audio channel and the productive mode could be considered speaking, while any language use in the visual channel and receptive mode would be reading.

Bachman and Palmer (1996: 75) assert that they believe that it takes very little reflection to discover the inadequacies of this approach. First, it would classify widely divergent language use tasks or activities together under a single "skill." Second, this approach fails to take into consideration the fact that language use is not simply a general phenomenon that takes place in a vacuum.

It is this conception of language use as the performance of specific situated language use tasks that provides, we believe, a much more useful means for characterising what have traditionally been called language skills. We would thus not consider language skills to be part of language ability at all, but to be the contextualised realisation of the ability to use language in the performance of specific language use

tasks. We would therefore argue that it is not useful to think in terms of “skills”, but to think in terms of specific activities or tasks in which language is used purposefully.

Thus, rather than attempting to define “speaking” as an abstract skill, we believe it is more useful to identify a specific language use task that involves the activity of speaking, and describe it in terms of its task characteristics and the areas of language ability it engages. Thus, we would argue that the concept that has been called “skill” could be much more usefully seen as a specific combination of language ability and task characteristics Bachman and Palmer (1996: 75) maintains.

2.1.8 Language ability in terms of test use

Davies (1990: 6) observes that language ability must be seen in terms of test use and test purpose. We may distinguish as least five uses: achievement, proficiency, aptitude, diagnosis, and pre-achievement. Two main test purposes may be distinguished, norm-referenced and criterion-referenced.

In achievement (or attainment) the concern is with measuring what has been learnt of what has been taught or what is in the syllabus, textbook, materials, etc. Achievement tests are, in other words, based on a clear and public indication of the instruction that has been given Davies (1990: 6) states.

An achievement test is related directly to classroom lessons, units, or even a total curriculum. Achievement tests are limited to particular material covered in a curriculum within a particular period.

Proficiency, on the other hand, is concerned not with publicity stated instruction but with the relationship between language control and a particular use of language, for example, with whether a testee has adequate language for academic study, for practising as a doctor, for working as a pilot, for driving a car, for being a ski instructor and so on Davies (1990: 7) asserts. If your aim in a test is to tap global competence in a language, then you are, in conventional terminology, testing proficiency. Proficiency tests have traditionally consisted of standardized multiple-choice items on grammar, vocabulary, reading comprehension, aural comprehension, and sometimes of a sample of writing. Such tests often have validity weaknesses:

they may confuse oral proficiency with literacy skills, or they may confuse knowledge about a language with ability to use a language.

Aptitude tests measure language-learning ability. In doing so they attempt to incorporate the dynamic (the process) of language learning that achievement and proficiency tests signally fail to do since they are (by their nature) static tests, testing ability and knowledge at one point in time. Aptitude tests on the other hand seek to indicate capacity for growth Davies (1990: 7) suggests.

A foreign language aptitude test is designed to measure a person's capacity or general ability to learn a foreign language and to be successful in that undertaking. Aptitude tests are considered independent of a particular foreign language, predicting success in the acquisition of any foreign language. An aptitude test biases both student and teacher. They are each led to believe that they will be successful or unsuccessful, depending on the aptitude test score, and a self-fulfilling prophecy occurs. It is better for teachers to be optimistic for students, and in the early stages of a student's process of language learning, to mentor styles and strategies carefully, leading the student toward strategies that will aid in the process of learning and away from those blocking factors that will hinder purpose of the testing context.

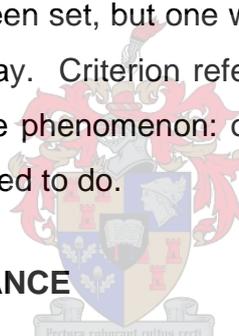
Diagnostic tests do exist in name, but it is not obvious that they differ seriously from proficiency tests. A 'true' diagnostic test seeks to plot learning to date (in certain specified areas) and sets out, through the student profile arrived at, to characterize what is needed in order to remedy error Davies (1990: 7) maintains.

A diagnostic test is designed to diagnose a particular aspect of a particular language. A diagnostic test in pronunciation might have the purpose of determining which particular phonological features of the language pose difficulty for a learner. Some proficiency tests can serve as diagnostic tests by isolating and analysing certain sets of items within the test. Certain proficiency tests and diagnostic tests can act in the role of placement tests whose purpose is to place a student in a particular level or section of a language curriculum or school.

Pre-achievement is in fact a special case of achievement testing, in which a test design is based not on the instruction that has been given, but on the instruction to come after. Many admissions tests are of this kind since they are content related to

the syllabus that will be followed. They operate as screening, level-finding tests and those whose performance is very good will of course be excluded from all following instruction Davies (1990: 7) asserts.

According to Davies (1990: 7) norm-referenced tests are based on the notion of the normal distribution of attributes (such as height) and abilities (such as intelligence, language). It is no doubt the case that the notion may be unreal, or that the tests that do distribute normally do so only because they have been designed to do so, not because they necessarily mirror reality. Norm referencing makes use of rank orders; it places individuals on a scale. Recent discussions of criterion referencing have seemed to suggest that it is a new development in that it replaces ranks by criteria or goals which are task related: the emphasis is now on discriminating tasks rather than individuals. But as with the analogies, high and long jumps, criterion referencing has always been a particular use of norm referencing in that a level of achievement or proficiency (a 'pass') has often been set, but one which did not necessarily select out a set proportion in a mindless way. Criterion referencing and norm referencing are essentially two sides of the same phenomenon: one side looks at what people can do, the other side at what they need to do.



2.2 MODELLING PERFORMANCE

2.2.1 Do we need a model of abilities?

Why, if at all, do we need a discussion of abstract models? In answer to this question, a number of arguments in favour of such models may be advanced.

1. The need for greater clarity

The existing discussion of essential terms in this area is seriously confused according to McNamara (1996: 54). There is a pressing need for definitional clarity, as the construct validity of performance-based language proficiency tests which claim to reflect the notion of communicative competence is by no means clear, despite the often-claimed 'directness' of such tests (Clark, 1975; cf. the discussion in Bachman, 1990, chs 7 and 8). According to Lantolf and Frailly (1988: 186), "A review of the

recent literature on proficiency and communicative competence demonstrates quite clearly that there is nothing even approaching a reasonable and unified theory of proficiency”.

The distinction (if any) between the term's *proficiency* and *communicative competence* is a case in point: it is a matter for strong disagreement. For example, according to Savignon (1983: 246), “Language proficiency is communicative competence and should be defined and evaluated as such” (emphasis in original). Ingram (1981: 124), on the other hand, states that the Australian Second Language Proficiency Ratings (ASLPR) oral proficiency interview and scale are “essentially concerned with language proficiency and not with communicative competence”. Moreover, the ASLPR is derived directly from the American FSI scale, itself described by Savignon (1983: 245) as “one of the first tests to attempt evaluate communicative competence”.

Other instances of statement, claims and definitions, which can only result in confusion for the reader, are not hard to find. For example, according to Richards (1985: 4), “language proficiency is hence not a global factor.” But according to Alderson *et al.* (1987: IV), “Proficiency is a global construct.”

An attempt is needed to resolve some of these apparent contradictions.

2. The problem of generalisation

The practical problem in performance assessment is the generalisation from one observed instance of behaviour to other unobserved instances. To recapitulate briefly the argument presented there: the threat to the validity of such generalisations has been pointed out by numerous writers on performance assessment (Fitzpatrick and Morrison, 1971; Linn *et al.* 1991; Messick, 1994). One approach to the problem has been to adopt a socio-linguistic framework for the specification of the dimensions of the communicative task; that is, to view the question as reducible to one of content validity (cf. Weir, 1988a; Kelly, 1978; Moller, 1982). However, Messick (1989: 17) explicitly rejects this view: “In a fundamental sense so-called content validity does not count as validity at all.”

Even a writer such as Morrow (1979: 150-1), who advocates a severely practical approach and is sceptical of theory, recognises the problem of what he calls “extrapolation” in performance testing. His solution is to espouse the notion of “enabling skills” proposed by Munby (1978). Morrow (1979: 152-3) comments on the function and status of these enabling skills in the following way:

The status of these enabling skill *vis-à-vis* competence: performance is interesting. They may be identified by an analysis of performance in operational terms, and thus they are clearly, ultimately performance-based. However, at the same time, their application extends far beyond any one instance of performance and in this creativity, they reflect an aspect of what is generally understood by competence. In this way, they offer a possible approach to the problem of extrapolation.

This notion of enabling skills has been important in the design of language tests, particularly tests of reading, as well as of instructional materials: for a discussion of this, and a critique of the assumptions behind such an approach, see Lumley (1993).

3. Implicit vs. explicit models

According to McNamara (1996: 52) even practical approaches, which try to eschew theory, imply a theoretical position. This is often to be found in the criteria for assessment, which embody an implicit view of the construct being assessed. Such criteria spell out the relative importance to be attributed to different aspects of the performance; in other words, they contain an implicit theory of performance. Two examples of this is the rich model of communication, including of underlying abilities in performance, implicit in the tasks and criteria for assessment proposed by Savignon (1972); and the approach to criteria in the work of Clark, who refers to “the ... performance criteria on which the test is conceptually based” (1972: 128).

4. The role of non-linguistic factors and the performance of native speakers

We need to understand the role of non-linguistic factors in performance, and their effect on inferences we may make about candidates’ proficiency. J.B. Carroll (1954:

38) formulated the issue as early as 1954, in a discussion of performance on an oral task (talking on a topic for two or three minutes) in a first and in a second language:

If we ask a group of examinees to discourse on a given theme in their *native* language, wide variation in performance will be noted.... There is a good likelihood, I think, that performance in a foreign language will reflect speaking ability in the native language. (Emphasis in original.)

Carroll himself was one of the first to develop some of the implications of Chomsky's competence / performance distinction for work in language assessment, including foreign language testing, and was clearly aware (Carroll, 1968: 50-1) of the issue of the role in second language performance contexts of non-linguistic variables which are not *specific* to such contexts:

The actual manifestation of linguistic competence ... in behaviour may be called linguistic performance, and is affected by a large number of non-linguistic variables.... The single most important problem confronted by the language tester is that he cannot test competence in any direct sense; he can measure it only through manifestations of it in performance. Inevitably, therefore, there is the danger that non-linguistic variables in performance will mask the manifestations of competence. (Emphasis in original.)

If these non-linguistic variables are not specific to the second language performance situation, then we cannot assume that native speakers will perform better than non-native speakers in the tasks on our tests, as native and non-native speakers may not easily be distinguished in terms of the non-linguistic performance capacities that are involved in the tasks.

5. Theory is needed to inform a research agenda

McNamara proposes (1996: 54) that we need a theory of the role of non-language-specific cognitive and affective variables in language performance settings in order to make sense of research on performance testing and to provide a general framework in which to formulate explicit hypotheses about the relationship between candidate and rater behaviour and test scores. Examples are research on such issues as the effect of gender variables in oral interviews (Porter, 1991), the effect of differing

pairings and groupings in task-based oral tests involving more than one candidate at a time (Iwashita, 1993), and raters, differing perceptions of the relevance of non-language skills (cf. Brown, 1995, on the difference between industry-based and education-based raters in the test of Japanese for tour guides, or Elder, 1993b, on ratings of the classroom proficiency of foreign-trained teachers of maths and science by maths-background and ESL-background raters).

2.2.2 What is theory of performance?

The most influential general discussion of language performance in applied linguistics is to be found in Hymes's theory of communicative competence (Hymes, 1967, 1972).

Most relevant in this context is Hymes's distinction between, on the one hand, actual instances of language use in real time (as, for example, in actual test performance) and, on the other, abstract models of the underlying knowledge and capacities involved in language use. The importance of a distinction of this general kind – between actual use and underlying models – is a development of Chomsky's (1965: 4) competence / performance distinction:

We thus make a fundamental distinction between competence (the speaker-hearer's knowledge of his language) and performance (the actual use of language in concrete situations: (Emphasis in original.)

Hymes (1972) early on pointed out that there is an ambiguity in Chomsky's use of the term *performance*. Hymes (1972: 280) distinguishes two uses of the term:

- 1) (underlying) competence v. (actual) performance;
- 2) (Underlying) grammatical competence v. (underlying) models / rules of performance.

Hymes distinguishes between *performance models* – ability as potential – and *actual use*, instances of the realisation of this potential – “actual performance” (Chomsky, 1965: 3). For Hymes, *ability for use* refers (Hornberger, 1989: 226) to “the individual's potential to realise a possible, feasible and appropriate speech act, not to

the realisation itself". Hymes (1989: 247) confirms that this is what he intended; *ability for use* is something "underlying," a "state." The term *performance* is therefore reserved by Hymes for "actual use and actual events" (1972: 283).

Further to the basic distinction between *competence* and *performance*, Chomsky (1980: 59) introduces a distinction between *grammatical competence* and *pragmatic competence*:

I assume that it is possible in principle to have full grammatical competence and no pragmatic competence, hence no ability to use a language appropriately, though its syntax and semantics are intact.

Chomsky's focus of interest was of course in knowledge of language, not the capacities underlying performance. In contrast, *communicative competence* for Hymes encompasses *both* aspects of *knowledge* (now extended to include, in particular, socio-linguistic knowledge) and aspects of *performance*, or what Hymes terms *ability for use*. That is, two areas of abstract modelling are proposed: both a *socio-linguistic* model of language knowledge *and* a broadly *psychological* model of language performance, all subsumed under a single term, *communicative competence*.

The scope of *ability for use* is broad, and includes a range of cognitive and non-cognitive factors, none of them exclusive to language performance. Hymes (1972: 283) includes *motivation*:

The specification of ability for use as part of competence allows for the role of non-cognitive factors, such as motivation, as partly determining competence (emphasis in original)

As well as range of other factors (explicated in detail by Goffman, 1967):

... Capacities in interaction such as courage, gaminess, gallantry, composure, presence of mind, dignity, stage confidence....

Hymes also points out that psycholinguistic performance models are models of aspects of his *ability for use* hence part of *communicative competence*. An example of such a model, addressing the mental processing of language, is that of the

psycholinguists Bialystock and Sharwood-Smith (1985); their work is typical of the strand of research in the area of second language acquisition, which attempts to understand how language processing occurs in the course of language use.

McNamara (1996: 59) observes that the main point is that we need to distinguish between *actual instances* of language use (in real time), what we might call (*actual*) *performance*, and the *potential* for that performance which is available to the speaker: this underlying potential is what Hymes calls *ability for use*. Making this distinction between potential for performance and actual instances of use is helpful when it comes to understanding discussions of communicative language testing, performance-based language tests, the approach to testing known as the Proficiency Movement, and the like.

2.2.3 Influential theoretical models of second language performance

According to Chalhoub-Deville (1995: 7) another componential interpretation of proficiency is Bachman's (1990a) CLA model. Bachman states that his model builds on the work by Hymes (1972; 1973), Munby (1978), Canale and Swain (1980), Canale (1983) and Savignon (1983). He also contends that his model is based on empirical work, primarily the Bachman and Palmer (1982) study. Bachman and Palmer report that the confirmatory factor analysis of a multitrait-multimethod study investigating the grammatical, socio-linguistic and pragmatic competence using a modified ACTFL interview, a writing sample, a multiple-choice test and a self-rating method revealed a general factor and two specific trait factors – grammatical and pragmatic competence. In short, Bachman proposes his CLA based on both theoretical and empirical work.

The CLA model includes three interacting components: language competence, strategic competence, and psycho physiological mechanisms. Language competence consists of two traits, organisational competence, and pragmatic competence. Organisational competence, in turn, includes grammatical and textual competence. Pragmatic competence subsumes illocutionary and socio-linguistic competence. These components are further broken down to provide a more detailed description of the construct (see components of language competence figure 1 Bachman, 1990a: 87). Strategic competence relates knowledge of language and

knowledge of the world to aspects of the context in which communication occurs. Finally, psycho physiological mechanisms include the auditory/visual channel and the receptive/productive mode, which are needed in language use.

As the above description shows, CLA provides an elaborate and comprehensive representation of language proficiency. Such a representation, however, presents a problem for those who need to construct tests. McNamara (1990: 69) contends, “despite advances such as those represented by, for example, the model of communicative language ability proposed by Bachman (1990), it may be difficult to fully apply such models in the actual implementation of tests’. It is important to note here that Bachman (1990a: 82) does not promote his model for undiscerning implementation, but advocates viewing it as “a guide, a pointer... to chart directions for research and development in language testing”. Other researchers regard Bachman’s model in this fashion also.

Skehan (1991:15) explains that because Bachman’s model outlines in detail the field’s current knowledge, it may be viewed as “God’s truth.” Skehan (1991: 15) writes, however, that “God’s truth” models come and go, and while the Bachman model is the best that we have at present, it is inevitable that it will be superseded and weaknesses will be revealed”. Skehan asserts that the value of Bachman’s model is in its weak version, which provides researchers with a useful organising structure within which systematic language testing investigations can be established. Although Bachman’s model can, indeed, be a useful tool for guiding researchers in their assessment development, the published discussion still lacks an explication of the nature of the association between CLA and test construction. Assessments that are typically constructed in a given context will not include all the features depicted in CLA; only those aspects highlighted by the variables operating in that context will be salient. Research is needed to explicate the relationship between theoretical models such as CLA, which depict the nature of language proficiency in general and the language construct as it appears in a given context.

2.2.4 The Proficiency Movement: proficiency versus communicative competence

McNamara suggests (1996: 76) that the best-known feature of the assessment procedure associated with the Proficiency Movement is the Oral Proficiency Interview (OPI) and associated rating scale for determining the level of the candidate's performance; the ACTFL scale derives from the original FSI scale, although there are differences between the two.

What do the Proficiency Movement and its proficiency scales say about knowledge of language? Lowe (1988: 12) offers the following definition of proficiency:

Proficiency equals achievement (functions, content, accuracy) plus functional evidence of internalised strategies for creativity expressed in a single global rating of general language ability over a wide range of functions and topics at any given level.

Bachman and Savignon (1986: 381), suggest that the Proficiency Movement adopt a definition of proficiency “that includes (at least) grammatical, discourse, and socio-linguistic competence”.

Cummins (1979; 1983) interprets proficiency as composed of two components: cognitive / academic language proficiency (CALP); and basic interpersonal and communicative skills (BICS). Cummins holds that CALP pertains to academic situations and that BICS encompasses those situations outside formal learning contexts.

❖ Oral Proficiency testing

One of the toughest challenges of communicative testing has been the construction of practical, reliable, and valid tests of oral production ability. Production, unlike comprehension, takes time, money and ingenuity to measure. The best tests of oral proficiency involve a one-on-one tester / testee relationship, “live” performance, a careful specification of tasks to be accomplished during the test, and a scoring rubric that is truly descriptive of ability. For several decades now, what was formerly the Foreign Service Institute's (FSI) Oral Proficiency Interview (OPI) has been widely

used across dozens of languages around the world. “FSI levels” (0-5) have become standard indicators within the profession of a person’s speaking proficiency in a given foreign language. In a series of structured tasks, the OPI is carefully designed to elicit pronunciation, fluency / integrative ability, socio-linguistic and cultural knowledge, grammar and vocabulary. Performance is judged by the interviewer, through a detailed checklist, to fall between level 0 (the interviewee cannot perform at all in the language) and level 5 (speaking proficiency equivalent to that of an educated native speaker). The OPI has come under harsh criticism in recent years from a large number of language testing specialists. In short, the OPI can only inform us of how learners can deal with an artificial social imposition rather than enabling us to predict how they would be likely to manage authentic linguistic interactions with target-language native speakers. Bachman also points out that the validity of the OPI simply cannot be demonstrated “because it confounds abilities with elicitation procedures in its design, and it provides only a single rating, which has no basis in either theory or research”.

Oral tests require one teacher to administer it to each student separately. Swain has been developing a test battery that includes a paper and pencil multiple-choice format as component of the three-part test; the other two parts measure oral communication skills and written proficiency. Each of these parts is subdivided into grammatical, discourse, and socio-linguistic skills. Of course, for a classroom teacher, and even more so for large test-administrations, this format takes time to administer because of the individualization involved, but time well invested in the interactive process that we profess to be teaching. For many years the FSI, oral interview has been a widely used interactive oral proficiency test. Its current scoring process involves a complex holistic evaluation.

❖ **Characteristics of discrete point and integrative tests as test methods.**

Discrete point tests were constructed on the assumption that language can be broken down into its component parts and those parts adequately tested. Those components are the “four skills” (listening, speaking, reading, writing), the various hierarchical units of language (phonology / graphology, morphology, lexicon, syntax) within each skill, and subcategories within those units. For example, a typical

proficiency test with its sets of multiple-choice questions divided into grammar, vocabulary, reading, and the like, with some items attending to smaller units and others to larger units, can measure these discrete points of language and, by adequate sampling of these units, can achieve validity. The discrete point approach met with some criticism largely from Oller, who argued that language competence is a unified set of interacting abilities that cannot be separated apart and tested adequately.

Two types of test have been held up as prime examples of integrative tests: cloze tests and dictations. A cloze test is a reading passage that has been “multilated” by the deletion of roughly every sixth or seventh word; the testee is required to supply words that fit into those blanks. The dictation test is a relatively short passage that is read by the teacher to students. The students then write down what they hear. Dictation, along with cloze, is said to be a potentially appropriate integrative test. Dictation test results tend to correlate strongly with other tests of proficiency. Reliability of scoring criteria is a problem that is not presented in multiple-choice, exact-word cloze test scoring. Dictations are quite impractical from a scoring standpoint.



2.2.5 A note on the theory of performance in Oller’s pragmatic tests

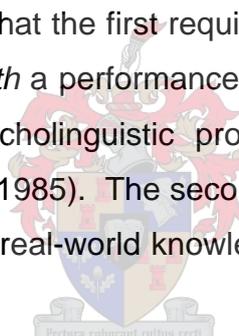
McNamara (1996: 79) states that it is necessary to say a word about the place of John Oller in this discussion of the nature of performance in second language tests. Oller (1979) advocated types of performance test (*pragmatic tests*) which focused not so much on replicating tasks candidates might face in the world, as on tasks which demanded of candidates the same processing of language (understood in psycholinguistic terms), as is required performances on relatively artificial tasks. Such as, filling in gaps in a choice passage or taking dictation were held to be as valid as performance tests involving more “authentic” tasks, which Oller developed an explicit theory of performance in these pragmatic (i.e. communicative) tests. Oller’s “pragmatic expectancy grammar” (Oller, 1979: 16) represents the first serious attempt to discuss the nature of performance in second language testing contexts (thus predating Canale and Swain):

The object of interest is language as it is used for communicative purposes ...
 The notion of *expectancy* is introduced as a key to understanding the nature of psychological real processes that underlie language use.

The “processes that underlie language use” clearly reflect Hymes’s concept of ability for use, defined, as we have seen earlier, as “(underlying) models / rules of performance”. There are two aspects of Oller’s view of performance: Oller refers to them (1979: 33) as “*pragmatic naturalness criteria*” for valid pragmatic tests:

First, they must cause the learner to process ... temporal sequences of elements in the language that conform to normal contextual constraints (linguistic and extra linguistic). Second, they must require the learner to understand the pragmatic interrelationship of linguistic contexts and extra linguistic contexts.

McNamara (1996: 80) proposes that the first requirements are of naturalness of real-time processing; it represents *both* a performance requirements (an *instance of use*) *and* a (relatively inexplicit) psycholinguistic processing model, such as that of Bialystock and Sharwood-Smith (1985). The second requirement is of the integration of linguistic and extralinguistic or real-world knowledge; different aspects of cognitive organisation are involved.



The relevance of this model to the “work sample” tradition of performance testing is suggested by Wesche (1985), who draws on Oller to characterise the nature of performance in such tests. Wesche sees “work sample” tests as a subcategory of pragmatic tests, in that they involve socio-linguistically precise specifications of the “extra linguistic contexts” relevant to the test use context.

❖ Support and refute the Unitary Trait Hypotheses

The unitary trait hypotheses, which suggested that language proficiency, are more unitary than the discrete point testers contended. That is, vocabulary, grammar, phonology, the “four skills”, and other discrete points of language cannot, in fact, be distinguished from each other. The unitary trait hypothesis contended that there is a general factor of language proficiency such, that not all the discrete points add up to that whole. Others argued strongly against the unitary trait position, noting, among

other things, that “language proficiency is one of the most poorly defined concepts in the field of language testing”. Finally, in the face of the evidence, Oller backed down, and admitted, “the unitary trait hypothesis was wrong”. Meanwhile, language-testing research was gradually pointing to the viability of the hypothesis that language proficiency consists of several distinct abilities. Researchers were – and continue to be – focused on the components of communicative competence in their effort to specify the multiple language traits that must be measured in a valid test. Along with the components of organizational (phonology, grammar, discourse) competence, language tests of the 90s are focussing on the pragmatic (socio-linguistic, functional) and strategic components of language ability. While the integrative properties of cloze and dictation testing still appear to hold some promise, most testing specialists now agree that valid tests of communicative competence must tap into all of these communicative components.

2.2.6 Communicative competence in first and second languages

A theme of this discussion has been the difficulty with which Hymes’s rich understanding of communicative competence, developed in a first language context, has been incorporated into thinking about the ability to communicate in a second language. However, why should communicative competence be understood differently for first and second languages? – Not that Hymes himself intended any such distinction.

McNamara proposes (1996: 81) that work settings in which first and second language learners share the same communicative tasks present *prima facie* cases for the appropriateness of a single model of communicative competence for first and second language speakers. The same issue arises in educational settings. The justification for distinct forms of assessment for native and non-native speakers in certain contexts – for example, on exiting from secondary education – is a matter of debate (cf. Davies, 1984 b). Why should the role-specific communicative competence of prospective undergraduates be a matter of concern for second language speakers, but not for first language speakers? (cf. discussion of Hamilton et al. (1993). As workforces and classrooms become increasingly linguistically diverse, the distinction

between first and second communicative ability becomes harder and harder to maintain.

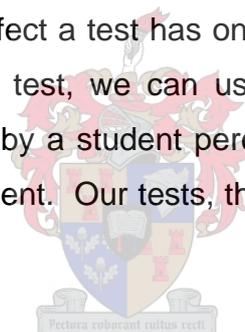
McNamara argues (1996: 84) that the incorporation of affective and other non-linguistic variables in a model of second language communicative ability is recognised by several authors as potentially problematic. We get such a rich picture of the assessment situation that it may be difficult to draw inferences confidently from it, as it contains too many variables. Widdowson (1989: 134-5) has pointed in general terms:

As soon as you talk about competence as ability, or what people can actually do with their language, you get into all kinds of difficulty ... The relationship between knowledge and ability is problematic for empirical work in the description of learner and user language. It is problematic too, of course, in language pedagogy. [Emphasis in original.]

❖ **Communicative test (Bachman & Swain)**

A communicative test has to meet some rather stringent criteria. It has to test for grammatical, discourse, socio-linguistic, and illocutionary competence as well as strategic competence. It has to be pragmatic in that it requires the learner to use language naturally for genuine communication and to relate to thoughts and feelings, in short, to put authentic language to use within a context. It should be direct (as opposed to indirect tests that may lose validity as they lose content validity). In addition, it should test the learner in a variety of language functions. Bachman offers four distinguishing characteristics: First, such tests create an “information gap”, requiring test takers to process complementary information through the use of multiple sources of input. A second characteristic is that of task dependency, with tasks in one section of the test building upon the content of earlier sections, including the test taker’s answers to those sections. Third, communicative tests can be characterised by their integration of test tasks and content within a given domain of discourse. Finally, communicative tests attempt to measure a much broader range of language abilities – including knowledge of cohesion, functions, and socio-linguistic appropriateness – than did earlier tests which tended to focus on the formal aspects of language – grammar, vocabulary, and pronunciation.

Swain recommended four criteria for the construction of communicative tests: (1) Start from somewhere. Tests should build on existing knowledge and principles. (2) Concentrate on content. Swain suggests that our tests need to have motivating, interesting, and substantive content. Tests that concentrate on content also need to be integrated and interactive. Integration provides context, that all-important element of communication. In addition, by eliciting opinions and the expression of one's own ideas, and interactive element can occur. (3) Bias for best. Our tests need to do everything possible to elicit the very best performance from students. Canale suggested that learners perform best if a test has four phases: (a) warm-up, where the testee gets accustomed to the test format; (b) level check, where items confirm that testees are where they are assumed to be; (c) probe, where the limits of one's abilities are ascertained through challenging, difficult items; and (d) wind down, where the learner is allowed to relax with some easier questions that set the mind at ease. These appear to be appropriate principles for "biasing for best". (4) Work for washback. "Washback" is the effect a test has on teaching in the classroom. While we should not teach "toward" a test, we can use tests as teaching tools. Tests become feedback devices whereby a student perceives elements of communicative performance that need improvement. Our tests, then, should be designed to provide that feedback.



2.2.7 The dimensions of an adequate model

McNamara (1996: 85) asserts that we first need a model that is rich enough for us to conceptualise any issue we might think is potentially relevant to understanding second language performance. There should be no limit in principle to the dimension of such a model as possible within the necessary limits of parsimony and clarity. All current discussions, as we have seen, are problematic.

A weakness of current models is that they focus too much in the individual candidate rather than the candidate in interaction. Given the interactional nature of performance assessment, we should be looking more to those in our field who are studying talk interaction.

Secondly, McNamara maintains (1996: 86) that we need a research agenda to investigate the significance for measurement of variables that our models may tell us is likely to be important.

Finally McNamara (1996: 87) states that we need to determine what is appropriate and possible to assess in a given test situation. This will raise both ideological (what position are we taking, and why?) and administrative issues (how feasible or practical is this assessment proposal?). In other words, we are forced into deciding precisely what it is that we are testing, and what it is that we are not. Candlin (1986: 53) makes a useful distinction between models of communicative competence on the one hand, and what he calls the “limits of testability” of such models on the other:

The answer [to the problems discussed in Candlin’s paper] lies elsewhere than in testing. They depend on improving the quality of our description of communicative competence, and then on making educational and administrative judgements on where the boundaries of communicative competence are to be drawn. Once we have made our particular peace between these two forces, we can tackle what must be the second order problem, namely the design and validation procedures we advocate as appropriate.



2.3 PERFORMANCE ASSESSMENT IN LANGUAGE TESTING

2.3.1 What are performance assessments?

Stiggins (1987: 33) gives the following definition for performance assessment:

“Achievement can be, and often is, measured by observation and professional judgement. This form of measurement is called performance assessment”.

Norris, Brown, Hudson, & Yoshioka (1998: 8) observes that what distinguishes performance assessments from other types of tests, then, appears to be that (a) examinees must perform tasks, (b) the tasks should be as authentic as possible, and (c) success or failure in the outcome of the tasks, because they are performances, must usually be rated by qualified judges. These three characteristics might just as

well serve as a working definition for performance assessments – a working definition that will help us to distinguish already existing performance assessments, such as, essays, interviews, extensive reading tasks, and so forth from integrative tests like dictations and cloze tests which do not fully meet any of the three criteria.

According to McNamara (1996: 6) a defining characteristic is that actual performances of relevant tasks are required of candidates, rather than more abstract demonstration of knowledge, often by means of pencil-and-paper tests. Performance assessment within second language contexts involves specific issues, largely because of the complex question of what communicative ability in a second language is understood to be. Performance assessment within second language contexts has meant something more, and more complex, than the application of techniques developed in non-language contexts.

Two traditions of second language performance assessment can thus be found:

- 1) A simple application of the techniques of performance assessment developed in non-language contexts, which we will call the work sample approach; and
- 2) A tradition that sees performance in a second language as a complex cognitive achievement, involving integration of a number of psycholinguistic processes.

In (2) assessment focuses on two issues: (a) the quality of the extension of the performance, rather as in music, ballet or gymnastics; and (b) what the performance reveals about the underlying state of language knowledge in the individual being assessed. This tradition tends to be more theory driven, or to be informed more by theory, although this is not always the case.

Jones (1985), following Slater (1980), discusses three main types of performance tests: direct assessment; work sample methods; and simulation techniques. In direct assessment, observation takes place directly in the workplace. In work sample methods, assessment again takes place in the workplace, but the tasks set are controlled to achieve standardization of assessment. Simulation techniques are distinguished by the fact that the tasks involve a degree of abstraction from workplace reality.

Wesche (1992) differentiated between performance testing in the workplace and in the instructional context. In the work-place context, tests are used for job certification and for prediction of post-training behaviour. In the instructional context, tests are used for washback, diagnostic feedback, and increasing students' motivation. Early introduction of performance tests can help communicate to learners the importance of language objectives, instructor's expectations, and criteria for judging performances. Texts and tasks, which are used in performance testing, also make very good instructional tasks, and ratings obtained from performance tests can be translated to diagnostic feedback in the form of profiles scores. Thus, performance tests can actually be introduced in the pre-instruction phase for placement, formative diagnosis, and achievement purposes; during the program itself, these tests can be used for achievement purposes, for summative testing at the end of a program, and for certification purposes. In instructional situations where the goals are based on an analysis of target language needs, there is a place in the curriculum for an evaluation system, which includes performance-type tasks.

2.3.2 What should performance assessment look like?

Shohamy (1995: 190) maintains that in constructing a performance test, a needs analysis is conducted in order to provide a detailed description of the specific context and tasks which learners will need to perform, the specific conditions under which these tasks will be performed, and the criteria against which the performance can be judge. The learners' performances can be judged over a range of tasks that need to be sampled, using a variety of instruments and procedures. The needs analysis will specify the context of the second language use; the type of interactions foreseen; the roles, discourse types, and language functions to be performed; and the basis on which successful fulfilment of the second language tasks is to be judged. It is with respect to these needs that the performance test is designed, texts and tasks are selected, and evaluation criteria are determined. These are then translated into appropriate test objectives and tasks, and later into actual test design and scoring. Performance tests are generally assessed with the aid of rating scales, which describe what a person can do with the language in specific situations.

Wiggins (1989) advocates widespread use of authentic tests for educational measurement. According to him, such tests should (a) have collaborative elements, (b) be contextualized and complex, (c) measure real-world tasks, and (d) have standards that are authentic and clear to students.

McNamara (1996: 9) argues that typically, the performance process (the nature of the performance task) is considered to be the distinctive feature of performance assessment.

Norris, Brown, Hudson, & Yoshioka (1998: 9) suggests that performance assessments will typically be based on tasks, which will be judged by raters on the basis of some form of rating scale.

1. The tasks should:

- a) Be based on needs analysis (including student input) in terms of rating criteria, content, and contexts
- b) Be as authentic as possible with the goal of measuring real-world activities
- c) Sometimes have collaborative elements that stimulate communicative interactions
- d) Be contextualised and complex
- e) Integrate skills with content
- f) Be appropriate in terms of number, timing, and frequency of assessment
- g) Be generally non-intrusive, that is, be aligned with the daily actions in the language classroom

2. Raters should be appropriate in terms of:

- a) Number of raters
- b) Overall expertise
- c) Familiarity and training in use of the scale

3. The rating scale should be based on appropriate:

- a) Categories of language learning and development
 - b) Appropriate breadth of information regarding learner performance abilities
 - c) Standards that are both authentic and clear to students
4. To enhance the reliability and validity of decisions as well as accountability, performance assessments should be combined with other methods of gathering information (for instance, self-assessments, portfolio's, conferences, classroom behaviours, and so forth)

2.3.3 How are performance assessment developed?

Shohamy (1992) points out that tests should not be used alone to implement change, but should instead be used as part of a dynamic system in which they can provide valuable information for affecting changes in instruction and learning.

Norris, Brown, Hudson, & Yoshioka proposes (1998: 12) that they have developed their own performance assessments at various stages of their careers, they would suggest a more comprehensive list of steps including at least those summarised in the following section.

1. Use needs analysis to define the purposes of the assessment in terms of the
 - a) Types of examinees who will be assessed
 - b) Types of decisions that will be made and their consequences (admissions, individual or group diagnosis, grading grouping, certification)
 - c) People who will make the decisions (not necessarily restricted to program administrators)
 - d) Uses to which the resulting scores/information will be put
 - e) The criteria by which a given performance will be judged (based on input from a variety of stakeholders)
2. Design the performances that will be required in terms of the

- a) Type of performance (e.g., a set of behaviours, or a performance product)
 - b) Knowledge, skill, and/or content focus, and how they will be elicited
 - c) Types of discourse, subject matter, and authenticity of tasks
 - d) Degree of warning (Will students be told in advance about assessments? Will assessment be in front of other students or done individually? Will covert observations be made without telling students until afterwards? Etc.)
 - e) Amount of evidence needed (based on the importance of the decisions, the need for representative performances, and the amount of time available) (How much language will be elicited?)
3. Describe how performance will be rated in terms of the
- a) Type of score needed (holistic or analytic)
 - b) People who will rate the performance (i.e., the student's teacher, other teachers, experts, peer raters, a combination of the above, etc.)
 - c) Training those raters will receive in the rating process (i.e., any rubrics they will be given, other information they will receive, practice ratings they will give, recalibration they will experience during the rating processes, etc.)
 - d) Performance criteria that will be applied (in a scoring rubric if appropriate) including description of the categories of language or behaviour that will be rated and description of what successful performance at each score level means (based on needs analysis)
 - e) Score recording method to be used (checklist, numerical scale, portfolio, etc.)
4. Administer the performance assessments, while insuring that
- a) Directions are clearly given for each performance to each student
 - b) Physical conditions (in terms of space, light, noise, temperature, etc.) are the same for each student's performance

- c) Administration conditions (in terms of timing, equipment, regalia, language stimulus, etc.) are the same for each student's performance
5. Rate the performances by
- a) Following the guidelines described in number 3 above
 - b) Training the raters
 - c) Using multiple raters
 - d) Checking the ratings as they are made for aberrant raters
 - e) Recalibrating the raters on a regular basis if necessary
 - f) Recording the results
 - g) Reporting the results to students and other interested parties in a manner that is thorough, easily understood, and conducive to productive change (e.g., in terms of emphasis on an individual's future studies)
6. Analyse the resulting data making sure to
- a) Estimate rater (and other types) of consistency/reliability
 - b) Consider (from multiple perspectives) the validity of the performance assessment procedures as well as the resulting decisions
 - c) Summarise and interpret the results in view of the types of decision involved
7. Initiate change based on the results, in terms of
- a) Making any appropriate curriculum revisions or policy reforms based on the results
 - b) Keeping in mind that performance assessments are only one source of information and that many other sources of information should be used in making such decisions (see J.D. Brown, 1995a p. 46 for a list of other such sources)

Shohamy (1995: 91) argues that there are a number of questions that need to be addressed in constructing performance tests: How can the evaluation criteria reflect the kinds of judgements and consequences that the performance would entail? What

relative weighting should be given to the different criteria? How can the scoring information be interpreted and presented to give maximum information back to the test users? There are also questions more generally related to the criteria by which the performance should be judged: What is the proportion of 'language' vs. 'domain knowledge' to be assessed? Who should be the judge of the performance – a native speaker, a domain specialist, or a teacher? Although most performance tests do use the native speaker as the top level of the scale (ACTFL 1986, Emmett 1985), this issue has been a topic of debate in the language testing literature for many years (Alderson 1980, Bachman 1990). Hamilton, *et al.* (1993) claim that performance on a test involves factors other than straight second language proficiency, and since these factors are included in the assessment, it is expected that there would be an overlap in the performance of native and non-native speakers. Therefore, the reference to native speaker performance unwarranted.

2.3.4 Why bother with performance assessment?

According to Miller and Legg (1993), performance assessment can compensate for the negative washback effect of standardised testing (that is, the negative impact on teaching practises and curriculum content produced by teaching to standardised tests.)



In language education, Jones (1985) argues that the value of language performance assessment is that it measures students' abilities to respond to real-life language tasks. In other words, unlike other types of tests, performance assessments can be used to approximate the conditions of a real task in a real-life situation. As a result, performance assessments have value in that their scores can be used to predict students' abilities in future, real-world situations, unlike other tests where scores are only very indirect predictors of ability to perform a real-life language task. We suggest that this potential for predicting or generalising to future, real-world language use is one of the key contributions that performance assessment might make as an alternative form of language assessment.

Short (1993) adds that non-standardised, alternative assessment (incorporating open-ended questions, portfolio's authentic assessments, and performance-based

measurements) offers a more accurate picture of student knowledge and ability than traditional (short answer, multiple-choice, and so forth) assessment.

Shohamy (1995: 191) asserts that in the past few years, performance testing has become a common form of assessment in the educational research context. It is associated with any procedure not employing paper-and-pencil multiple choice items, and it includes a variety of assessment alternatives such as open ended responses, constructed responses, problem solving tasks, essays, hands-on science problems, computer simulations of real-world problems, exhibits, and portfolios of students' work (Linn, Baker and Dunbar 1991).

Benefits of using performance assessments suggested by Norris, Brown, Hudson, & Yoshioka (1998: 15)

1. Performance assessments can compensate for the following negative aspects of standardised testing (note that Mehrens, 1992 argued against these):

- a) Negative washback (i.e., any negative effects of testing on teaching, curriculum development, or educational policies)
- b) Lake Wobegon effects (i.e., teaching too closely to the test, or excessive focus on the teaching of abilities needed to score well on standardised tests)
- c) Bias in testing
- d) Irrelevant content
- e) Delimited domain coverage in multiple-choice tests (lack of inclusiveness of test content)
- f) Lack of relationship with curriculum goals and objectives
- g) Multiple-choice tests measure ability to recognise only and cannot measure higher order thinking skills
- h) Multiple-choice tests lack obvious and real-world criteria for selection and scoring

2. Performance assessments can have positive washback effects by:

- a) Providing diagnostic information in functional, or task-based curriculums
 - b) Supplying achievement information in functional, or task-based curriculums
 - c) Documenting critical thought, creativity, and self-reflection
 - d) Encouraging critical thought, creativity, and self-reflection
 - e) Aligning classroom assessment and instructional activities with authentic, real-life activities
 - f) Showing students' strengths and weaknesses in detailed and real-world terms
3. Performance assessments approximate the conditions of real-life tasks so they can:
- a) Measure abilities to respond to real-life language tasks
 - b) Create more accurate assessments of student's knowledge and ability than traditional multiple-choice tests
 - c) Predict students' abilities in future, real-world situations
 - d) Be more valid than traditional tests in terms of predicting students' abilities to use language in the future real-life situations

2.3.5 The nature and extent of variability in performance assessment

McNamara (1996: 121) asks how extensive the variability associated with aspects of the assessment setting, such as rater and task factors are? Take, for example, the assessment of writing. We can identify at least three main sources of variability in the scores obtained when candidates are assessed on writing they have produced. Not all candidates will get the same scores, for a number of reasons. First, the relative ability of the candidates will differ, and unless the test involves a simple task within the competence of all candidates, this variation in ability will be reflected in the scores. Secondly, there may be variability associated with the task: if there is a choice of task, then candidates may gain different scores depending on which task they have chosen. Thirdly, there is variability associated with the raters, so that if a

candidate had had a different rater, he or she might have gained a different score for the same performance.

As far as variability associated with task demands is concerned, Linn *et al.* (1991: 19) reports as follows: 'Generalizability studies of direct writing assessments that manipulate tasks... indicate that the variance component for the sampling of tasks tends to be greater than that for the sampling of raters.'

McNamara (1996: 122) states that in short, variability associated with both raters and tasks is extensive and is a fact of life that must be dealt with if we are to derive stable and fair estimates of how well learners can manage relevant tasks. In order to deal with this issue, however, we need to understand more about the sources of variability. The relevant characteristics of judges are of particular interest: what is it about judges and what they do in the rating situation that is associated with this instability in scoring?

❖ **Nature of variability among judges**

- 1) Two raters may simply differ in their overall leniency.
- 2) Raters may display particular patterns of harshness or leniency in relation to only one group of candidates, not others, or in relation to particular tasks, not others. That is, there may be an interaction involving a rater and some other aspect of the rating situation.
- 3) Raters may differ from each other in the way they interpret the rating scale they are using.
- 4) Finally, and rather more obviously, raters may differ in terms of their consistency (or inconsistency); that is, the extent of the random error associated with their ratings.

❖ **Attempts to deal with differences between judges**

McNamara (1996: 125) observes that of course it would be simpler if judges differed less, and in particular, random error is potentially a serious problem. Attempts are usually made to reduce the variability of judges' behaviour. One of these is rater

training, in which raters are introduced to the assessment criteria and asked to rate a series of carefully selected performances, usually illustrating a range of abilities and characteristic issues arising in the assessment. Ratings are carried out independently, and raters are shown the extent to which they are in line with other raters and thus achieving a common interpretation of the rating criteria.

Recent research (McIntyre, 1993; Weigle, 1994; Shohamy *et al.*, 1992) has examined its effects, and has demonstrated the following:

- 1) Rater training is successful in making raters more self-consistent.
- 2) Rater training can reduce but by no means eliminate the extent of rater variability in terms of overall severity.

Constable and Andrich (1984) raise the issue if perfect agreement is even desirable:

It is usually required to have two or more raters who are trained to agree on independent ratings of the same performance. It is suggests that such a requirement may produce a paradox of attenuation associated with item analysis, in which too high a correlation between items, while enhancing reliability, decreases validity.

McNamara (1996: 127) sums it up that raters display certain characteristics in their participation in the rating process, and these characteristics are a source of potentially considerable variability in the ratings of a performance. Traditional methods have tried to eliminate 'undesirable' rater characteristics through processes of training and accreditation.

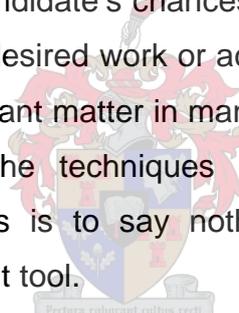
❖ **Other facets of the assessment setting**

McNamara (1996: 127) asserts that we have discussed the effect of rater and task variability on score outcomes. These are but two aspects of the performance setting. For example, in a performance assessment of speaking where the candidate is in direct face-to-face interaction with the assessor, a number of other factors may come into play (it might be the time of day, the physical setting, a gender effect etc.).

McNamara (1996: 128) maintains that each of these aspects of the setting can be called a facet, and information on the effect of any of these facets, or any combination of them, can be sought. Once this information is available, it can then be incorporated automatically into our estimates of the ability of the candidate; in other words, this measure of the candidate's ability results from an automatic adjustment of the candidate's raw score to take account of what is known about the influence of these facets. Exactly what kind of adjustment or allowance should be made, and how it is made, are problems addressed by multi-faceted measurement theory.

❖ **How important is it to model facets of the assessment setting in determining candidates' abilities?**

Briefly, McNamara (1996: 129) argues that typical differences in rater severity following training may affect a candidate's chances of getting a given rating (say, the rating needed for entry to some desired work or academic setting) by as much as 40 percent. This is clearly an important matter in many contexts. Thus, from a practical point of view, knowledge of the techniques and procedures of multi-faceted measurement is worthwhile; this is to say nothing of the outstanding research potential of this new measurement tool.



2.3.6 What problems occur in performance assessment?

According to McNamara, Aschbacher (1991) points out, with respect to large-scale performance assessments, that state governments in the United States are worried primarily about costs, logistics, technical reliability and validity, and support for implementation.

The difficulties and problems in developing, implementing, scoring, and interpreting performance assessments are many. Norris, Brown, Hudson, & Yoshioka (1998: 19) argues that based on their review of the literature and their experiences with performance assessment, they would therefore characterise the problems as summarised in the following section:

1. Considerable increased costs in:

- a) Developing the tests
 - b) Administering the tests
 - c) Training teachers and raters
 - d) Transporting tests (for large-scale assessments)
 - e) Conducting extensive rating sessions
 - f) Reporting scores
 - g) Educating the public education about the procedures
 - h) Many other factors not be immediately evident
2. Increased logistical problems of:
- a) Dealing with large amounts of collected material
 - b) Providing special equipment
 - c) Storing and transporting the tests
 - d) Increasing time commitments
 - e) Providing test security
3. Reliability may be problematic because it depends on:
- a) Rater consistency
 - b) Task-specific variance
 - c) A limited number of observations
 - d) Subjectivity in the scoring process
4. Validity of performance tests is subject to the following threats:
- a) Construct under representation – In language testing terms, this means the problem of generalising from one or a few observations of language behaviour to other real-life instances
 - b) Construct-irrelevant variance (i.e., performance attributes that have little to do with language ability per se)
 - c) Inadequate content coverage
 - d) Lack of construct generalizability



- e) Sensitivity of performance tests to test method, task type, and scoring criteria considerations especially in terms of their validity for measuring higher-order thinking skills
 - f) Differing views on what constitutes evidence for validity (including the perspectives of and interactions between teachers, teacher trainers, psychometricians, students, the general public, and any other stakeholders in the assessment process)
5. Test security may be problematic because of:
- a) The small number of test items which endangers test security and leads to the need for ongoing item development
 - b) The difficulty of creating new items year after year
 - c) Potential affects of “teaching to the test”
6. Political considerations that must be taken into account include the facts that:
- a) It may prove difficult to marshal support for implementation of new alternative procedures when they are compared to traditional, inexpensive, easy-to-administer, norm-referenced tests
 - b) Increased costs will negatively affect public acceptance
 - c) Decisions may have to be defended to multiple stake holders
 - d) Safety, ethics, and legal defensibility of performance tests could become problematic
 - e) The equivalence of the various testing formats and tasks may be difficult to create and defend
 - f) Fairness may be an issue because tasks lack relevancy or because of offensiveness or bias for certain populations
7. Language performance testing, in particular, may prove problematic because:
- a) In trying to create reliable procedures, it may be difficult to:

- i. Create the descriptors of language behaviour which must serve as the criteria for assessment
 - ii. Judge real-time language performance because it moves quickly
 - iii. Deal with differences between level descriptors and actual rater decisions
 - iv. Deal with discrepancies between level descriptors and actual second language performance
 - v. Deal with excessive influence of accuracy on ratings
 - vi. Decide who is an expert and who should rate
- b) In trying to create valid procedures, it may be difficult to:
- i. Account for the role of non-linguistic factors
 - ii. Deal with the fact that native speaker norms cannot be taken to be representative of the highest levels of language performance
 - iii. Implement language performance tests in light of the general lack of necessary theoretical background
 - iv. Determine what constitutes authentic measurement, and for whom? (contextual basis)
 - v. Assess receptive skills because they can only be observed indirectly
 - vi. Decide whether to use live testing or other methods
 - vii. Understand the test situation effect and interviewer effect
 - viii. Use native speakers as models as experts

2.3.7 What steps can be taken to avoid performance assessment problems?

Stiggins (1998: 365) suggests that the “two keys to success of this kind of judgement-based assessment are developing clear, explicit performance criteria and using systematic procedures for rating performances”.

Based on literature review and Norris, Brown, Hudson, & Yoshioka's experience with developing performance assessments, they would conclude that a number of steps can be taken to avoid the problems of performance assessment as summarised in the following section (1998: 24).

1. Considerable increased costs of performance tests

- a) Make sure that any comparisons of the costs of performance testing and the standardised tests to which they will inevitably be compared include *all* of the implementation costs of the standardised tests, including time, staff costs, reusability of items, and so forth.
- b) Use existing resources to the greatest degree possible; for example, do the testing during school time using school facilities, supplies, and teachers

2. Increased logistical problems

- a) Designate a single person responsible for collecting and organising materials and reward that person in some way for doing the task
- b) Write grant proposals for grants to pay for any necessary special equipment
- c) Store and transport the tests using school space and school transportation
- d) Decrease teacher's time commitments to other less essential work related items and tasks (e.g., needless, repetitive paperwork)
- e) Use the same test security precautions that would be used in a standardised test
- f) Use an item bank of performance tasks from which various forms of the test (in various years) can be drawn

3. Reliability can be improved by doing the following:

- a) Create clear, explicitly explained performance criteria
 - i. Produce them early in the test development process
 - ii. Obtain professional advice in criteria

- iii. Examine already existing performance test criteria from the literature
 - iv. Pilot criteria with a special focus on comparing real student work with performance levels found on pilot test
- b) Be systematic in using procedures for rating students' performances
- i. Use clear and detailed scoring criteria in written form
 - ii. Provide explicit criteria for different categories of language if an analytic approach is to be used.
 - iii. Provide explicit criteria for different score levels (within categories of language if appropriate)
 - iv. Give raters example work samples at various levels to help them get a feel for the relationship between the criteria and real work
 - v. Train the raters intensively
 - vi. Retrain the raters as necessary
 - vii. Periodically monitor rater agreement
- c) Use the following general guidelines to improve the reliability of performance assessments:
- i. Provide clear task descriptions and directions
 - ii. Use standard scoring procedures
 - iii. Use multiple tasks (generally, the more, the better)
 - iv. Marshal and utilise multiple sources of information in decision making
4. Threats to the validity of performance assessment can be minimised as follows:
- a) In developing performance assessment, focus either on constructs or on tasks:
- i. Begin construct-based test development by focusing on the construct of interest and then develop tasks based on the performance attributes of the construct, score uses, scoring constraints, and so forth.

- ii. Begin task-centered test development by deciding which performances are the desired ones. Then, score uses, scoring constraints, and so forth.
- b) Carefully study the validity of the performance test, especially with regard to domain representativeness (i.e., the degree to which a task or a few tasks in a testing situation represent the many real-world tasks that will be required later)
 - c) Increase the breadth and depth of coverage in performance tests by using numerous tasks and subtests of varied content
 - d) Receptive-response and productive-response item prompts and task descriptions, as well as various types of formats (for instance, multiple-choice items, checklists, self-ratings, essays, etc.), can effectively be combined to maximize acceptable content coverage and score generalizability
 - e) Complex, interrelated, and integrated skills should be assessed in addition to the discrete component skills (component processes as well as complex, integrated performances are all part of authentic situations, hence they should all be tested)
 - f) Several of the authors suggest novel ways that validity should be viewed if it is to be adequate for assessing performance test and comparing them to other forms of testing
 - i. Significance, that is, scoring criteria should represent a sample of knowledge and strategies from the real-world target domain
 - ii. Fidelity, that is, tasks should involve real-world activities, conditions, purposes, and expectations
 - iii. Generalisability
 - a. Scoring criteria should be representative of specific domains
 - b. Scoring criteria should also include several domains when that is appropriate
 - c. Scoring criteria should represent instructional practices

- d. Raters, teachers, and students should share a common understanding of what the scoring criteria mean
 - e. Scoring criteria should be based on input from multiple perspectives drawn from all stakeholders in the decision-making process
- iv. Developmental appropriateness, that is, scoring criteria should be consistent with theoretical stages of development and should be described in terms of those bands of development (while emphasising accomplishments at each level rather than weaknesses)
 - v. Accessibility, that is, scoring criteria should be written in a clear and accessible writing style so that, at minimum, students, teachers, parents, and even administrators will be able to understand them
 - vi. Utility, that is, scoring criteria should concentrate on aspects of performance that instruction can realistically address within the instructional time available
 - vii. Finally, the consequences, costs, and efficiency of the performance testing must also be considered in thinking about validity
5. Test security can be maximised by:
- a) Using numerous and varied tasks and task prompts
 - b) Minimising teaching-to-the-test by making sure that all parties involved understand the scoring criteria
 - c) Considering that it is *not* possible to compromise the scoring criteria by using rote memorisation strategies
 - d) Maintaining test clarity and meaningfulness (i.e., insuring that the performance tests are meaningful educational experiences that motivate and guide learning)
6. Political problems can be avoided in part by making sure that:
- a) Various testing formats and tasks are equated

- b) Fairness is established and maintained by comparing task relevancy for different populations and by performing periodic reviews for bias and offensiveness
- c) The safety, ethics, and legality of the performance tests are considered prior to implementation
- d) Extensive documentation is maintained throughout the process
- e) Test results are thoroughly studied and documented on a periodic basis
- f) Rating criteria are made explicit from the outset and are in line with classroom activities and real-world objectives

2.3.8 What examples exist of actual performance assessment projects?

Shohamy (1995: 192) observes that over the years, a wide variety of performance tests have been developed in the language-testing field. The early English for Special Purpose (ESP) test (Carroll 1980) included specific performances expected of students in academic situations. Cole and Neufeld (1991) report on a test used by the Canadian Public Service to certify the second language skills of employees in bilingual positions. Modesto, Cahill and Frey (1983) report on the Canadian Public Service's Communicative English at Work instructional program for Francophone government employees in bilingual positions. Sajavaara (1992) describes the Finnish Foreign Language Diploma for Professional Purposes that offers information for employers to use in screening new professionals in terms of foreign language skills. McNamara (1990) reports on a performance test used for health professionals in Australia – the Occupational English Test (OET) – which is used for immigrants and refugees who want to be medical practitioners (nurses, physiotherapists, occupational therapists, dentists, speech pathologists, veterinarians, and surgeons).

Bailey (1985) describes a foreign teaching assistant performance assessment, while Clark and Grognet (1985) discuss a survival skills performance assessment. Other examples include McNamara (1990), which examines an occupational English test for health professionals, and Paltridge (1992), which describes an EAP placement test development project. Fulcher (1996) describes the development of task-based

oral assessments using groups of EFL students in Cyprus according to Norris, Brown, Hudson, & Yoshioka (1998: 30).

2.4 MEASURING PERFORMANCE

2.4.1 Definition of terms: measurement, test, evaluation

Bachman (1990: 18) observes that the terms 'measurement', 'test', and 'evaluation' are often used synonymously; indeed they may, in practice, refer to the same activity. When we ask for an evaluation of an individual's language proficiency, for example, we are frequently given a test score. This attention to the superficial similarities among these terms, however, tends to obscure the distinctive characteristics of each, and I believe that an understanding of the distinctions among the terms is vital to the proper development and use of language tests.

Measurement

Bachman (1990: 18) suggests that measurement in the social sciences is the process of quantifying the characteristics of persons according to explicit procedures and rules. This definition includes three distinguishing features: quantification, characteristics, and explicit rules and procedures.

❖ Quantification

According to Bachman (1990: 18), quantification involves the assigning of numbers, and this distinguishes measures from qualitative descriptions such as verbal accounts or non-verbal, visual representations. Non-numerical categories or rankings such as letter grades ('A, B, C...'), or labels (for example, 'excellent, good, average...') may have the characteristics of measurement scales'. However, when we actually use categories or rankings such as these, we frequently assign numbers to them in order to analyse and interpret them, and technically, it is not until we do this that they constitute measurement.

❖ Characteristics

Bachman (1990: 19) proposes that we can assign numbers to both physical and mental characteristics of persons. Physical attributes such as height and weight can be observed directly. In testing, however, we are usually interested in quantifying mental attributes and abilities, sometimes called traits or constructs, which can only be observed indirectly. These mental attributes include characteristics such as aptitude, intelligence, motivation, field dependence/independence, attitude, native language, fluency in speaking, and achievement in reading comprehension.

The precise definition of 'ability' is a complex undertaking. In a very general sense, 'ability' refers to being able to do something, but the circularity of this general definition provides little help for measurement unless we can clarify what the 'something' is.

John B. Carroll (1983c, 1987a) has proposed defining ability with respect to a particular class of cognitive or mental tasks that an individual is required to perform, and 'mental ability' thus refers to performance on a set of mental tasks (Carroll 1987a: 268). We generally assume that there are degrees of ability and that these are associated with tasks of performances of increasing difficulty or complexity (Carroll 1980, 1987a). Thus, individuals with higher probability of correct performance on tasks of lower difficulty or complexity, and a lower probability of correct performance on tasks of greater difficulty or complexity.

Bachman (1990: 19) argues that whatever attributes or abilities we measure, it is important to understand that it is these attributes or abilities and *not* the response themselves that we are measuring. That is, we are far from being able to claim that a single measure or even a battery of measures can adequately characterise individual human beings in all their complexity.

❖ Rules and procedures

Bachman (1990: 20) asserts that the third distinguishing characteristic of measurement is that quantification must be done according to explicit rules and procedures. That is, the 'blind' or haphazard assignment of numbers to characteristics of individuals cannot be regarded as measurement. In order to be considered a measure, an observation of an attribute must be replicable, for other

observers, in other contexts and with other individuals. Measures then are distinguished from such 'pseudomeasures' by the explicit procedures and rules upon which they are based. There are many different types of measures in the social sciences, including rankings, rating scales, and tests.

Test

Carroll (1986) provides the following definition of a test:

A psychological or educational test is a procedure designed to elicit certain behaviour form, which one can make inferences about certain characteristics of an individual. Carroll (1968: 46)

Bachman (1990: 20) maintains that from this definition, it follows that a test is a measurement instrument designed to elicit a specific sample of an individual's behaviour. As one type of measurement, a test necessarily quantifies characteristics of individuals according to explicit procedures. What distinguishes a test from other types of measurement is that it is designed to obtain a specific sample of behaviour.

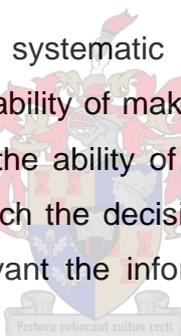
Bachman (1990: 21) states that the inferences and uses we make of language test scores depend upon the sample of language use obtained. Language tests can thus provide the means for more carefully focussing on the specific language abilities that are of interest. As such, they could be viewed as supplemental to other methods of measurement. Given the limitations on measurement, and the potentially large effect of elicitation procedures on test performance, however, language tests can more appropriately be viewed as the best means of assuring that the sample of language obtained is sufficient for the intended measurement proposes, even if we are interested in very general or global abilities. That is, carefully designed elicitation procedures such as those of the ILR oral interview, those for measuring writing ability described by Jacobs *et al.* (1981), or those of multiple-choice tests such as the *Test of English as a Foreign Language* (TOEFL), provide the best assurance that scores from language tests will be reliable, meaningful, and useful. While measurement is frequently based on the naturalistic observation of behaviour over a period, such as in teacher rankings or grades, such naturalistic observations might not include samples of behaviour that manifest specific abilities or attributes. This is not to imply

that other measures are less valuable than tests, but to make the point that the value of tests lies in their capability for eliciting the specific kinds of behaviour that the test user can interpret as evidence of the attributes or abilities which are of interest.

McNamara (1996: 150) observes that the influence of test item characteristics on estimates of candidate abilities has always been recognized; it is a matter of common knowledge. People know that the degree of difficulty of the test is important in understanding the meaning of their achievement. The difficulty of creating equivalent tests has also made the charting of the progress of learners or groups of learners difficult. It would be very helpful if candidate ability measures could be freed from dependence on particular test items, and if the relative difficulty of test items on different tests could be somehow established.

Evaluation

Evaluation can be defined as the systematic gathering of information for making decisions (Weiss 1972). The probability of making the correct decision in any given situation is a function not only of the ability of the decision-maker, but also of the quality of the information upon which the decision is based. Everything else being equal, the more reliable and relevant the information, the better the likelihood of making the correct decision.



Bachman (1990: 22) observes that evaluation, therefore, does not necessarily entail testing. Similarly, tests in and of themselves are not evaluative. Tests are often used for pedagogical purposes, either as a means of motivating students to study, or as a means of reviewing material taught, in which case no evaluative decision is made on the basis of the test results. Test may also be used for purely descriptive purposes. It is only when the results of tests are used as a basis for making a decision that evaluation is involved. Again, this may seem a moot point, but it places the burden for much of the stigma that surrounds testing squarely upon the test user, rather than on the test itself. Since by far the majority of tests are used for the purpose of making decisions about individuals, I believe it is important to distinguish the information-providing function of measurement from the decision making function of evaluation.

2.4.2 Item Response Theory and Generalizability Theory

McNamara (1996: 257) states that Item Response Theory (IRT) is a powerful general measurement theory which was developed in the 1950s and 1960s independently, it seems, in two different locations: by Alan Birnbaum in the United States and by the Danish mathematician George Rasch in Denmark. Rasch's work was promoted and extended by an American, Ben Wright, who attended a series of invitational lectures given by Rasch in Chicago in 1960 and became his pupil and the advocate of his ideas in North America (cf. Wright's account of the history of this relationship in Wright, 1988). Two main branches of Item Response Theory (or Latent Trait Theory as it is sometimes still known), stemming from these two developmental traditions, are recognized. (Thorndike, 1982 gives a summary of the field.) They differ theoretically and practically. The essential feature of both is that they attempt to model statistically patterns in data from performances by candidates on test items, in order to draw conclusions about the underlying difficulty of items and the underlying ability of candidates. They differ mainly in the number of item parameters (characteristics of the interaction between a test taker and a test item) being estimated in the analysis: Rasch analysis considers one item parameter (item difficulty), while other models consider one or more further parameters (item discrimination, and a guessing factor). They also differ theoretically in relation to the question of 'specific objectivity'.

Douglas (1995: 174) proposes that if language test analysis in the 1970s and 1980s could be characterised by reference to factor analytic studies (e.g., Bachman and Palmer 1982, Carroll 1983, Davidson 1988, Oller 1978, Sang, *et al.* 1986), the 1990s might be said to be the decade of Item Response Theory and Generalisability Theory. Item Response Theory (IRT), also known as Latent Trait Theory, has been the vehicle for a number of studies on the relationship between test item difficulty and test taker's ability. In fact, IRT is based on the assumption that "an individual's expected performance on a particular test question, or item, is a function of *both* the level of difficulty of the item and the individual's level of ability" (Bachman 1990a: 202)

Douglas (1995: 174) observes that there are many approaches and applications of IRT to language test analysis, and these have been the subjects of a number of recent studies (Boldt 1992, Choi and Bachman 1992, Hudson 1991, Reynolds,

Perkins and Brutten 1994, Sasaki 1991). A major concern in the use of IRT models for the analysis of language test data has been the so-called 'unidimensionality assumption' of IRT – the assumption that all the items test a single ability. This assumption is problematic in light of the now widely accepted notion that communicative language ability is multidimensional (Bachman 1990a). However, it has been argued (Henning 1992) that the *psychometric* assumption of unidimensionality does not preclude the use of IRT models on tests, which assume the *psychological* multidimensionality of communicative language ability, and the IRT approach has been shown to be "robust" with respect to the unidimensionality assumption (McNamara 1990; 1991). Nevertheless, doubts about the applicability of the model remain (e.g., Buck 1994), and research in the 1990s will no doubt continue to explore the theoretical assumptions of the IRT approach.

Douglas (1996: 174) asserts that other studies have used IRT to collect evidence for test validation (de Jong 1991, Henning 1990; 1992, Lowe, Janczewski and Cascallar 1991, McNamara 1990; 1991, McNamara, *et al.* 1992, Perkins and Brutten 1993, Perkins and Henning 1992). IRT has also been used to produce information allowing for item banking, a useful method for equating test forms by matching item response characteristics (Boutin, *et al.* 1993, ETS 1991, Henning, *et al.* 1994, Jones 1994, Mets 1993). Other IRT-based research includes the following issues: 1) language-test item bias related to subjects' field of specialisation (Henning 1990), 2) the potential of IRT analysis to provide validity evidence for the posited constructs in a communicative test language ability (McNamara 1990), 3) validity issues in a test of listening comprehension (McNamara 1990), and 4) the identification of misfitting items in the use of proficiency tests for placement purposes (Perkins and Brutten 1993). It will be interesting to see whether the use of IRT techniques progresses in these directions in the future, as well as how current reservations about their uses are resolved.

Generalisability Theory (G-Theory)

Bachman (1990: 187) states that a broad model for investigating the relative effects of different sources of variance in test scores has been developed by Cronbach and his colleagues (Cronbach *et al.* 1963; Gleser *et al.* 1965; Cronbach *et al.* 1972). This

model, which they call generalizability theory (G-Theory), is grounded in the framework of factorial design and the analysis of variance. It constitutes a theory and set of procedures for specifying and estimating the relative effects of different factors on observed test scores, and thus provides a means for relating the uses or interpretations to be made of test scores to the way test users specify and interpret different factors as either abilities or sources of error.

In using G-Theory, the researcher carefully specifies various factors that may have an effect on performance, including the trait to be measured (that factor which should have the greatest effect on test performance), the test method factors, and random (or “left over”) effects. By estimating the contribution of each of the factors (and the interaction among them to the test score, the “generalisability coefficient” provides a basis for generalising from the test to performance in other areas. Recent uses of generalisability theory in language test analysis include the use of G-Theory to examine the new English as a Second language Placement Examination (NESLPE) (Kunnan 1992), and the use of G-Theory to compare the Oral Proficiency Interview and the Simulated Oral Proficiency Interview (Stansfield and Kenyon 1992). G-Theory has great potential promise in the analysis of test reliability as part of the validation process, helping pinpoint the sources of error in test performance, which detract from the validity of score interpretations according to Douglas (1995: 175).

Fulcher (1996: 35) suggests that ‘performance based’ refers not only to the test format but also to the scoring methods. This immediately limits the generalizability of any scores to the types of task used in the test (Messick, 1994), as the scoring system adopted for the test would contain descriptors, which were directly related to the test situation itself. This approach thus fuses inextricably together the testing method and what it is we wish to test. It unites trait and test method facet.

Fulcher (1996: 37) maintains that if ability is to be distinguished from behaviour (Bachman, 1990:308) and trait from test method facet (Bachman and Savignon, 1986), as I believe they have to be for the study of construct validity, then whatever ability or abilities one hypothesizes underlie test performance, they must contribute to performance in a variety of situations and tasks. The problem of generalizability from one task to another is thus one, which should be tackled through the development of the scoring system and not necessarily the design of the task.

McNamara (1996: 51) observes that the practical problem in performance assessment is the generalization from one observed instance of behaviour to other unobserved instances. The threat to the validity of such generalizations has been pointed out by numerous writers on performance assessment (Fitzpatrick and Morrison, 1971; Linn *et al.* 1991; Messick, 1994). One approach to the problem has been to adopt a socio-linguistic framework for the specification of the dimensions of the communicative task; that is, to view the question as reducible to one of content validity (cf. Weir, 1988a; Kelly, 1978; Moller, 1982). But Messick (1989: 17) explicitly rejects this view: 'In a fundamental sense so-called content validity does not count as validity at all.'

Even a writer such as Morrow (1979: 150-1), who advocates a severely practical approach and is sceptical of theory, recognizes the problem of what he calls 'extrapolation' in performance testing. His solution is to espouse the notion of 'enabling skills' proposed by Munby (1978). Morrow (1979: 152-3) comments on the function and status of these enabling skills in the following way:

The status of these enabling skills *vis-à-vis* competence: performance is interesting. They may be identified by an analysis of performance in operational terms, and thus they are clearly, ultimately performance-based. However, at the same time, their application extends far beyond any one instance of performance and in this creativity they reflect an aspect of what is generally understood by competence. In this way, they offer a possible approach to the problem of extrapolation.

2.4.3 Rasch model as a research model and multi-faceted Rasch analysis

The advent of Item Response Model (IRM) to the field of language testing (e.g. Henning 1984; Henning, Hudson and Turner 1985; Griffin 1985; Woods and Baker 1985; Pollitt and Hutchinson 1987; Choi and Bachman 1992) has been among the most important developments in the recent history of the discipline. IRM has given language testing a rigorous basis for the measurement dimension. The catch, though, is that IRM has been conceived as a measurement model with little or no immediate implications for language testing research. Specifically, the unidimensionality assumption in IRM has been an initial obstacle for many language

testers. It is argued that, if language were inherently complex, it would be straight jacketing language-testing research by forcing the unidimensional condition onto all language data. (See Bachman 1990 for an interesting discussion.)

Theoretical discussions of IRM within language testing (e.g. Reckase 1979; Henning, Hudson and Turner 1985, Henning 1992; Choi and Bachman 1992) have helped to define the scope of the unidimensionality assumption and to resolve the apparent dilemma. In addition, research designs encompassing an IRM component and development within IRM itself has made it a true research tool.

Rasch model as a research model

Wright and Masters (1982: 2) maintain that the unidimensionality assumption is a 'universal characteristic of all measurement'. This, however, should not in theory preclude analyses over and above an IRM based analysis. Jensen (1978), for example, warns of '...a flagrant conceptual and scientific blunder ... to *apply* orthogonal rotation of principle components are factors without first extracting the general factor (i.e. the first principal component or first principal factor)'. Indeed, IRM can easily be conceptualised as a rigorous way to extract the general factor. The standardised residuals from an IRM analysis would provide data for further analysis as envisaged by Jensen. Lee (1993) analyses the residuals in a study on the construct validity of an ESL reading test.

Multi-faceted Rasch analysis

Multi-faceted Rasch analysis (Linacre 1989a) is the expansion of the one-parameter Rasch model to encompass analyses of facets in the data. This has enabled IRM to be employed in diverse research design and analysis configurations and data collection schedules.

Linacre (1989b) argues and demonstrates the possibility of extending the initial one-parameter (or two-facet) Rasch model to n-facet models. This is an interesting development. Constituents within complex human behavioural context can now be accommodated within the same IRM model for analysis. Lee (1996: 266) states that typically facets can include judges of human performance (e.g. in a writing test), or

sub-groupings of subjects/candidates, or item groups. With the flexibility introduced, research designs can now be developed which would do greater justice to unavoidable features in human behaviour (e.g. varying degrees of severity of judges, cultural and/or economic background of subjects). Within a many-faceted model, all facets entered into an analysis are combined to produce the calibrations and themselves calibrated regarding consistency. In addition, FACETS (Linacre and Wright 1990), which is the software implementation of multifaceted Rasch analysis, can generate interaction analysis of the facets.

According to McNamara (1996: 129) multi-faceted measurement provides us with information on the influence of rater and task characteristics on ability estimates in performance assessment settings. Each rating can be seen as a function of the interaction of three facets or factors: the ability of the candidate; the difficulty of the item; and the characteristics of the rater. The model states that the likelihood of a particular rating on an item from a particular rater for a particular candidate can be predicted mathematically from the ability of the candidate, the difficulty of the item and the severity of the rater.

❖ **Feasibility under operational conditions**

McNamara (1996: 140) suggests that it is worth commenting on the use of multi-faceted analysis under operational conditions. First, the design of the data collection is an important issue, as it is important to plan for some multiple rating each element of each facet under each of the conditions of assessment of interest. Secondly, designs incorporating many facets will need large data sets to ensure stability of estimate of facets. There is a need to limit the number of facets to be analysed in any single analysis, unless very large data sets are available. Thirdly, fresh analysis needs to be conducted for each test administration; we cannot assume that rater characteristics will remain constant across administrations. Fourthly, it is possible to enlarge data sets by combining data from old and new analyses provided the data matrices have some overlap; this overlap is the basis for a procedure of linking across occasions known as anchoring, whereby the values of elements in a facet derived from one analysis are specified for the subsequent analysis, thus allowing comparability of output from the two analyses. Fifthly, sufficient time needs to be

built in between test scoring and the reporting of results to allow the necessary analysis and transformation of raw scores into measures to take place, although with experience this can become an efficient process. In general, the routine use of multi-faceted analysis will be restricted to high-stakes test settings where resources and the necessary expertise are available. One disadvantage of multi-faceted measurement is that analysis requires relatively large data sets; an advantage however, is that economies in the extent of multiple rating are possible because of the flexible data requirements of the analysis. Certainly, use of these procedures under operational test conditions has now become routine for a number of high-stakes language performance tests, and their feasibility has been amply demonstrated.

❖ **Bias analysis**

McNamara (1996: 141) points out that he has considered the way in which multi-faceted measurement can characterize raters in terms of their overall behaviour, across all items and all candidates. Of course, it is possible that raters will display particular patterns of harshness or leniency in relation to only one group of candidates, not others, or in relation to particular tasks, not others. That is, there may be an interaction involving a rater and some other aspect of the rating situation. The identification of these systematic sub-patterns of behaviour is achieved in multi-faceted Rasch measurement in so-called bias analyses.

McNamara (1996: 141) asserts that multi-faceted analysis compares expected and observed values in a set of data. Differences between expected and observed values are known technically as residuals. Fit statistics (for raters, items, persons, and other facets) summarize for each rater, item, person, etc. the extent to which the difference between expected and observed values are within a normal range. The basic idea in bias analysis is to further analyse the residuals to see if any further sub-pattern emerges. The study of bias is thus the study of interaction effects, e.g. systematic interaction between particular raters and particular candidates, or between particular raters and particular tasks / item etc. There will be some unexplained random error left over at the end of this analysis; this represents unexplained random variation.

2.4.4 Properties of measurement scales

Bachman (1990: 26) argues that if we want to measure an attribute or ability of an individual, we need to determine what set of numbers will provide the best measurement. When we measure the loudness of someone's voice, for example, we use decibels, but when we measure temperature, we use degrees Centigrade or Fahrenheit. The sets of numbers used for measurements must be appropriate to the ability or attribute measured, and the different ways of organising these sets of numbers constitute *scales of measurement*.

Unlike physical attributes, such as height, weight, voice pitch, and temperature, we cannot directly observe intrinsic attributes or abilities, and we therefore must establish our measurement scale by definition, rather than by direct comparison.

Bachman (1990: 26) maintains that the scales we define can be distinguished in terms of four properties. A measure has the property of *distinctiveness* if different numbers are assigned to persons with different values on the attribute, and is *ordered in magnitude* if larger numbers indicate larger amounts of the attribute. If equal differences between ability levels are indicated by equal differences in numbers, the measure has *equal intervals*, and if a value of zero indicates the absence of the attribute, the measure has an *absolute zero point*.

Bachman (1990: 26) states that ideally, we would like the scales we use to have all these properties, since each property represents a different type of information, and the more information our scale includes, the more useful it will be for measurement. However, because of the nature of the abilities we wish to measure, as well as the limitations on defining and observing the behaviour that we believe to be indicative of those abilities, we were not able to use scales that possess all four properties for measuring every ability. That is, not every attribute we want to measure, or quantify, fits on the same scale, and not every procedure we use for observing and quantifying behaviour yields the same scale, so that it is necessary to use different scales of measurement, according to the characteristics of attribute we wish to measure and the type of measurement procedure we use. Ratings, for example, might be considered the most appropriate way to quantify observations of speech from an oral interview, while we might believe that the number of items answered correctly on a

multiple-choice test is the best way to measure knowledge of grammar. These abilities are different; as are the measurement procedures used, and consequently, the scales they yield have different properties. The way we interpret and use scores from our measures is determined, largely, by the properties that characterise the measurement scales we use, and it is thus essential for both the development and the use of language tests to understand these properties and the different measurement scales they define. Measurement specialists have defined four types of measurement scales – *nominal*, *ordinal*, *interval*, and *ratio* – according to how many of these four properties they possess.

Nominal scale

Bachman (1990: 27) observes that as its name suggests, a nominal scale comprises numbers that are used to 'name' the classes or categories of a given attribute. That is, we can use numbers as a shorthand code to identify different categories. If we quantified the attribute 'native language', for example, we would have a nominal scale. We would assign different code numbers to individuals with different native language backgrounds, (for example, Xhosa = 1, Sesotho = 2, Venda = 3, Tswana = 4 etc.) and thus create a nominal scale for this attribute. The numbers we assign are arbitrary, since it makes no difference what number we assign to what category, so long as each category has a unique number. The distinguishing characteristic of a nominal scale is that while the categories to which we assign numbers are distinct, they are *not ordered* with respect to each other. In the example above, although '1' (Xhosa) is *not equal* to '2' (Sesotho), it is neither greater than nor less than '2'. Nominal scales thus possess the property of distinctiveness. Because they quantify categories, nominal scales are also sometimes referred to as 'categorical' scales. A special case of a nominal scale is a *dichotomous scale*, in which the attribute has only two categories, such as 'sex' (male or female), or 'status to answer' (right or wrong) on some types of tests.

Ordinal scales

Bachman (1990: 28) suggests that an ordinal scale, as its name suggests, comprise the numbering of different levels of an attribute that are ordered with respect to each

other. The most common example of an ordinal scale is a ranking, in which individuals are ranked 'first', 'second', 'third', and so on, according to some attribute or ability. A rating based on definitions of different levels of ability is another measurement procedure that typically yields scores that constitute an ordinal scale. The points, or levels, on an ordinal scale can be characterised as 'greater than' or 'less than' each other, and ordinal scales thus possess, in addition to the property of distinctiveness, the property of ordering.

Interval Scale

Bachman (1990: 28) proposes that an interval scale is a numbering of different levels in which the distances, or intervals, between the levels are equal. That is, in addition to the ordering that characterises ordinal scales, interval scales consist of equal distances or intervals between ordered levels. Interval scales thus possess the properties of distinctiveness, ordering, and equal intervals.

Ratio scale

Bachman (1990: 29) argues that none of the scales discussed thus far has an absolute zero point, which is the distinguishing characteristic of a ratio scale. Most of the scales that are used for measuring physical characteristics have true zero points. If we looked at a bathroom scale with nothing on it, for example, we should see the pointer at zero, indicating the absence of weight on the scale. The reason we call a scale with an absolute zero point a ratio scale is that we can make comparisons in terms of ratios with such scales. For example, if I have two pounds of coffee and you have five pounds, you have twice as much coffee (by weight) as I have, and if one room is ten feet long another thirty, the second room is three times as long as the first.

2.4.5 Characteristics that limit measurement

Bachman (1990: 30) argues that as test developers and test users, we all sincerely want our tests to be the best measures possible. Thus there is always the temptation to interpret test results as absolute, that is, as unimpeachable evidence of the extent

to which a given individual possess the language ability in question. This is understandable, since it would certainly make educational decisions more clear-cut and research results more convincing. However, we know that our tests are not perfect indicators of the ability we want to measure and that test results must always be interpreted with caution.

Limitations in specification

Bachman (1990: 30) asserts that in any language testing situation, as with any non-test situation in which language use is involved, the performance of an individual will be affected by a large number of factors, such as the testing context, the type of test tasks required, and the time of day, as well as her mental alertness at the time of the test, and her cognitive and personality characteristics. The most important factor that affects test performance, with respect to language testing, of course, is the individual's language ability, since it is language ability in which we are interested.

Bachman (1990: 31) maintains that in order to measure a given language ability, we must be able to specify what it is, and this specification generally is at two levels. At the theoretical level, we can consider the ability as a type, and need to define it to clearly distinguish it *both* from other language abilities and from other factors in which we are not interested, but which may affect test performance. Thus, as the theoretical level we need to specify the ability in relation to or in contrast to, other language abilities, and other factors that may affect test performance. Given the large number of different individual characteristics – cognitive, affective, physical – that could potentially affect test performance, this would be a nearly impossible task, even if all these factors were independent of each other. How much more so, given the fact that not only are the various language abilities probably interrelated, but that these interact with other abilities and factors in the testing context as well. At the operational level, we need to specify the instances of language performance that we are willing to interpret as indicators, or tokens, of the ability we wish to measure. This level of specification, then, defines the relationship between the ability and the test score, between type and token.

Limitations in observation and quantification

In addition to the limitations related to the under specification of factors that affect test performance, there are characteristics of the processes of observation and quantification that limit our interpretations of test results. These derive from the fact that all measures of mental ability are necessarily *indirect, incomplete, imprecise, subjective, and relative*.

❖ **Indirectness**

Bachman (1990: 32) states that in the majority of situations where language tests are used, we are interested in measuring the test taker's underlying competence, or ability, rather than his performance on a particular occasion. That is, we are generally not interested so much, in how an individual performs on a given test on a given day, as his ability to use language at different times in a wide range of contexts. Thus, although our measures are necessarily based on one or more individual observation of performance, or behaviour, we interpret them as indicators of a more long-standing ability or competence.

❖ **Incompleteness**

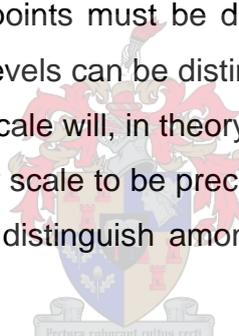
Bachman (1990: 33) observes that in measuring language abilities, we are never able to observe or elicit an individual's total performance in a given language. This could only be accomplished by following an individual around with a tape recorder 24 hours a day for his entire life, which is clearly an impossible task. That is, given the extent and the variation, that characterise language use, it simply is not possible for us to observe and measure every instance of an individual's use of a given language. For this reason, our measures must be based on the observation of a part of an individual's total language use. In other words, the performance we observe and measure in a language test is a *sample* of an individual's total performance in that language.

❖ **Imprecision**

Bachman (1990: 35) suggests that the accuracy or precisions of our measurements are a function of both the representativeness and the number of tasks or units with

which we define our scales. In measuring height, for example, we assume that the yardstick we use is a valid instance, or token, of the standard yardstick. The precision of our measurement of height with this yardstick will increase as we increase the number of units we use from feet to inches to fractions of inches. In measuring language abilities, where we are not dealing with direct physical comparisons, the units of our measurement scales must be defined, and precision, or reliability, becomes, in part, a function of how we define these units. The units that we use to define scales in language tests are most commonly of two kinds: levels or bands, such as in scales for rating compositions, and the number of tasks correctly completed, as in multiple-choice tests.

Bachman (1990: 36) proposes that in developing a rating scale, the points on the scale are typically defined in terms of either the types of language performance or the levels of abilities that are considered distinctive at different scale points. In order to be rated precisely, these scale points must be defined to clearly represent distinct levels, or bands. Assuming the levels can be distinctly defined, the more levels there are, the more precise our rating scale will, in theory, tend to be. However, in order for actual ratings made based on our scale to be precise, it must also be possible for our raters, human beings, to clearly distinguish among all the different levels we have defined.



❖ **Subjectivity**

As Pilliner (1968) noted, language tests are subjective in nearly all aspects. Test developers make subjective decisions in the design of tests and test writers make subjective decisions in producing test items. Furthermore, test takers make subjective judgements in taking tests, and, with the exception of objectively scored tests; scorers make subjective judgements in scoring tests. Language tests are based on either course syllabi or theories of language proficiency, both of which represent the subjective judgements of the individuals who devise them, and tests are therefore neither infallible nor exhaustive definitions of the abilities we may want to test. Furthermore, test developers subjectively select the specific abilities to be included in a given test. For these reasons, a test score should be interpreted as a

reflection of a given test developer's selection of the abilities included in a given syllabus or theory.

Bachman (1990: 38) argues that the procedures by which we arrive at test scores are also determined subjectively. The decision to use an open-ended cloze test rather than a reading passage followed by multiple-choice questions, for example, are subjective, as is the setting of time limits and other administrative procedures. Finally, interpretations regarding the level of ability or correctness of the performance on the test may be subjective. All of these subjective decisions can affect both the reliability and validity of test results to the extent that they are sources of bias and random variation in testing procedures.

❖ *Relativeness*

Bachman (1990: 38) asserts that the last limitation on measures of language ability is the potential relativeness of the levels of performance or ability we wish to measure. When we base test content on domains of language use, or on the actual performance of individuals, the presence or absence of language abilities is impossible to define in an absolute sense. The concept 'zero' language ability is a complex one, since in attempting to define it we must inevitably consider language ability as a cognitive ability, its relationship to other cognitive abilities, and whether these have true zero points. This is further complicated with respect to ability in a second or foreign language by the question of whether there are elements of the native language that are either universal to all languages or shared with the second language.

Bachman (1990: 38) maintains that at the other end of the spectrum, the individual with complete language ability does not exist. From the perspective of language history, it could be argued that given the constant change that characterises any language system, no such system is ever static or 'complete'. From a cognitive perspective, it might be argued that cognitive abilities are constantly developing, so that no cognitive ability is ever 'complete'.

2.4.6 Steps in measurement

Interpreting a language test score as an indication of a given level of language ability involves being able to infer, on the basis of an observation of that individual's language performance, the degree to which the ability is present in the individual. The limitations discussed above restrict our ability to make such inferences. A major concern of language test development, therefore, is to minimise the effects of these limitations. To accomplish this, the development of language tests needs to be based on a logical sequence of procedures linking the putative ability, or construct, to the observed performance. This sequence includes three steps: (1) identifying and defining the construct theoretically; (2) defining the construct operationally, and (3) establish procedures for quantifying observations (Thorndike and Hagen 1977).

Defining constructs theoretically

Bachman (1990: 41) maintains that physical characteristics such as height, weight, and eye colour can be experienced directly through the senses, and can therefore be defined by direct comparison with a directly observable standard. Mental abilities such as language proficiency, however, cannot be observed directly. We cannot *experience* grammatical competence, for example, in the same way as we experience eye colour. We infer grammatical ability through observing behaviour that we presume to be influenced by grammatical ability. The first step in the measurement of a given language ability, therefore, is to distinguish the construct we wish to measure from other similar constructs by defining clearly, precisely, and unambiguously. This can be accomplished by determining what specific characteristics are relevant to the given construct.

Defining constructs operationally

The second step in measurement, according to Bachman (1990: 42), defining constructs operationally, enables us to relate the constructs we have defined theoretically to our observations of behaviour. This step involves, in essence, determining how to isolate the construct and make it observable. Even if it were possible to examine our subjects' brains directly, we would see little that would help

us determine their levels of language ability. We must therefore decide what specific procedures, or operations, we will follow to elicit the kind of performance that will indicate the degree to which definition it will suggest relevant operations. For example, in order to elicit language performance that would indicate a degree of pragmatic competence as defined above, we would have to design a test that would require the subject to process discourse and that would involve both the performance of illocutionary acts and adherence to socio-linguistic rules of appropriateness.

Quantifying observations

The third step in measurement, according to Bachman (1990: 44), is to establish procedures for quantifying or scaling our observations of performance. As indicated above, physical characteristics such as height and weight can be observed directly and compared directly with established standard scales. In measuring mental constructs, however, our observations are indirect and no such standards exist for defining the units of measurement. The primary concern in establishing scales for measuring mental abilities, therefore, is defining the units of measurement. The units of measurement of language tests are typically defined in two ways. One way is to define points or levels of language performance of language ability on a scale. Another common way of defining units of measurement is to count the number of tasks successfully completed, as is done in most multiple-choice tests, where an individual's score is the number of items answered correctly.

Relevance of steps to the development of language tests

Bachman (1990: 45) proposes that these general steps in measurement provide a framework both for the development of language tests and for the interpretation of language test results, in that they provide the essential linkage between the unobservable ability or construct we are interested in measuring and the observation of performance, or the behavioural manifestation, of that construct in the form of a test score.

Relevance of steps to the interpretation of test results

Bachman (1990: 48) states that the steps in measurement discussed above also relate to virtually all concerns regarding the interpretation of test results. The first step, defining constructs theoretically, provides the basis for evaluating the validity of the uses of test scores. The second step, defining constructs operationally, is also related to test validity, in that the observed relationship among different measures of the same theoretical construct provide the basis for investigating concurrent relatedness. The appropriateness of our operational definitions, or testing methods, will also affect the authenticity of the test tasks, and the way the test is perceived by test takers and test users. Finally, the third step, how we quantify our observations, is directly related to reliability.

2.4.7 Research

Douglas (1995: 175) observes that a great deal of research in language testing has been carried out during the first half of the decade. At least one ambitious testing program has been developed which makes direct use of SLA findings on developmental sequences (Mackey, *et al.* 1991, Pienemann 1992), although the system has been criticised as relying too heavily on unproven acquisition sequences and a too narrow view of second language acquisition (Hudson 1993b). In the area of test reliability, Swain (1993) has argued that estimates of internal test consistency are incompatible with current understandings of the nature of second language communicative ability, and she presents data to illustrate the difficulty. Anderson, Bachman, Perkins and Cohen (1991) also argue that multiple sources of data are necessary in the test validation process, and they employ a “triangulation” approach, taking account of test-taking strategies, content analysis, and performance data in their analysis. Similar multi-faceted approaches to test analysis are developed and employed by Chapelle and Douglas (1993) and Douglas (1994).

Douglas (1995: 175) states that research has also been conducted on norms and rating scales in testing. For example, Davidson (1994) and Hamilton, *et al.* (1993) have both studied the appropriacy of using native-speaker groups to set standards of second language performance. With regard to scales, Chalhoub-Deville (1994),

Lazaraton (1993), and Wigglesworth (1993) have investigated sources of bias in rating scales.

Douglas (1995: 176) maintains that studies such as these are representative of a current general interest in understanding sources of error in trait measurement. Studies of the effect of test-method characteristics on interpretations of performance are also concerned with identifying non-trait components of test results. For example, Bachman, Davidson and Milanovic (1991) have studied the method characteristics of TOEFL and the Cambridge First Certificate; similarly Clapham (1991) analyses performance on the reading section of the IELTS. Studies by Bradshaw (1990), Brown (1993), and Cohen (1994b) are all interested in testees' understanding and interpretation of the test-taking process; these studies employ think-aloud and test taker feedback techniques to investigate test-takers' strategies and to understand their effect on test performance. Taking a slightly different approach to the same questions, Allan (1992) developed a "test-wiseness" assessment scale. In a complementary approach, Amer (1993) taught learners test-taking strategies and investigated their effects on test performance.

2.4.8 Interlanguage and error analysis

Brown (1994) argues that the *markedness* theory distinguishes members of a pair of related forms or structures by assuming that the marked member of a pair contains at least one more feature than the unmarked one. The addition, the unmarked member of the pair is the one with the wider range of distribution than the marked one. He refers to Eckman, who showed that marked items in a language will be more difficult to acquire than unmarked, and that degrees of markedness will correspond to degrees of difficulty.

Universal grammar (UG) is the 'rules' that are shared by all languages. Such rules are a set of limitations or parameters of language. Different languages set their parameters differently. According to Brown, markedness theory and UG perspectives provide a more sophisticated understanding of difficulty in learning a second language than we had previously from the early formulations of the CAH. But describing and predicting difficulty amidst all the variables of human learning is still an elusive process. As in the case with virtually every problem in linguistic analysis,

our scientific methodological capacities are currently inadequate to give a complete account.

Interlanguage – According to Brown, the CAH stressed the interfering effects of the first language on second language learning and claimed, in its strong form, that second language learning is primarily, if not exclusively, a process of acquiring whatever items are different from the first language. Second language is a creative process of constructing a system in which learners are consciously testing hypotheses about the target language from a number of possible sources of knowledge. The interlanguage hypothesis led to a whole new area of second language research and teaching and presented a significant breakthrough from the shackles of the contrastive analysis hypothesis.

Error analysis – Brown maintains that a mistake reverts to a performance error that is either a random guess or a 'slip', in that it is a failure to utilize a known system correctly. All people make mistakes, in both native and second language situations. Native speakers are normally capable of recognizing and correcting such 'lapses' or mistakes. An error is a noticeable deviation from the adult grammar of a native speaker, reflecting the interlanguage competence of the learner. Errors – overt manifestations of learners' systems – arise from several possible general sources: interlingual errors of interference from the native language, intralingual errors within the target language, the sociolinguistic context of communication, psycholinguistic or cognitive strategies, and no doubt countless affective variables.

Errors in error analysis – The classroom foreign language teacher can become so preoccupied with noticing errors that the correct utterances in the second language go unnoticed. We must beware of placing too much attention on errors and not lose sight of the value of positive reinforcement of clear, free communication. The ultimate goal of second language learning is the attainment of communicative fluency in a language. We do well in the analysis of learners' interlanguage errors, to engage in 'performance analysis' – or perhaps more simply, 'interlanguage analysis', a less restrictive concept that places a healthy investigation of errors within the larger perspective of the learner's total interlanguage performance.

Identifying and describing errors – One of the common difficulties in understanding the linguistic systems of both first and second language learners is the fact that such systems cannot be directly observed. They must be inferred by means of analyzing production and comprehension data. What makes the task even thornier, however, is the instability of learners' systems. Systems are in a constant state of flux as new information flows in and, through the process of subsumption, causes existing structures to be revised.

Overt and covert errors – Overtly erroneous / utterances are unquestionably ungrammatical at the sentence level. Covertly erroneous utterances are grammatically well-formed at the sentence level but are not interpretable within the context of communication. Overt error, for example: O sa phela? – Re sa phela. (Ke sa phela). Covert error, for example: Lebitso la hao ke mang? – Ke *mosadi* (Ke *Nthati*).

Stages of interlanguage development.

The first is a stage of random errors, a stage that Corder called "presystematic", in which the learner is only vaguely aware that there is some systematic order to a particular class of items. The learner is making rather wild guesses at what to write.

The second, or emergent, stage of interlanguage finds the learner growing in consistency in linguistic production. The learner has begun to discern a system and to internalize certain rules. Generally the learner is still, at this stage, unable to correct errors when they are pointed out by someone else.

A third stage is, according to Brown, a truly systematic stage in which the learner is now able to manifest more consistency in producing the second language. The most salient difference between the second and third stage is the ability of learners to correct their errors when they are pointed out to them.

A final stage, according to Brown, is the stabilization stage in the development of interlanguage systems, is akin to what Corder called a 'postsystematic' stage. Here the learner has relatively few errors and has mastered the system to the point that fluency and intended meanings are not problematic. The learner has the ability to self-correct.

Brown states that one simply must expect a good proportion of interlanguage data to fall beyond our capacity for systematic categorization. It should also be made clear that the four stages of systematicity outlined above do not describe a learner's total second language system.

Brown identifies the following sources of error:

Interlingual transfer – The beginning stages of learning a second language are characterized by a good deal of interlingual transfer from the native language, or interference. In these early stages, before the system of the second language is familiar, the native language is the only linguistic system in previous experience upon which the learner is the only linguistic system in previous experience upon which the learner can draw. These errors are attributable to negative interlingual transfer.

Intralingual transfer – Intralingual transfer (within the target language itself) is a major factor in second language learning. Researchers have found that the early stages of language learning are characterized by a predominance of interference (interlingual transfer), but once learners have begun to acquire parts of the new system, more and more intralingual transfer – generalization within the target language – is manifested.

Context of learning – The context of learning overlaps both types of transfer. In a classroom context the teacher of the textbook can lead the learner to make faulty hypotheses about the language. The socio-linguistic context of natural, untutored language acquisition can give rise to certain dialect acquisition that may itself be a source of error.

Communication strategies – Communication strategies are, according to Brown, related to learning styles. Learners obviously use production strategies in order to enhance getting their message across, but at times these techniques can themselves become a source of error.

Fossilization – The relatively permanent incorporation of incorrect linguistic forms into a person's second language competence has been referred to as fossilization. Fossilization items, according to Brown, are those ungrammatical or incorrect items in the speech of a learner that gain first positive affective feedback (I like it), then positive cognitive feedback (I understand), reinforcing an incorrect form of language.

Learners with fossilization items have acquired them through the same positive feedback and reinforcement with which they acquired correct items.

Error correction – Brown states that one of the keys, but not the only key, to successful learning lies in the feedback that a learner receives from others. Fossilization may be the result of too many green lights when there should have been some yellow or red lights. The task of the teacher is to discern the optimal tension between positive and negative cognitive feedback: providing enough green lights to encourage continued communication, but not so many that crucial errors go unnoticed, and providing enough red lights to call attention to those crucial errors, but not so many that the learner is discouraged from attempting to speak at all. Bailey gives seven 'basic options' that are complemented by eight 'possible features' within each option. According to Brown, the basic options are to treat or ignore, to treat immediately or to delay, to transfer treatment or not, to transfer another individual, a subgroup, or the whole class, to return, or not, to original error maker after treatment, to permit other learners to initiate treatment and to test for the efficiency of the treatment. The possible features are fact or error indicated, location indicated, opportunity for new attempt given, model provided, error type indicated, remedy indicated, improvement indicated and praise indicated. Brown states that one general conclusion that can be drawn from the study of errors in the interlanguage systems of learners is that learners are indeed creatively operating on a second language, a system for understanding and producing utterances in the language. The system should rather be looked upon as a variable, dynamic, approximate system, reasonable to a great degree in the mind of the learners, albeit idiosyncratic. The teacher's task, according to Brown, is to value learners, prize their attempts to communicate, and then to provide optimal feedback for the system to evolve in successive stages until learners are communicating meaningfully and unambiguously in the second language.

CHAPTER THREE

TASK-BASED LANGUAGE TEACHING AND ASSESSMENT

3.1 WHAT ARE TASKS?

According to McDonough and Shaw the notion of task usually refers to a specialized form of technique or series of techniques, closely allied with communicative curricula, and as such must minimally have communicative goals. The common thread running through half a dozen definitions of tasks is its focus on the authentic use of language for meaningful communicative purposes beyond the language classroom. McDonough and Shaw state that the term 'activity' may refer to virtually anything that learners actually do in the classroom. Because an activity implies some sort of active performance on the part of the learners, it is generally not used to refer to certain teacher behaviours like saying "good morning". Procedure is the actual moment-to-moment techniques, practices, and behaviours that operate in teaching a language according to a particular method. McDonough and Shaw state that the term 'technique' is accepted as a superordinate term to various activities that either teachers or learners perform in the classroom. Techniques include all tasks and activities. They are usually planned and deliberate. You can think of a lesson as consisting of a number of techniques.

Long (1985: 89) defines task generically as:

A piece of work undertaken for oneself or for other, freely or for some reward. Thus examples of tasks include painting a fence, dressing a child, filling out a form, buying a pair of shoes, making an airline reservation, borrowing a library book, taking a driving test, typing a letter, weighing a patient, sorting letters, taking a hotel reservation, writing a check, finding a street destination and helping someone across a road. In other words, by 'task' is meant the hundred and one things people *do* in everyday life, at work, at play, and in between.

For Long (1985), then, a task is fundamentally something that is accomplished or done, and language is involved insofar as it is inherently necessary for accomplishment of the task.

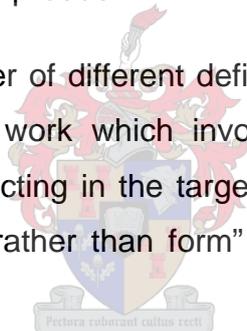
Crookes (1986: 1) also offers a broad definition of tasks in the real world, but he specifically includes the possibility of classroom and laboratory applications:

A piece of work or an activity, usually with a specified objective, undertaken as part of an educational course, at work, or used to elicit data for research.

Doyle (1983: 161) has the following definition for academic tasks:

The term “task” focuses attention on three aspects of students’ work: (a) the products students are to formulate... (b) the operations that are to be used to generate the product... and (c) the “givens” or resources available to students while they are generating a product.

Nunan (1989) draws on a number of different definitions to define a communicative task as “a piece of classroom work which involves learners in comprehending, manipulating, producing or interacting in the target language while their attention is principally focused on meaning rather than form” (p. 10). Later he discusses two categories of tasks:



- 1) Real-world tasks, which are designed to practice or rehearse those tasks which are found to be important in a needs analysis and turn out to be important and useful in the real world, and
- 2) Pedagogic tasks, which have a psycholinguistic basis in SLA theory and research but do not necessarily, reflect real-world tasks. (p. 40 – 41)

Michael Breen’s definition of task seems to capture its essence: “any structured language learning endeavour which has a particular objective, appropriate content, a specified working procedure and a range of outcomes for those who undertake the task”. A task is really a special form of technique. In some cases, task and technique may be synonymous. Tasks are usually “bigger” in their ultimate ends than techniques. It views the learning process as a set of communicative tasks that are directly linked to the curricular goals they serve. Research on task-based

learning attempts to identify types of tasks that enhance learning. Task based learning is a perspective that you can take within a CLT framework that forces you to carefully consider all of the techniques that you use in the classroom in terms of a number of important pedagogical purposes:

- a) Do they ultimately point learners beyond the forms of language alone to real-world contexts?
- b) Do they specifically contribute to communicative goals?
- c) Are their elements carefully designed and not simply haphazardly or idiosyncratically thrown together?
- d) Are their objectives well specified so that you can at some later point accurately determine the success of one technique over another?
- e) Do they engage learners in some form of genuine problem solving activity?

Target tasks are what the students must accomplish beyond the classroom. Target tasks are much more specific and more explicitly related to classroom instruction.

Pedagogical tasks are which form the nucleus of the classroom activity. Pedagogical tasks include any of a series of techniques designed ultimately to teach students to perform the target task. Pedagogical tasks are distinguished by their specific goals that point beyond the language classroom to the target task.

A task-based curriculum, then, specifies what a learner needs to *do* with the Sesotho language in terms of target tasks and organizes a series of pedagogical tasks intended to reach those goals.

3.2 WHAT IS THE ROLE OF NEEDS ANALYSIS IN TASK SELECTION?

McNamara raises the question of the reasons for the design of the specific test. This involves clearly stating the specific inferences about language ability or capacity for language use we intend to make based on the results, and any specific decisions, which will be based upon these inferences. This will constrain the test specifications

in basic ways. Educational, social, and political contexts motivate tests. McNamara states that as a preliminary to investigating the domain to be covered by the test, and then creating test specifications, it will be helpful to establish for the particular test development project,

Who want to know what about whom, for what purpose?

McNamara further poses the question: How are relevant tasks to be identified and sampled for the purposes of assessing second language performance?

Norris, Brown, Hudson & Yoshioka (1998: 34) find that a thorough analysis of the real-world needs that learners have in terms of tasks, task elements, and task types plays a central role in selecting pedagogic tasks. Long and Crookes (1992: 37) suggests that:

It is impossible for anyone to verify the appropriacy of particular pedagogic tasks for a given group of learners without objective evaluation criteria, one of which must surely be relevance to learner needs.

Indeed, an effective needs analysis should indicate both pedagogic-oriented and assessment-oriented tasks. Long and Crookes (1993) suggest that, once identified, target tasks serve more as pedagogic goals than as the actual tasks students will undertake in the classroom, mainly because target tasks “would often be too difficult, inefficient in terms of class time, logistically impossible, and irrelevant for some learners in heterogeneous classes when students’ future needs vary” (p. 40). Instead, in task-based syllabus construction, target tasks are often collapsed and classified into more general task types. “To take a simple example, serving breakfast, serving lunch, serving dinner and serving snacks and refreshments, might be classified into ‘serving food and beverages’ in a course for trainee flight attendants” (p. 40). Once organised, such task types can then provide the basis for developing pedagogic tasks, which are later sequenced to create a task-based syllabus – all of which Long and Crookes suggest to be quite a complex task in itself.

However, according to Norris, Brown, Hudson & Yoshioka (1998: 36), in terms of task-based performance assessments, which are designed to test the real-world outcomes of such task-based pedagogy, use of the actual tasks in authentic

situations (to the extent that it is feasible) is called for. Here, then, assessment tasks must diverge from pedagogic tasks. We content that it is imperative for assessment tasks to be preserved with as much authenticity as possible, in order to legitimately reflect learner abilities to perform under real-world conditions. Therefore, needs analysis is crucial for determining pedagogic sequencing and emphasis and for establishing desired outcomes in terms of task performance (expert judgements of what constitutes success or failure, minimal criteria to be met in accomplishing the task, etc.)

In Long and Crookes' approach, target tasks, once identified, would, for implementational reasons, be broken down into component 'pedagogic' tasks, and sequenced according to some notion of complexity or sequencing potential.

3.3 WHAT ARE THE FACTORS THAT AFFECT TASK DIFFICULTY AND SEQUENCING?

Skehan (1998: 100) maintains that giving learners tasks to do, even when these tasks are motivating and engage natural communicational abilities, is not enough: they also have to form part of a larger pedagogic plan. According to Norris, Brown, Hudson & Yoshioka (1998: 39) setting task parameter (through task definition) and selecting appropriate tasks (through needs analysis) are integral, feasible, and well-motivated components of performance assessment design. The final task-related factor to be considered from the TBLT perspective involves the notion of differentiating among tasks based on their difficulty.

Skehan (1998: 134) asserts that tasks of appropriate difficulty are required to respond to reasonable challenges, which are achievable if there is effort on their part. If the appropriate level of task difficulty is chosen so that balanced, language performance will result. Task-based approaches, therefore, need to focus on task difficulty as a precondition for any task work. Weir (1993:22) points out that if there is a variety of tasks/items testing a particular ability, try and put easier tasks/items first, as this will encourage all students to try their hardest and show their best. If you start with the most difficult task the weakest will soon give up. Within tasks it may not always be possible to begin with the easier items as one would want to preserve the sequential ordering of items in relation to their occurrence in the text.

Crookes (1986), citing Long (1985), suggests that within task-based syllabus design, “difficulty is a prime consideration”, and he lists the following possible contributors:

- 1) Number of steps needed
- 2) Number of parties involved
- 3) Presupposed knowledge
- 4) Intellectual challenge
- 5) Spatio-temporal displacement (p. 24)

Crookes also notes that these possible contributors to task difficulty are hypothesised rather than founded on empirical evidence.

According to Nunan (1989), once a number of related tasks have been designed, they need to be graded (ranked in order of difficulty) so they can later be sequenced and appropriately integrated into the syllabus. Such grading, he maintains, is not an easy endeavour because it is affected by a great number of factors. The level of difficulty of the input, for example, as well as the following:

...the length of the text, the prepositional density (how much information it contains and the extent to which this information is recycled), the amount of low frequency vocabulary, the speed of spoken texts and the number of speakers involved, the explicitness of the information, the discourse structure and the clarity with which this is signalled (for example paragraphs in which the main point is buried away will probably be more difficult to process than those in which the main idea is clearly presented in the opening sentence). (p. 98)

Furthermore, Nunan asserts that the way a text may be modified and the genre it is written in can also affect the difficulty of the input. Nunan (1989) points to factors other than input that can influence the level of task difficulty:

- 1) *Learner factors*, including their confidence, motivation, prior learning experience, learning pace, observed ability in language skills, cultural knowledge are awareness, linguistic knowledge (pp. 102 – 103); as well as

- 2) *Activity factors*, including relevance, complexity, amount of context provided prior to task, process ability of the language of the task, amount of help available to the learner, degree of grammatical accuracy/contextual appropriacy, time available to learner, cognitive load, communicative stress, particularly and generalisability, code complexity and interpretive density, and process continuity. (pp. 109 – 111)

Nunan (1993) acknowledges the *difficulty* with task difficulty:

Grading and sequencing are carried out with reference to priority of learner needs and also with reference to notions of difficulty. Determining difficulty is a major problem because of the number of factors involved... In addition, these factors interact. (p. 60)

Nunan refers to Honeyfield (1993), who takes a different approach to pedagogic task difficulty. He discusses general “instructional strategies” that can be used to design tasks that make new demands on learners while doing so in a feasible way. One such instructional strategy deals mainly with task difficulty and how learner factors and task factors can be combined and modified to make a given task easier or more difficult depending on the need. As a simple illustration, Honeyfield provides a list, which includes the following considerations:

- 1) Procedure, or what the learners have to do to derive output from input
- 2) Input text
- 3) Output required [note that point a. -d. below may need to be considered for both input and output]
 - a. Language items: vocabulary, structures, discourse structures, process ability, and so forth
 - b. Skills, both macra-skills and sub-skills
 - c. World knowledge or “topic content”
 - d. Text handling or conversation strategies
- 4) Amount and types of help given
- 5) Roles of teacher and learners

- 6) Time allowed
- 7) Motivation
- 8) Confidence
- 9) Learning styles (p. 129)

Robinson (1996a) discusses a similar approach to the classification and grading of tasks for syllabus design purposes, but he advises important conceptual differences in using the terms *task difficulty* and *task complexity*. He summarises these differences as follows:

Task *difficulty* is determined by learner variables (affect, e.g., confidence, and physical abilities, e.g., eyesight); task *complexity* is determined by task factors (its point along a dimension of complexity, e.g., that of planning time, and task conditions, e.g., the one way or two way direction of information flow). (p. 4)

Robinson refers to Perkins and Linnville (1987), who propose that they found that word length (number of letters, number of syllables); frequency and abstractness were the only characteristics that consistently predicted item difficulty across three different levels of English proficiency.

Fulcher's (1996) results indicate that group oral discussion as a task type offers potentially more enjoyable, anxiety-reducing, perceptively valid alternative for students, regardless of ability level, than one-on-one interviews in oral testing. This study represents an important first step toward investigating the variety of task types that can be used in performance assessment, but more empirical research, of course, needs to be conducted.

Skehan (1996: 53) asserts that tasks should be analysed, compared, and, best of all, sequenced according to some principled basis. The rewards, if tasks are well chosen are that an effective balance between fluency, accuracy, and the opportunity for previous restructuring to be applied. Tasks which are too difficult are likely to over-emphasize fluency, as learners only have the attentional capacity to convey meanings, using production strategies, lexicalised language, and making meaning primary. If the tasks are too easy, they will present no challenge, and are not likely to extend any other goals of restructuring, accuracy, or fluency in any effective way.

Degree of structure

Skehan (1998) proposes that this variable has an effect upon the fluency and accuracy of performance. He suggests that tasks which contain clear structure, especially sequential structure, facilitate task performance by clarifying the macrostructure of the speech event. Bygate, Skehan and Swain (2001) asserts that we can summarise the results in this section by saying that there is a fairly consistent pattern that tasks based on more structured information seem to be associated with greater fluency. There are some indications that accuracy might also be enhanced, but the evidence is, to say the least, mixed, and so it would be unwise at this stage to make any claims in this direction. If on-line planning and attentional availability are facilitated by structured tasks, these are directed towards fluency.

3.4 EVALUATING ENGLISH LANGUAGE TEACHING (ELT) MATERIALS

McDonough and Shaw state that for some teachers the selection of a good textbook can be valuable, particularly in contexts where the assimilation of stimulating, authentic materials can be difficult to organize. Teachers may well be interested in the evaluation as a useful process in its own right, giving insight into the organizational principals of the materials and helping them to keep up with developments in the field. We thus examine criteria in two stages; an external evaluation which offers a brief “overview” of the materials from the outside which is then followed by a closer and more detailed internal evaluation. The evaluation process is never static. When materials are deemed appropriate for a particular course after a preliminary evaluation, their ultimate success or failure may only be determined after a certain amount of classroom use.

External evaluation: According to McDonough and Shaw, in this central stage of the model, we have included criteria, which will provide a comprehensive, external overview of how the materials have been organized. Our aim is examining the organization of the materials by looking at: (a) the “blurb”, or the claims made on the cover of the book and (b) the introduction and table of contents. The table of contents often represents a “bridge” between the external claim made for the materials and what will actually be presented “inside” the materials themselves. At

this stage, we need to consider why the materials have been produced. During this external evaluation stage we have examined the claims made for the materials by the author / publisher with respect to: the intended audience, the proficiency level, the context and presentation of language items, whether the materials are to be core or supplementary, the role and availability of a teachers' book, the inclusion of a vocabulary list / index, the table of contents, the use of visuals and presentation, the cultural specificity of the materials, the provision of audio / video material and inclusion of tests. After completing this external evaluation we can arrive at a decision as to materials' appropriacy for adoption / selection purposes. If our evaluation shows the materials to be potentially appropriate and worthy of a more detailed inspection then we can continue with our internal or more detailed evaluation.

Internal evaluation: According to McDonough and Shaw, the internal evaluation procedure performs an in-depth investigation into the materials. The essential issue at this stage is for us to analyse the extent to which the aforementioned factors in the external evaluation stage actually match up with the internal consistency and organization of the materials as stated by the author / publisher – for strong claims are often made for these materials. In order to perform an effective internal inspection of materials, we need to examine at least two units of a book or set of materials to investigate the following factors. The treatment and presentation of the skills, the sequencing and grading of materials, the type of reading, listening, speaking and writing contained in the materials, appropriacy of tests and exercises, self study provision and teacher – learner “balance” in use of the materials.

Overall evaluation: at this stage, McDonough and Shaw state, we may now make an overall assessment as to the suitability of the materials by considering the following parameters: the usability factor, the generalisability factor, the adaptability factor and the flexibility factor. The aim of this final stage is, according to McDonough and Shaw, intended to enable the evaluator to decide the extent to which the materials have realized their stated objectives. Even after the internal evaluation, we still have the option of not selecting the materials if we so wish. We must bear in mind that their ultimate success or failure can only be determined after trying them in the classroom with real learners.

3.5 HOW DO WE ASSESS TASK-BASED PERFORMANCE?

According to Skehan (1998: 290) communicative tests have been seen as those, which meet a set of performance conditions, but there has been little attempt to conceptualise the dimensions of performance in any convincing way. As a result, abilities-based approaches 'clothed' by communicative tests have been regarded as the most effective way to assess.

A test, in plain, according to Skehan, is a method of measuring a person's ability or knowledge in a given area. A test measures a person's ability or knowledge. Care must be taken in any test to understand who the testees are. Also being measured in a test is ability or knowledge – that is, competence. A test samples performance but infers certain competence. A language test samples language behaviour and infers general ability in a language. A test of reading comprehension may consist of some questions following one or two paragraphs, a tiny sample of a second language learner's total reading behaviour. From the results of that test, the examiner infers a certain level of general reading ability. Finally, Skehan states, a test measures a given area. In the case of a proficiency test, even though the actual performance on the test involves only a sampling of skills, that area is overall proficiency in a language – general competence in all skills of a language. Other tests may have criteria that are more specific. A test of pronunciation might well be a test only of a particular phonemic minimal pair in a language. One of the biggest obstacles to overcome in constructing adequate tests is to measure the criterion and not inadvertently something else.

Norris, Brown, Hudson & Yoshioka (1998: 53) observes that several approaches to task-based testing can be found within the TBLT literature, and, as with any kind of language assessment, these various approaches are tied to a range of decision-making purposes. Each of these approaches also incorporates specific procedures for the analysis and/or evaluation of task performance.

One approach to task-based testing involves the assessment of task outcomes in terms of learner/examinee success or failure to accomplish the task. Such an outcomes-referenced approach would be utilised for the purposes of certifying that learners can accomplish particular tasks that have been identified (through a needs

analysis) as seminal within a particular curriculum. This approach draws from occupational performance testing for certification purposes, and it has parallels in language testing within the English for Specific Purposes literature (see Clark & Grognet, 1985 for an example of task-based outcomes testing within an English for survival purposes program). Robinson (1996b) also notes that such tests have been utilised for some time in order to make exit decisions in vocational training programs, and he cites similar efforts at task-based exit testing in language programs (e.g., Hauptman, LeBlanc, & Wesche, 1985).

Performances on such outcomes-referenced tests are, according to Robinson, evaluated in terms of simple yes/no decisions; that is, the pivotal question is, *Did the examinee accomplish the goal of the task or not?* Finally, the tasks that are utilised in outcomes-referenced tests are necessarily drawn directly from an analysis of the real-world needs that learners have in terms of language tasks that they will be faced with upon exit or certification. Criteria for the yes/no decision also issue from the needs analysis, and they are based on those aspects specific to a given task that prove to be essential in its accomplishment.

Robinson (1996b) has noted the limitations of outcomes-referenced testing for second language pedagogical purposes:

- 1) They are often difficult to design and administer.
- 2) They lack generalisability across programs.
- 3) They are considered uneconomical. (p. 96)

Norris, Brown, Hudson & Yoshioka (1998: 54) elaborate on these pedagogic limitations by suggesting that outcomes-referenced testing does not provide the learner/teacher with useful feedback in terms of the particular aspects of a task that prove more or less difficult for an examinee, nor does it necessarily reflect the efficiency with which a learner utilises the L2 for task accomplishment.

Norris, Brown, Hudson & Yoshioka (1998: 54) state that with respect to L2 performance assessment, they occur with Robinson that task-based testing in general is difficult and uneconomical to design and administer, although we feel that (a) these problems can be solved, and (b) task-based testing is nonetheless requisite

for claims regarding learner's abilities to use a given L2 under real-world conditions. However, Robinson's point number 2 above, which poses the greatest limitation for outcomes-referenced testing, is also one of the central issues for any approach to task-based testing that seeks to do more than certify learner abilities with respect to isolated real-world tasks. We turn therefore to other suggestions from the literature in search of solutions to the problem of generalisability in task-based assessment.

Robinson (1996b) distinguishes between performance-referenced and system-referenced testing as two possible approaches to task-based assessment related to a task-based language curriculum (see also Robinson & Ross, 1996). Drawing on Baker (1990), Robinson indicates that system-referenced tests attempt to tap a particular psychological construct (e.g., developmental level), which underlies a language task, without analysing the accomplishment of the task itself. Although these tests do not, according to Robinson, provide feedback regarding real-world task performances, it is held that they can provide generalisable information regarding some component of a learner's language ability that might underlie the accomplishment of any number of different tasks. Performance-referenced tests, on the other hand, approximate as closely as possible the conditions of a future language task, and they therefore retain high face validity. However, generalising to other task performance or L2 abilities from such performance-referenced tests is problematic.

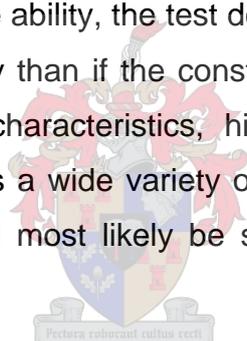
Skehan (1998: 287) suggests that performance itself is far more interesting and important than it has been given credit for within testing, and that it is essential, if effective assessment is to proceed, that a range of processing conditions are built in to the test formats which are used. Skehan (1998) comments that task difficulty factors will interact with rating scales. He suggests, therefore, that an approach to the evaluation of task performances should stress those "areas, which have some linkage to the nature of processing" (p. 274).

3.6 WHAT ARE THE FACTORS THAT AFFECT TASK-BASED ASSESSMENT RELIABILITY?

Backman and Palmer state that a reliable test is a test that is consistent and dependable. Sources of unreliability may lie in the test itself or in the scoring of the

test, known respectively as test reliability and rater (or scorer) reliability. If you give the same test to the same subject or matched subjects on two different occasions, the test itself should yield similar results: it should have test reliability. Scorer reliability is the consistency of scoring by two or more scorers. If very subjective techniques were employed in the scoring test, one would not expect to find high scorer reliability.

Bachman & Palmer (1996: 135) asserts that probably the most important consideration in setting a minimum acceptable level of reliability is the purpose for which the test is intended. While the test developer will want to set the minimum acceptable level of reliability as high as possible, this level needs to be realistic. Two other considerations need to be kept in mind – the way the construct has been defined and the nature of the test tasks – as these will affect the level of reliability that one can expect to achieve. If the construct definition focuses on a relatively narrow range of components of language ability, the test developer can reasonably expect to achieve higher levels of reliability than if the construct is complex. If the test tasks are relatively uniform in their characteristics, higher levels of reliability can be expected than if the test includes a wide variety of types of test tasks. A minimum acceptable level of reliability will most likely be stated in terms of an appropriate statistical estimate of reliability.



Alderson, Clapham and Wall (1995: 87) argues that although it is usually impossible to achieve a perfectly reliable test, test constructors must make their tests as reliable as possible. They can do this by reducing the causes of unsystematic variation to a minimum. They should ensure, for example, that the test is administered and marked consistently, that the test instructions are clear, and that there are no ambiguous items. As we have seen from the section on discrimination indices, ambiguous or faulty items have low discrimination indices, and a test, which contains many such items, will tend to be unreliable.

According to Alderson et al, inter-rater reliability refers to the degree of similarity between different examiners: can two or more examiners, without influencing one another, give the same marks to the same set of scripts or oral performances? It would not be realistic to expect all examiners to match one another all the time; however, it is essential that each examiner try to match the standard all the time.

Inter-rater reliability is usually measured by means of a correlation coefficient, or through some form of analysis of variance according to Alderson, Clapham, and Wall (1995: 290).

Intra-rater reliability – An examiner is judged to have intra-rater reliability if he or she gives the same marks to the same set of scripts or oral performances on two different occasions. The examiner may still be considered reliable even if some of the marks are different; however, not much variation can be allowed before the reliability becomes questionable. Intra-rater reliability is usually measured by means of a correlation coefficient, or through some form of analysis of variance according to Alderson, Clapham, and Wall (1995: 290).

Alderson et al, refer to Brindley (1994), who refers to task-based assessment as task-centered assessment (TCA). He points out that the reliability of ratings is a major problem in task-centered assessment for the classroom.

TCA relies heavily on teacher's subjective judgements of language performance. In the interests of fairness to learners, it is important that these judgements are seen to be reliable. As more and more rating tools are developed to assess productive task performance, teachers will need to be trained to interpret and apply assessment instruments in a consistent way. Rater training involving familiarisation with the rating criteria and practice in applying them to sample of performances across a range of ability levels has long been standard practice with proficiency rating scales and it has been claimed that high levels of inter-rater agreement can be obtained in this way (e.g., Dandonoli and Henning, 1991: 84).

Although rater training may be the procedure usually followed to insure reliability, there appears to be a chink in the armor. According to various citations in Brindley's article, a rater's tendency for severity or leniency in judgements seems to remain unchanged despite having rater training. Since this factor cannot be easily accounted for or eliminated from the judgements, it introduces a worrisome element in the calculations that could affect inter-rater reliability. Luckily, new measurement technology offers new hope. According to Brindley (1994), one such tool is item response theory:

The Rasch model is one of a family of techniques known as latent trait theory or item response theory (IRT), which have been developed by psychometricians over the last three decades or so. One of the strengths of the theory is that it allows candidate ability and item difficulty to be estimated independently and reported on a common scale, thus avoiding many of the problems associated with sample-dependent classical measurement techniques (Henning 1987). The multi-faceted Rasch model extends previous Rasch models to include rater characteristics... The program adjusts candidate ability estimates to take account of rater's' tendency to rate either harshly or leniently. (p. 85)

Brindley also discusses another tool, FACETS, a software application, that not only performs the above multi-faceted, adjustable Rasch analysis but also has the added benefit of informing the raters of how they are rating (i.e., too leniently, too harshly), thus, serving in a way as a rating self-awareness program. Although such new measurement tools are quite useful, they may unfortunately be out of reach to many people due to costs or availability.

Fulcher (1996) raises another reliability problem: the potential lack of generalisability from one task performance to another. However, he notes that this lack of generalisability may be due to the scoring system employed (especially if it does not refer to test method facets in its descriptors) that to the design of the task. In fact, such a lack of generalisability of one task to another is not only related to the consistency or reliability across tasks but also to the validity of those tasks.

Bachman (1990: 220) argues that long tests are generally more reliable than short ones. If we can assume that all the items or tasks included are representative indicators of the ability being measured, the more we include, the more adequate our sample of that ability. Not only does a longer test tend to provide a more adequate sample of language use, it also tends to minimize the influence of various sources of bias. A test cannot be valid unless it is reliable. On the other hand, it is quite possible for a test to be reliable but invalid.

3.7 WHAT ARE THE FACTORS THAT AFFECT TASK-BASED ASSESSMENT VALIDITY?

According to Bachman and Palmer, validity is the degree to which the test actually measures what it is intended to measure. Validity can only be established by observation and theoretical justification. If a test actually samples the class of situations, that is, the universe of subject matter about which conclusions are to be drawn, it is said to have content validity. The test actually involves the testee in a sample of the behaviour that is being measured. You can usually determine content validity observationally, if you can clearly define the achievement that you are measuring.

According to Bachman and Palmer, face validity asks the question: does the test, on the “face” of it, appear to test what it is designed to test? To achieve “peak” performance on a test, a learner needs to be convinced that the test is indeed testing what it claims to test. Face validity is almost always perceived in terms of content: if the test samples the actual content of what the learner has achieved or expects to achieve, then face validity will be perceived.

One way to look at construct validity is, according to Bachman and Palmer, to ask the question: does this test actually tap into the theoretical construct as it has been defined? Tests are, in a manner of speaking, operational definitions of such constructs, in that they operationalize the entity that is being measured. It is important for the teacher to know that if a learner is doing well in a test, that the test is indeed testing what it claims to test. Some learners can be upset because tests, on the face of it, cannot appear to them to test their true abilities in Sesotho.

Bachman & Palmer (1996: 29) propose that the relationship between interactivensness and construct validity is a function of the relative involvement of areas of language knowledge, strategic competence, or metacognitive strategies, and topical knowledge. That is, the extent to which high interactivensness corresponds to construct validity will depend on how we have defined the construct and on the characteristics of the test takers. A test task that requires the processing of a great deal of non-verbal input in the form of graphs and charts, for example might be quite interactive, in that it involves the test taker’s metacognitive strategies and topical

knowledge. However, if it requires very little involvement of areas of language knowledge, it may not provide a valid measure of language knowledge.

Bachman and Palmer refer to Brindley (1994), who states that one of the assumptions made about task-centered assessment is that, since the communicative tasks used for assessment are based on authentic language use, they are automatically valid. This assumption has been criticised for a number of reasons.

First, there is a misinterpretation in terms of authenticity; that is, an assessment situation, although possibly employing authentic language, is still an artificial situation, not an authentic one.

Second, Fulcher states, there is the issue of sufficiency. Even though an assessment task may be authentic (reflecting a real-life situation), it does not mean that the sampling of language involved is of a sufficient amount for assessment purposes and that it can be generalised to other language use situations. In other words, a task may be authentic but may be impoverished in terms of what it reveals about the learner's language. Authenticity is not a panacea for validity. What Brindley suggests instead is moving away from loose authenticity/validity-based assertions to a more theory-based construction of validity.

Fulcher (1996) points to the artificiality in oral tests when he states, "Claims that test tasks replicate natural contexts and real-life situations which encourage natural language use remain the cornerstone of the claim to validity in oral test design."

In addition to the authenticity issue, Brindley (1994) points to another problem area in the validity of task-centered assessment: how to define the criteria for assessment, upon which student performances will be ranked or scored. According to Brindley, a variety of approaches has been tried in the past, but none of them has been particularly satisfactory.

Bachman and Palmer refer to a study of a "bottom up consultative approach" used in Australian primary and secondary ESL (English Second Language) education as reported by Griffin and McKay (1992: 20) who states:

Limitations of this approach include the difficulties involved in obtaining appropriate descriptions of language behaviour from practitioners. It is often

the case that practitioners' observations are limited by a lack of knowledge of theoretical models, by inadequate observation skills and/or an inability to articulate descriptions of independent student language behaviour. The developer of the scales has to make decisions about the need to use the imprecise language of the practitioner, and perhaps lose some of the definitive nature of the theoretical model, or to use a specialist terminology and run the risk of practitioner misinterpretation and rejection. (p. 78)

Brindley (1994) further points out that disagreement is not only between test developers and teachers. Even a group of teachers themselves may be lost in terms of assessment criteria:

Studies aimed at investigating how expert judgements are made, however, cast some doubt on the ability of expert judges to agree on any of these issues. Alderson and Lukmani (1989), for example, in an examination of item content in EFL (English Foreign Language) reading tests, found that judges were unable to agree not only on what particular items they were testing but also on the level of difficulty of items or skills and the assignment of these to a particular level. (pp. 78-79)

Brindley's third approach for defining assessment criteria is the genre-based approach, which has the slight advantage of being driven by theory:

One way of obtaining detailed assessment information at the level of the individual task is represented by genre-based approaches to assessment, which derive from the analysis of spoken and written genres within the framework of systematic-functional linguistic theory (Halliday, 1985). Within this approach, the genres (such as argument, describing a procedure etc.) are carefully described in terms of their structural organisation and linguistic features. These features are then used as the basis for the implementation of a teaching-learning cycle and serve as the criteria for assessment of overall task performance. (Brindley, 1994: 80).

The main problem with this approach, however, is according to Bachman and Palmer, that comprehensive descriptions of the different genres or structures are yet unavailable, making the development of such an approach quite difficult.

Measurement of many of these features also constitutes a tedious process that renders them unwieldy at best and highly unsuitable for classroom assessment.

Brindley's fourth and final approach calls for databased assessment criteria that are consistent with current theories of second language acquisition and use. Research into discourse analysis, consideration of processing dimensions such as analysis of linguistic knowledge and control of processing, and documenting real-world language use, for example, could produce better constructed, better justified, and more satisfying assessment tasks than previous approaches. Bachman and Palmer state that more research on validity clearly needs to be done, but at least Brindley seems to have illuminated the path for future research.

3.8 WHAT ARE THE FACTORS THAT AFFECT TASK-BASED ASSESSMENT PRACTICALITY?

Bachman and Palmer state that a test ought to be practical – within the means of financial limitations, time constraints, ease of administration, and scoring and interpretation. One important aspect of practicality that testing researchers have pointed out is that a test ought to have what Oller called instructional value, that is, “it ought to be possible to use the test to enhance the delivery of instruction in student populations”. Teachers need to be able to make clear and useful interpretations of test data in order to understand their students better.

According to Bachman and Palmer, Brindley (1994) points to a number of practical issues that task-centered assessment designers must face:

- 1) It is a time-consuming method of assessment because it involves eliciting, evaluating, and scoring student performances one at a time over a long period instead of all at once like paper-and-pencil tests. As a result, task-centered testing is necessarily more expensive than traditional forms of assessment.
- 2) A large amount of time must also be spent to train teachers so they can carry out task-centered assessment competently and confidently.
- 3) Public acceptability may also be a problem because many people might consider such an assessment system to be less than rigorous.

Bachman & Palmer (1996: 137) assert that practicality pertains primarily to the ways in which the test will be implemented, and, to a large degree, whether it will be developed and used at all. That is, for any given situation, if the resources required for implementing the test exceed the resources available, the test will be impractical, and will not be used unless resources can be allocated more effectively, or unless additional resources can be allocated. In designing a test, Bachman and Palmer state, we try to achieve the optimum balance among the qualities of reliability, construct validity, authenticity, interactiveness, and impact for our particular testing situation. In addition, we must determine the resources required to achieve this balance, in relationship to the resources that are available. Thus, Bachman and Palmer state, determining the practicality of a given test involves the consideration of (1) the resources that will be required to develop an operational test that has the balance of qualities we want, and (2) the allocation and management of the resources that are available. Bachman & Palmer can define practicality as the relationship between the resources that will be required in the design, development, and use of the test and the resources that will be available for these activities. Practicality is a matter of the extent to which the demands of the particular test specifications can be met within the limits of existing resources.

Bachman & Palmer (1996: 148) proposes that a test's practicality is a function of the amount of resources required and available during the different stages of the test development process. To evaluate practicality, we can ask the following questions:

- 1) What type and relative amounts of resources are required for: (a) the design stage, (b) the operationalization stage, and (c) the administration stage?
- 2) What resources will be available for carrying out (a), (b), and (c) above?

3.9 PRINCIPLES TO CREATING INTRINSICALLY MOTIVATING TESTS

Bachman and Palmer posit the following principles:

The principle of giving students advance preparation – This may sound simple, but much too often teachers do little to help students to prepare for a test. Tests, by their very nature, are anxiety-raising experiences. Students don't know what to expect. And they may not be aware of test-taking strategies that could help them. So, your

first task in creating intrinsically motivating tests is to be an ally in the preparation process. You can do the following: Provide information about the general format of a test and about types of items that will appear. Give students opportunities to practice certain item types. Encourage a thorough review of material to be covered. Offer advice on strategies for test preparation on strategies to use during the test itself. Give anxiety-lowering reassurance.

The principle of face validity – Sometimes students don't know what is being tested when they tackle a test. Sometimes they feel, for a variety of reasons, that a test isn't testing what it is 'supposed' to test. Face validity means that the students, as they perceive the test, feel that it is valid. You can help to foster that perception with: A carefully constructed, well thought-out format. Items that are clear and uncomplicated. Directions that are crystal clear. Tasks that are familiar and relate to their course work. A difficult level that is appropriate for your students. Test conditions that are *biased for best*, and therefore bring out students' best performance.

The principle of authenticity - Make sure that the language in your test is as natural and authentic as possible. Also, try to give language some context so that items aren't just a string of unrelated language samples. Thematic organization of items may help in this regard. Or, consider a storyline that may run through your items. Also, the tasks themselves need to be tasks that they have participated in and feel comfortable with. A classroom test is not the time to introduce brand new tasks because you won't know if student difficulty is a factor of the task itself or of the language you are testing.

The principle of washback – 'Washback' is, according to Bachman and Palmer, the benefit that tests offer to learning. When students take a test, they should be able, within a reasonably short period of time, to utilize the information about their competence that test feedback offers. Formal tests must therefore be learning devices through which students can receive a diagnosis of areas of strength and weakness. Their incorrect responses can become windows of insight about further work. Your prompt return of written tests with your feedback is therefore very important to intrinsic motivation. One way to enhance washback is to provide narrative evaluations of test performance. When you return a (written) test, give more back than a number of grade. Give praise for strengths as well as constructive

criticism of weaknesses. Give strategic hints on how a student might improve certain elements of performance. In other words, take some time to make the test performance an intrinsically motivating experience through which a student will feel a sense of accomplishment and challenge.

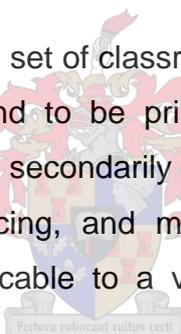
3.10 CURRENT APPROACHES TO MATERIALS DESIGN

McDonough and Shaw make the following distinctions:

Methodology - The study of pedagogical practices in general (including theoretical underpinnings and related research). Whatever considerations are involved in 'how to teach' are methodological.

Approach - Theoretical positions and beliefs about the nature of language the nature of language learning and the applicability of both to pedagogical settings.

Method – A generalized, prescribed set of classroom specifications for accomplishing linguistic objectives. Methods tend to be primarily concerned with teacher and student roles and behaviours, and secondarily with such features as linguistic and subject-matter objectives, sequencing, and materials. They are almost always thought of as being broadly applicable to a variety of audiences in a variety of contexts.



McDonough and Shaw furthermore make the following distinctions:

Curriculum / syllabus – Designs for carrying out a particular language program. Features include a primary concern with the specification of linguistic and subject-matter objectives, sequencing, and materials to meet the designs of a designated group of learners in a defined context.

Technique – Any of a wide variety of exercises, activities, or devices used in the language classroom for realizing lesson objectives.

Multi-syllabus – One can see multi-syllabus in terms of a merging of two broad approaches. One of these is concerned with a view of language use, and includes categories of function, context and language skill. The other is a version of a more formal linguistic syllabus, and is comprised of elements of grammar, pronunciation,

and vocabulary. What a multi-syllabus does is to build on a range of communicative criteria at the same time as acknowledging the need to provide systematic practice in the formal properties of language.

Lexical syllabus – It is typical of many current coursebooks to base teaching on an understanding of the psychological mechanisms whereby people learn and remember lexical items. Both the philosophy and the database of the dictionary have led to an approach to materials design that is usually termed the 'lexical syllabus'. Its central claim is that the lexical database provides 'a rich input to real language', thus giving authenticity and context to the tasks and exercises.

Process syllabus – The most common labels attached to this kind of syllabus design proposal are 'process', 'task based' and 'procedure'. Research focus upon the language learning process and the contributions of the learner to that process. It contains the far-reaching implication that syllabuses cannot be fully worked out in advance but must evolve as learners' problems and developing competence gradually emerge. Prabhu refers to this as a 'simple', not a 'sophisticated' syllabus.

McDonough and Shaw state that the profession had learned some profound lessons from our past wanderings. We had learned to be cautiously eclectic in making enlightened choices of teaching practices that were solidly grounded in the best of what we knew about second language learning and teaching. We needed to unify our approach to language teaching and designing effective tasks and techniques that are informed by that approach. We recognize that the complexity of language learners in multiple worldwide contexts demands an eclectic blend of tasks each tailored for a particular group of learners in a particular place, studying for particular purposes in a given amount of time. Your approach takes on great importance. Your approach to language teaching methodology is your theoretical rationale that underlies everything that you do in the classroom. Your approach should include most if not all of the principles. Your approach is inspired by the interconnection of all your reading and observing and discussing and teaching. Your approach is a dynamic composite of energies within you that changes with your experiences in your own learning and teaching. The interaction between your approach and your classroom practice is the key to dynamic teaching.

3.11 WHAT ARE THE STEPS INVOLVED IN DEVELOPING TASK-BASED ASSESSMENT?

Norris, Brown, Hudson, & Yoshioka (1998: 65) suggest that though it is not specifically designed for language teachers, *A Practical Guide to Alternative Assessment* by Herman, Aschbacher, and Winters (1992) offers a comprehensive guide for teachers and curriculum planners interested in developing a system of alternative assessment in the schools. It not only delineates the many important factors that must be considered in the development and decision-making process (i.e., instructional goals, task selection, rating scales, reliability, etc.) but also provides numerous samples of actual systems or rubrics that have been created and implemented, mainly in K-12 content classes. For the purposes of this review, chapters 4 (“Selecting Assessment Tasks”), 5 (“Setting Criteria”), and 6 (“Ensuring Reliable Scoring”) are of particular interest.

Herman, Aschbacher, and Winters (1992), referred to by Norris et al, lay out a five-step- plan for selecting assessment tasks:

- 1) Establish what the teacher’s specific instructional goals are because it is important that the chosen assessment task actually matches the instructional outcome(s) it is designed to measure.
- 2) Identify the specific, discipline-based content and skills that students are expected to attain and determine whether the task adequately represents or utilises them.
- 3) Insure that the task is fair and free of bias, allowing students to demonstrate their true progress and abilities without being disadvantaged by some extraneous element in the task lack of prior knowledge, unequal access to resources or materials, and so forth.
- 4) Decide which of the three possible forms the tasks will take (the choice should depend mainly on the type of skills and content that needs to be covered):
 - a. Authentic, real-world tasks (which have the advantage of generating greater motivation and offering greater transferability than traditional tasks),

- b. Interdisciplinary tasks (which economically combine and utilise the content and skills from various disciplines in performing the task, i.e., math, science, and reporting skills), or
 - c. Multi-dimensional tasks (which consist of a “mega-task” composed of smaller tasks that need to be tackled to complete the larger task).
- 5) Describe the assessment task so that others can understand and use it in other settings. Such a description should detail the intended outcomes, the content covered, the work and roles in the task, the materials and instructions involved, the rating system, and so on. Other areas of consideration in task selection include determining whether the task is teachable, feasible, credible, meaningful, and so on.

In addition to task selection and design, Herman, Aschbacher, and Winters (1992) discuss the development of scoring criteria (chapter 5) for evaluating student performance on a task, which function to:

- 1) Help teachers define excellence and plan how to help students achieve it
 - 2) Communicate to students what constitutes excellence and how to evaluate their own work
 - 3) Communicate goals and results to parents and others
 - 4) Help teachers or other raters be accurate, unbiased, and consistent in scoring
 - 5) Document the procedures used in making important judgments about students
- (p. 48)

According to Herman, Aschbacher, and Winters (1992), such scoring criteria will often have four elements:

- 1) Include dimensions used to judge the student’s work, which will generally express the qualities or characteristics essential in the performance of a given task. Such dimensions should reflect the instructional goals of the task and reflect what the teacher expects to see in terms of behaviour or characteristics if the task is done well, satisfactorily, or poorly;
- 2) Provide definitions, models, examples, or questions which will help to clarify more explicitly just what those dimensions are;

- 3) Offer a rating scale of some type, be it a yes-no checklist, a numerical scale, a qualitative scale (either descriptive or evaluative), or a combination numerical-qualitative scale, to help assess student performance of the dimensions; and
- 4) Standards of excellence (criterion-referenced, norm-referenced, or both) should be specified along with models or examples of each level.

This final step seems to bring together all the previous elements and completes the construction of a rating scale or rubric.

At this point, authors provide a guide to help teachers begin developing their own scoring criteria:

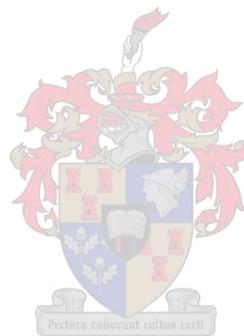
- 1) Investigate how the assessed discipline defines quality performance
- 2) Gather sample rubrics for assessing writing, speech, the arts, and so on as models to adapt for your purposes
- 3) Gather samples of students' and experts' work that demonstrate the range of performance from ineffective to very effective
- 4) Discuss with others the characteristics of these models that distinguish the effective ones from the ineffective ones
- 5) Write descriptors for the important characteristics
- 6) Gather another sample of students' work
- 7) Try out criteria to see if they help you make accurate judgments about students
- 8) Revise your criteria
- 9) Try it again until the rubric score captures the 'quality' of the work

Once scoring criteria have been developed, using the guidelines mentioned above, the last step according to Herman, Aschbacher, and Winters (1992) is to evaluate the scoring criteria. The authors suggest checking to make sure the scoring criteria are:

- 1) Keyed to important outcomes
- 2) Sensitive to purpose
- 3) Meaningful, clear, and credible

- 4) Fair and unbiased
- 5) Feasible and
- 6) Generalisable (pp. 76 – 79)

Although not covered in detail here, chapter 6 of Herman, Aschbacher, and Winters (1992) is also important because it pursues the next step, insuring reliable and consistent scoring through rater training and education, sample runs, and reliability studies performed on the actual tests.



CHAPTER FOUR

CONCEPTUALISATION IN THE DESIGN OF A TASK-BASED PERFORMANCE TEST FOR SESOTHO

4.1 INTRODUCTION

Robinson and Ross (1996) state:

“Task-based tests require candidates to perform an activity which simulates performances they will have to engage in outside the test situation, as, for example, in driving tests. Driving tests are an example of strictly “performance-referenced” task-based tests, in which the criterion of success or failure is ability to perform the task. Task-based tests can also be “system-referenced” if the task is used simply to elicit samples of the candidate’s linguistic knowledge or generalised verbal ability, as in oral proficiency interviews or multiple-choice reading comprehension tests based on texts adapted for testing purposes.”

4.2 PERFORMANCE REFERENCING AND TASK-BASED TEST DESIGN

According to Robinson and Ross (1996), system-referenced tests are “designed to evaluate language mastery as a psychological construct without specific reference to any particular use of it”. They evaluate, for example, whether a student can control certain tenses, understand certain low-frequency words, or can skim and scan accurately under specified time constraints.

Robinson and Ross (1996) suggest that the advantages of such tests are that they are generalisable to many different test samples and are relatively easy to construct and administer. Their disadvantage is that they seem artificial and lack face validity, since most language teaching programs aim to develop holistic procedural skill in the language of instruction. To the extent that system-references tests alone are used to assess the development of procedural skill they also lack construct validity. Since such tests can only tangentially evaluate the achievement of procedural and

communicative aims. Since they require the test developer to examine component skills in isolation from each other, or to determine the extent to the test-taker's access to fragmentary, piecemeal knowledge of the language system.

In contrast, Robinson and Ross state, greater face and construct validity are the main advantages of performance-referenced tests. Such tests involve near simulation of some future potential activity and results can be used to predict that candidate's ability to perform that or similar tasks in the future. Examples of these are reading tests which require candidates to give timetable information from actual train or bus timetable schedules, or, at the more advanced level, to find, and then take notes from, a cited journal article in a university library. The disadvantage of such tests are that they lack generalisability because they are developed with the goals and content domains of a particular test population's needs in mind: they are also difficult to administer since they involve simulating, with great attention to detail, the actual conditions of the potential future performance.

Another distinction according to Robinson and Ross (1996), which cuts across the system- versus performance-referenced distinction, is useful for identifying the components of our testing procedure- the distinction between direct and indirect test. In direct tests, the test procedure is equal to, or very similar to, the criterion or "target" procedure. Indirect tests, however, are abstractions, or incomplete versions of the target criterion procedure.

According to Robinson and Ross, direct system-referenced tests involve the elicitation of a language sample to demonstrate a skill, for example via oral interview or written composition, which is then examined or analysed with reference to its component parts, e.g. the grammar or vocabulary Robinson and Ross observes (1996). Reading tests that require candidates to read a text and answer comprehension questions that demonstrate knowledge of cause and effect relations, or to perform skilled operations like skimming or scanning, are examples of such tests. Indirect system-referenced tests require candidates to demonstrate knowledge of specific aspects of the system, e.g. via multiple-choice grammar questions.

Robinson and Ross (1996) propose that in indirect performance-referenced tests, the test performance (or what the candidate has to do during the test) exactly simulates

the criterion performance (or what the candidate would have to do in the 'real world' for which the course of instruction is preparing her/him).

In indirect performance-referenced tests, the criterion and test performances are not isomorphic. This is usually the result of breaking down the criterion performance into more manageable subtasks, or component steps, which are then examined separately.

4.3 DOUGLAS BROWN'S PRINCIPLES ON SECOND LANGUAGE LEARNING AND TEACHING

The following twelve important principles of second language teaching according to Douglas Brown, are central to language learning, these principles also need to be taken into account in test construction.

Cognitive principles

Automaticity – According to Brown, it is better to acquire language subconsciously – without overtly analyzing the forms of language. Barry McLaughlin called it automatic processing with peripheral attention to language forms. Automaticity includes the importance of: (1) subconscious absorption of language through meaningful use, (2) efficient and rapid movement away from a focus on the forms of language to the focus on the purposes to which language is put, (3) efficient and rapid movement away from a capacity-limited-control of a few bits and pieces to a relatively unlimited automatic mode of processing language forms and, (4) resistance to the temptation to analyse language forms. Efficient second language learning involves a timely movement of the control of a few language forms into the automatic processing of a relatively unlimited number of language forms. Overanalysing language, thinking too much about its forms, and consciously lingering on rules of language all tend to impede this graduation to automaticity.

Meaningful learning – Brown argues that meaningful learning implicates the following in the classroom. (1) Capitalize on the power of meaningful learning by appealing to students' interests, academic goals and career goals. (2) Whenever a new topic or concept is introduced, attempt to anchor it in students' existing

knowledge and background so that it gets associated with something they already know. (3) Avoid the pitfalls of rote learning: (a) too much grammar explanations, (b) too many abstract principles and theories, (c) too much drilling and / or memorization, (d) activities whose purposes are not clear, (e) activities that do not contribute to accomplishing the goals of lesson or unit or course, (f) techniques that are so mechanical or tricky that it get centered on the mechanics instead of the language or meanings. Meaningful learning will lead to better long-term retention than rote learning.

The anticipation of reward – There is, according to Brown, virtually nothing we do that is not inspired and driven by a sense of purpose or goal, and, according to Skinner, the anticipation of reward is the most powerful factor in directing one's behaviour. The following constructive classroom implications may be drawn: (1) Provide an optimal degree of immediate verbal praise and encouragement to students as a form of short term reward. (2) Encourage students to reward each other with compliments and supportive action. (3) Short term reminders of progress may help students to perceive their development (for example, progress charts and graphs...). (4) Display enthusiasm and excitement yourself in the classroom. (5) Try to get learners to see the long-term rewards in learning Sesotho (by pointing things out to them).



The intrinsic motivation principle – The most powerful rewards, according to Brown, are those that are intrinsically motivated within the learner. Because the behaviour stems from needs, wants, desires within oneself, the behaviour itself is self-rewarding; therefore, no externally administered reward is necessary at all. Teachers can be wonderful if they design classroom tasks that feed into the students' intrinsic drives.

Strategic investment – Brown states that successful mastery of the second language will be due to a large extent to a learner's own personal 'investment' of time, effort, and attention to the second language in the form of an individualized battery of strategies for comprehending and producing the language. Two major pedagogical implications of the principle: (1) the importance of recognizing and dealing with the wide variety of styles and strategies that learners successfully bring

to the learning process, and, therefore (2) the need for attention to each separate individual in the classroom.

Affective principles

Language ego – Brown explains that, as human beings learn to use a second language, they also develop a new mode of thinking, feeling and acting – a second identity. The new ‘language ego’, intertwined with the second language, can easily create within the learner a sense of fragility, a defensiveness, and a raising of inhibitions. All second language learners need to be treated with affective tender loving care (patience, empathy, understanding).

Self-confidence – The eventual success, according to Brown, that learners attain in a task is at least partially a factor of their belief that they indeed are fully capable of accomplishing the task. At heart of all learning, according to Brown, is the condition that a person believes in his or her own ability to accomplish the task. Classroom applications: (1) Give ample verbal and non-verbal assurance to students. It helps a student to hear a teacher affirm a belief in the student’s ability. (2) Sequence techniques from easier to more difficult.

Risk-taking – Brown argues that successful language learners, in their realistic appraisal of themselves as vulnerable beings yet capable of accomplishing tasks, must be willing to become ‘gamblers’ in the game of language, to attempt to produce and to interpret language that is a bit beyond their absolute certainty. It is important to get learners to take calculated risks in attempting to use language.

The language-culture connection – Whenever you teach a language, Brown maintains, you also teach a complex system of cultural customs, values, and ways of thinking, feeling, and acting. Especially in ‘second’ language learning contexts, the success with which learners adapt to a new cultural milieu will affect their language acquisition success, and vis-à-vis, in some possibly significant ways.

Linguistic principles

The native language effect – Brown argues that the native language of learners will be a highly significant system on which learners will rely to predict the target language system. While that native system will exercise both facilitating and interfering effects on the production and comprehension of the new language, the interfering effects are likely to be the most salient.

Interlanguage – Brown maintains that second language learners tend to go through a systematic or quasisystematic developmental process as they progress to full competence in the target language. Successful interlanguage language development is partially a factor of utilizing feedback from others.

Communicative competence – Given that communicative competence is the goal of a language classroom, Brown states, the instruction needs to point toward all of its components: organizational, pragmatic, strategic, and psychomotor. Communicative goals are best achieved by giving due attention to language use and not just usage, to fluency and not just accuracy, to authentic language and contexts, and to students' eventual need to apply classroom learning to heretofore unrehearsed contexts in the real world.

Learner-centered teaching – According to Brown, this term applies to curricula as well as to specific techniques. Learner-centered instruction includes: (a) techniques that focus on or account for learner's needs, styles and goals, (b) techniques that give some control to the student, (c) curricula that include the consultation and input of students and that do not presuppose objectives in advance, (d) techniques that allow for student creativity and innovation, techniques that enhance a student's sense of competence and self-worth.

Cooperative learning – Brown maintains that a curriculum or classroom that is cooperative – and therefore not competitive – usually involves the above learner-centered characteristics. An added connotation to the term 'cooperative', however, is its emphasis on collaborative efforts of students and teachers working together to pursue goals and objectives. Collaboration may be among students or the student and teacher.

Interactive learning – According to Brown, the communicative purpose of language compels us to create opportunities for genuine interaction in the classroom.

Interactive classes will most likely be found: (a) doing a significant amount of pair work and group work, (b) receiving authentic language input in real-world contexts, (c) producing language of genuine, meaningful communication, (d) performing classroom tasks that prepare them for actual language use 'out there', (e) practising oral communication through the give and take and spontaneity of actual conversations, (f) writing to and for audiences, not contrived ones.

Teaching children and adults a second language

According to Brown, the difference between children and adults lies primarily in the contrast between the child's spontaneous, peripheral attention to language forms and the adult's overt, focal awareness and attention to those forms. Adults have superior cognitive abilities that can render them a bit more successful in certain classroom endeavours.

Intellectual development – Brown argues that rules, explanations, and other even slightly abstract talk about language must be approached with extreme caution. Children are centered on the 'here and now', on the functional purposes of language. They have little appreciation for our adult notions of 'correctness', and they certainly cannot grasp the metalanguage we use to describe and explain linguistic concepts. Don't explain grammar using 'big' terms. Rules that are stated in abstract terms should be avoided. Some grammatical concepts can be called to learner's attention by showing them certain patterns and examples. Certain more difficult concepts or patterns require more repetition than adults need. Children must understand the meaning and relevance of repetitions. Adults are more readily able to handle abstract rules and concepts. Too much abstract generalization about usage and not enough real-live language use can be deadly for adults, too.

Attention span – Brown maintains that one of the most salient differences between adults and children is attention span. Language lessons must be interesting, lively and fun to get a larger attention span. Because children are focused on the immediate here and now, activities should be designed to capture their immediate interest. A lesson needs to have a variety of activities to keep interest and attention alive. A teacher needs to be animated, lively and enthusiastic about the subject matter. While you may think that you are overdoing it, children need this

exaggeration to keep spirits buoyed and minds alert. A sense of humour will go a long way to keep children laughing and learning. Since children's humour is quite different from adults', remember to put yourself in their shoes. Children have a lot of natural curiosity – use it. Adults have longer attention spans for material that may not be intrinsically interesting to them. However, the rule of keeping your activities brief applies to adult age teaching.

Affective factors – According to Brown, the ego's of children are still being shaped, and therefore the slightest nuances of communication can be negatively interpreted. Teachers need to help them to overcome such potential barriers to learning. Help your students to laugh with each other at various mistakes that they all make. Be patient and supportive to build self-esteem. Be firm in your expectations of students. Elicit as oral participation as possible from students. We should never underestimate the emotional factors that may be attendant to adult second language learning. Their level of shyness can be equal to or greater than children.

Authentic, meaningful language – Brown argues that children are focussed on what this new language can actually be used for here and now – language that hold immediate rewards for them. Children are good at sensing language that is not authentic. Language needs to be firmly context embedded. Storylines, familiar situations and characters, real-life conversations are meaningful purposes to improve attention and retention. A whole language approach is essential. Do not break up language into too many bits and pieces or students will not see the relationship to the whole. In addition, stress the interrelationships among the various skills otherwise, they will not see the important connections. Adults, with their more developed abstract thinking ability, are better able to take a context-reduced segment of language and understand it. Authenticity and meaningfulness are of course still highly important.

Students' cognitive learning processes

Beginning – Brown argues that you can expect in plenty of repetition of a limited number of words, phrases, and sentences. In the first few days of class, you can coax your students into some peripheral processing by getting them to use practised language for genuinely meaningful purposes.

Intermediate – At the intermediate stage, Brown maintains, some automatic processing has taken hold. Phrases, sentences, structures and conversational rules have been practiced and are increasing in number, forcing the mental processes to automatize. One of your principle goals at this level is to get students to continue to automatize, to continue to allow bits and pieces of language that might clutter the mind to be relegated to automaticity.

Advanced – Competence in language continues to build. Students can realize the full spectrum of processing, assigning larger and larger chunks to automatic modes and gaining the confidence. Therefore, your task at this level is to assist in that ongoing attempt to automatize language and in that delicate interplay between focal and peripheral attention to selected aspects of language.

The role of the teacher

Beginning – Brown states that beginning students are highly dependent on the teacher for models of language, and so a teacher-centered and teacher-fronted classroom is appropriate for some of your classroom time. In a second language context where instruction is carried out in the target language (Sesotho), virtually all of your class time will be teacher-controlled. In a foreign language situation, where your students speak the same native language, some negotiation might be possible in the native language, allowing for a small amount of student control.

Intermediate – According to Brown, students should be encouraged to ask questions, make comments, and negotiate certain options in learning where appropriate. More student-student interaction can now take place in pairs, small groups, and whole class activity. Learner-centered work is now possible for more sustained lengths of time, as students' language is able to maintain topics of discussion and focus.

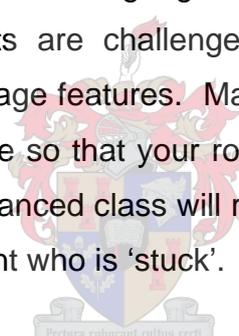
Advanced – You can sit back and let the students' questions and self-generated curiosity take over. The same independence that students have acquired must be cleverly channelled into classroom routines that benefit most of the students most of the time. A direct role on your part can create effective learning opportunities even within a predominantly learner-centered classroom.

Teacher talk

Beginning – Your input in the class is crucial. Your own Sesotho needs to be very clearly articulated. It is appropriate to slow speech somewhat for easier student comprehension. You do not need to talk any louder to beginners than to advanced students. Use simple vocabulary and structures that are at or just slightly beyond their level.

Intermediate – According to Brown, most of your oral production can be sustained at a natural pace, as long as articulation is clear. Teacher talk should not occupy the major proportion of a class hour; otherwise, you are probably not giving students enough opportunity to talk. You should be using less of the native language of the learners at this level, but some situations may still demand it.

Advanced – Brown argues that natural language at natural speed is necessary at this level. Make sure your students are challenged by your choice of vocabulary, structures, idiom and other language features. Make sure your students have ample opportunities to produce language so that your role as a provider of feedback takes prominence. A teacher of an advanced class will resort to a word or two in the native language in order to help a student who is ‘stuck’.



Authenticity of language

Beginning – Brown argues that the language that you expose your students to should be authentic language, not stilted just because students are beginners. Simple greetings and introductions are authentic and yet manageable. Make sure utterances are limited to short, simple phrases. At times, such language may appear to be artificial because of all the repetition needed at this stage.

Intermediate – At this level, Brown states, students sometimes are overly concerned about grammatical correctness and may want to wonder into esoteric discussions of grammatical details. This penchant for analysis might get them too far afield from authentic, real language. Make sure they stay on the track.

Advanced – According to Brown, everything from academic prose to literature to idiomatic conversation becomes a legitimate resource for the classroom. Virtually no

authentic language material ought to be summarily disqualified at this stage. Certain restrictions may come to bear, depending on how advanced your class is, of course.

Fluency and accuracy

Beginning – The ‘flow’ of language is, according to Brown, important to establish, from the beginning in reasonably short segments of language. Attention to accuracy should center on the particular grammatical, phonological or discourse elements that are being practiced. In teaching speaking skills, it is extremely important at this stage that you are very sensitive to students’ need to practice freely and openly without fear of being corrected at every minor flaw. You need to give some treatment of selected grammatical and phonological errors.

Intermediate – The dichotomy between fluency and accuracy is, according to Brown, a crucial concern here, more so than at either of the other ends of the proficiency spectrum. Some students are likely to become overly concerned about accuracy and others may slide into a self-satisfied rut in which they actually become quite fluent but in which they become very difficult to comprehend. Fluency exercises are necessary at this level. A big part of your task with most students is to maintain their ‘flow’ with just enough attention to error to keep them growing.

Advanced – At this level, Brown states, most, if not all, of your students are ‘fluent’ in that they have passed beyond that ‘breakthrough’ stage where they are no longer thinking about every word or structure they are producing or comprehending. At issue is a handful or two of problems that need attention. If errors are relatively rare, an occasional treatment from you or from peers may be quite helpful.

Techniques

Brown proposes the following techniques for learning at various levels:

Beginning – Short, simple techniques must be used. Some mechanical techniques are appropriate – choral repetition and other drilling. A good many teacher-initiated questions dominate at this level. Group and pair activities are excellent techniques

as long as they are structured and very clearly defined with specific objectives. A variety of techniques is important because of limited language capacity.

Intermediate – Because of the increasing language capacities of your students, techniques can increase in complexity. Common interactive techniques for intermediates include chain stories, surveys and polls, paired interviews, group problem solving, role-plays, storytelling and many others.

Advanced – Techniques can now tap into a full range of socio-linguistic and pragmatic competencies. Typical of this level are activities like group debates etc. Often at this, level students have specific purposes for which they are planning to use Sesotho. Focus on those purposes as much as possible.

Listening and speaking goals

Brown posits the following goals for different levels of learners:

Beginning – Listening and speaking for beginners are meaningful and authentic communication tasks. They are limited more by grammar, vocabulary, and length of utterance than by communicative function. It is surprising how many language functions can be achieved with much uncomplicated language.

Intermediate – The linguistic complexity of communicative listening – speaking goals increases steadily. Along with the creation of novel utterances, students can participate in short conversations. The functions themselves may not be intrinsically more ‘complex’, but the forms they use are.

Advanced – At this level students can focus more carefully on all the socio-linguistic nuances of language. Pragmatic constraints are common areas needing work as students finely tune their production and comprehension.

Advantages of groupwork in the second language classroom

Brown proposes the following advantages of group work:

Group work generates interactive language – Small groups provide opportunities for: student initiation, for face to face give and take, for practice in negotiation of

meaning, for extended conversational exchanges, and for student adoption of roles that would otherwise be impossible.

Group work offers an embracing affective climate – in small groups, reticent students become vocal participants. The small group becomes a community of learners cooperating with each other in pursuit of common goals. Small group work is an increase in student motivation.

Group work promotes learner responsibility and autonomy – Group work places responsibility for action and progress upon each of the members of the group somewhat equally.

Group work is a step toward individualizing instruction – Small groups can help students with varying abilities to accomplish separate goals.

Main types of tasks for group work

Brown proposes the following types of group work:

Games – A game could be anything that formalizes a technique into units that can be scored in some way. Several of the other group tasks outlined below could thus become ‘games’. Guessing games are common language classroom activities.

Role-play – minimally involves (a) giving a role to one or more members of a group and (b) assigning an objective or purpose that participants must accomplish.

Simulations – usually involve a more complex structure and often-larger groups (of 6 – 20), where the entire group is working through an imaginary situation as a social unit, the object of which is to solve some specific problem.

Drama – Drama is a more formalized form of role-play or simulation, with a preplanned story line and script. Sometimes small groups may prepare their own short dramatization of some event, writing the script and rehearsing the scene as a group.

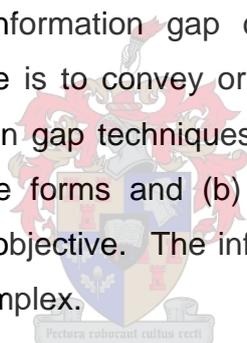
Projects – For learners of all ages, but perhaps especially for younger learners who can greatly benefit from hands-on approaches to language, certain projects can be

rewarding indeed. If you were to adopt an environmental awareness theme in your class, for example, various small groups could each be doing different things.

Interview – A popular activity for pair work, but also appropriate for group work, interviews are useful at all levels of proficiency. At the lower levels, interviews can be structured, in terms of both the information that is sought and the grammatical difficulty and variety. At the higher levels, interviews can probe more complex facts, opinions, ideas, and feelings.

Brainstorming – Brainstorming is a technique whose purpose is to initiate some sort of thinking process. It gets students' "creative juices" flowing without necessarily focussing on specific problems or decisions or values. Brainstorming involves students in a rapid-fire, free-association listing of concepts or ideas or facts or feelings relevant to some topic or context.

Information gap – The term information gap covers a tremendous variety of techniques in which the objective is to convey or to request information. The two focal characteristics of information gap techniques are (a) their primary attention to information and not to language forms and (b) the necessity of communicative interaction in order to reach the objective. The information that students must seek can range from very simple to complex.



Jigsaw – Jigsaw techniques are a special form of information gap in which each member of a group is given some specific information and the goal is to pool all information to achieve some objective.

Problem solving – Problem solving group techniques focus on the group's solution of a specified problem. They might or might not involve jigsaw characteristics, and the problem itself might be relatively simple to quite complex.

Decision making – Decision making techniques are simply one kind of problem solving where the ultimate goal is for students to make a decision. Some of the problem solving techniques does not involve a decision about what to do and others do involve such questions.

Opinion exchange - An opinion is usually a belief or feeling that might not be founded on empirical data or that others could plausibly take issue with. Opinions are difficult

for students to deal with at the beginning levels of proficiency, but by the intermediate level, certain techniques can effectively include the exchange of various opinions.

Various techniques for planning group work

Littlewood draws the following distinctions for group work:

Introduce the technique – The introduction may simply be a brief explanation. The introduction usually should include a statement of the ultimate purpose so that students can apply all the other directions to that objective.

Justify the use of small groups for the technique – Remind your students that they will get an opportunity to practice certain language forms or functions, and that if they are reluctant to speak up in front of the whole class, now is their chance to do so in the security of a small group.

Model the technique – For a new and potentially complex task, it never hurts to be too explicit in making sure students know what they are supposed to do. After students get into their groups, you might for example, show them what you exactly want.

Give explicit detailed instructions – Give the students specific instructions on what they are to do, a restatement of the purpose, rules they are to follow, establish a time frame and assign roles.

Divide the class into groups – To ensure participation or control you may want to preassign groups in order to account for one or two of the following: native language, proficiency levels, age or gender differences, culture or subcultural group, personality types, cognitive style preferences, cognitive / developmental stages, interests, prior learning experience or target language goals.

Check for clarification – Before students start moving into their groups, check to make sure they all understand their assignment. Test out certain elements of your lead-in by asking questions like, 'Keiso, explain the purpose of this activity'.

Set the task in motion – This part should now simply be a matter of saying something like ‘Okay, get into your groups and get started right away on your task’. Some facilitation may be necessary to ensure smooth logistics.

The main considerations relevant to the learning of a second language as a skill

Littkewood elaborates as follows on language learning as skills learning:

Cognitive habits

There was a heated debate between supporters of the ‘behaviourist’ and ‘cognitive’ approach. John Carroll summed up a compromise with the term ‘cognitive habit-formation’: yes, the basis for using language creatively must be a system of mental rules: but yes, too, if a person wishes to use the language fluently in communication situations, these rules have to be applied automatically, like a set of mental habits. Language learning involves developing a set of habits (i.e. automatized skills), but these skills have their basis in the mind.

Performing a skill

This compromise view coincides nicely with accounts of other kinds of skilled performance, where it is generally agreed that both cognitive and behavioural aspects are involved. Underlying the effective performance of any skill is the ability to form ‘cognitive plans’ which direct the behaviour itself. These plans operate at different levels but are integrated when performance takes place. Speech pass through four ‘levels of representation’ before it is actually produced as sound.

Adaptability and automaticity

An important characteristic of the plans that produce skilled performance is that they can be adopted to changing circumstances. A second important feature of many of the plans, which underlie skilled performance is that once they are selected, they operate automatically. Adaptability and automaticity are connected with each other. So long as the lower-level plans operate automatically, we can devote maximum

attention to controlling performance by forming effective higher-level plans. We can adopt while the performance is in progress.

Learning to perform a skill

There are three main aspects, which belong to the learning of a skill:

1. Learners have to become aware of the key features of the target performance, so that they can create the mental plans, which are necessary for producing it themselves.
2. They have to practice converting these plans into actual behaviour, so that in due course the lower-level plans can operate automatically, in response to higher-level decisions.
3. They must learn to start form a higher-level plan and select lower-level plans, which are appropriate for carrying it out.

Failures in performance

1. The most obvious failure in performance occurs when a learner conceives a communicative purpose for which his or her language repertoire simple does not contain lower-level plans needed for carrying it out efficiently.
2. A further source of weakness is when a learner's repertoire contains the necessary lower-level plans to carry out a desired communicative intention but these plans are not yet sufficiently automated to operate without conscious attention and effort.
3. A learner may have a large number of individual automated plans in his or her repertoire. However, he or she has not had sufficient whole-task experience to be able to select and adopt the plans appropriately to particular communicative purposes as these arise during interaction.

Learning language as a natural process

Littlewood outlines elements of language learning as a natural process as follows:

Creative construction in first language learning

Littlewood states that many psychologists and linguists come o challenge the behaviourist view of language learning as 'habit-formation' and emphasized that learning must consist in developing rules for creating and understanding new utterances. These studies let to the view that natural language learning is a process of 'creative construction'. On the basis of what they hear children 'construct' rules of there own to account for the regularities in the language. Littlewood states that these rules constitute the children's internal 'grammar', which begins as very simple systems but gradually develops until it is the same as the grammar of adult speakers of the language. When we look at language in terms of developing, it is no longer appropriate to talk about 'errors' in the normal sense. Children produce a large number of forms, which do not correspond to the forms than an adult would produce. However, they are a reflection of the children's grammatical system at that point in time rather than simply imperfect attempts to match the adult's system. This view that errors are a natural element in the learning process has had considerable influence in the field of foreign-language learning.

Creative construction by second language learning

Littlewood discusses the notion of creative construction in second language learning as follows:



Many similarities have been found between the sequences involved in natural second language learning and those found in first-language acquisition. The main extra dimension is that second-language learners sometimes transfer rules from their first language. Foreign-language learners like first-language learners, construct rules creatively rather than simply modelling the forms that are presented to them. They also show how these natural processes of rule-construction are often more powerful than the learning processes those teachers tried to control.

Conditions for natural learning.

Two of the essential conditions for natural language learning relate to the provision of adequate learning opportunities: exposure to the language and interaction with other people. Two are connected with motivation: a need to use the language in real communication and a more generalized desire to make contact with the people who

speak the language. The last condition concerns the extent to which the learner's emotional state enables him or her to be open to the new learning experiences.

Integrating skill-learning and natural learning.

Littlewood views the integration of skill learning and natural learning as follows:

Should a skill-learning and natural learning model exclude one another?

Both the skill-learning model and the natural-learning have thus provided the foundations for teaching approaches, which have been successful in the classroom. However, neither has led to the almost universal success, which we ought to be able to achieve. Each kind of learning has its own useful contribution to make in the classroom and we should therefore look for ways of integrating them within a broader framework.

Integrating the two kinds of learning

Conscious learning provides rules, which can at first be applied only with conscious effort and which, therefore, the learner needs to automate by means of practice.

Learning and acquisition – Littlewood argues that this model solves the question by if we possess two sets of language-learning mechanisms, which operate separately from each other. One set of mechanisms enables us to learn in a conscious way: to learn rules, memorize vocabulary, benefit from drills etc. A different set enables us to learn in a natural AUBCINSCIOUS WAY. Krashen uses the term 'learning' for our conscious processes and 'acquisition' for our subconscious processes. Subconscious acquisition provides rules, which are already automated when they enter the learner's communication system.

Learning and communication – According to Littlewood, there are different ways in which has been learnt – whether consciously or subconsciously – can contribute to communication. When learners communicate, they have different kinds of rule at their disposal: (1) Rules, which have emerged in an already automated form, because of natural acquisition. (2) Non-automated rules, which have recently been learnt and can only be used accurately in favourable situations. (3) Rules, which

were initially learnt consciously but have become automated as a result of frequent practice and use.

Conscious learning as reinforcement for natural learning – Littlewood refers to Piennemann, who argues that a new language structure is only ‘teachable’ through explicit instruction if it is taught at a time when the learner’s natural processing mechanisms are ready to receive it. Piennemann’s ‘teachability hypothesis’ does, however, encourage us to follow a path which common sense would also suggest: to try to organize courses in such a way that formal practice activities support, rather than contradict, the sequence of learning that comes naturally to the learners.

The implications for learning and teaching

Language learning as a skill-learning process: The main input to this process, according to Littlewood, comes through organized instruction in which we pre-select the items that we ask the learners to internalise.

Language learning as a natural process: The methodological implications of this model are in principle more straightforward than those of the skill-learning model. The task of the teacher is simply to create environments in which learners can communicate through the language and obtain the kind of input that is needed for natural learning to take place. The natural-learning model thus serves to highlight the importance of communication in the overall framework.

Specific Purpose Sesotho for personnel in Small Business Corporation.

The national language policy of South Africa is that there are 11 official languages. The national policy plays an important role because African languages are now widely introduced in schools – especially the African language of the region, broadcasting on TV, radio etc.

The fact that the African languages are official languages should have the consequence that people have a high incentive to learn these languages as second language since this can be a direct benefit to future career prospects – multilingualism in South Africa.

Subtractive bilingualism refers to a native language if it is considered detrimental to the learning of the second language. Additive bilingualism occurs when the community or society holds the native language in prestige.

In South Africa, the learning of an African language by English and Afrikaans speakers is held in high esteem for social as well as professional purposes. The acceptance of the multilingual nature of South Africa resulted in the view that competence in African languages (by non-first language speakers of African languages) is strongly recommended and valued in the social as well as professional arenas.

Sesotho is more spoken in the Gauteng area and in the Free State region. In Sesotho, it is not often found that a region has a dialect. In the regions where Sesotho are spoken it is a lot used in everyday speaking especially by the mother tongue speakers who are proud of their African language. One thing that you do find in Sesotho is that it is spoken in a slightly different way than it is written.

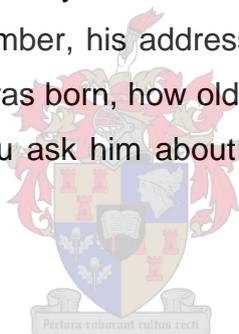


Task 1:**Sesotho:**

Kajeno o bua le monghadi mme o mmotsa dipotso tse ngata. O qala ho mmotsa lebitso, fane, diinisheale, o dula kae mme o tswa kae. Hape o mmotsa nomoro ya lebokoso la poso, Nomoro ya mohala, aterese, nomoro ya boitsebiso le nomoro ya paseporoto. O botsa na o hlahetse kae, o dilemo tse kae, o hlahetse neng mmeo mmotsa na o nyetswe le kapa o na le bana. O mmotsa ka boahi, kereke, sekolo, yunibesiti, kgato, mosebetsi le testamente.

English:

Today you talk to a sir and you ask him a lot of questions. You start by asking his name, surname, initials, where he stay and where he comes from. You ask him his postal address, his telephone number, his address, his ID number and his passport number too. You ask where he was born, how old he is, when he was born and if he is married and has children. You ask him about his nationality, church, university, degree, work and testament.

***Narrative:***

There may be narrative in this dialogue with every question asked. With every question asked the person might feel the need to tell you a short story to answer your question, for example of something that might have happened, or why the person wants to do something in a particular way, or if the person might want to explain something to you, or you might find that you have common interests and discuss that etc.

Opinion exchange:

An opinion is usually a belief or feeling that might not be founded on emperical data or that others could plausibly take issue with. Opinions are difficult for students to deal with at the beginning levels of proficiency, but by the intermediate level, certain techniques can effectively include the exchange of various opinions.

One warning: You play an important and sensitive role when you ask students to discuss their beliefs. Some beliefs are deeply ingrained from childhood rearing or from religious training, among other factors. So, it is easy for a student to be offended by what another student says. In such exchanges, do everything you can to assure everyone in your class that, while there may be disagreement on issues, all opinions are to be valued, not scorned, to be respected, and not ridiculed.

These questions might lead to an answer based on belief / feeling:

- How are you? (Wena o kae?)
- What is your surname?
- Where do you live?
- Where do you come from?
- Where were you born?
- How old are you?
- Are you married?
- Do you have children?
- What is your nationality?
- What church do you go to?
- To which school did you go?
- To which university did you go?
- Where do you work?
- Do you have a testament?



Information gap:

The term information gap covers a tremendous variety of techniques in which the objective is to convey or to request information. The two focal characteristics of information gap techniques are (1) their primary attention to information and not to language forms, and (2) the necessity of communicative interaction in order to reach the objective. The information that students must seek can range from very simple to complex.

Jigsaw:

Jigsaw techniques are a special form of information gap in which, each member of the group is given some specific information and the goal is to pool all information to achieve some objective.

Information gap and jigsaw can be represented in the following questions:

- How are you? (Wena o kae?)
- What is your name?
- What is your surname?
- What are your initials?
- Where do you live?
- Where do you come from?
- What is your postal address?
- What is your telephone number?
- What is your address?
- What is your ID number?
- What is your passport number?
- Where were you born?
- How old are you?
- When were you born?
- Are you married?
- Do you have children?
- What is your nationality?
- What church do you go to?
- To which school did you go?
- What standard did you pass at school?
- To which university did you go?
- Which degree did you finish at University?
- What is your profession?
- Where do you work?
- Do you have a testament?



Dialogue:

- You: Dumela, monghadi.
- English: Good day, sir (mister).
- Client: Ee, dumela ntate. Wena o kae?
- English: Yes, good day man. How are you? (Where are you?)
- You: Nna ke teng. Ha ke tsebe wena?
- English: I am fine. I do not know about you?
- Client: Nna ke teng.
- English: I am fine.
- You: Lebitso la hao ke mang, monghadi?
- English: What is your name, sir (mister)?
- Client: Lebitso la ka ke Willie.
- English: My name is Willie.
- You: Willie, fane ya hao ke mang?
- English: Willie, what is your surname?
- Client: Fane ya ka ke Lombaard.
- English: My surname is Lombaard.
- You: Diinisheale tsa hao ke mang?
- English: What are your initials?
- Client: Diinisheale tsa ka ke W. S.
- English: My initials are W. S.
- You: O dula kae, Willie?
- English: Where do you live, Willie?
- Client: Ke dula ka Wepener.
- English: I live in Wepener.
- You: O tswa kae?
- English: Where do you come from?
- Client: Ke tswa Foreisetata.
- English: I come from the Free State.
- You: Nomoro ya lebokoso la poso ke mang?
- English: What is your postal address?
- Client: Nomoro ya lebokoso la poso ke 63, Wepener, 9944.

- English: My postal address is 63, Wepener, 9944.
- You: Nomoro ya mohala ke mang?
- English: What is the telephone number?
- Client: Nomoro ya mohala ke 051-5831484/0/1.
- English: The telephone number is 051-5831484/0/1.
- You: Aterese ya hao ke mang?
- English: What is your address?
- Client: Aterese ya ka ke 8 van Riet seterata, Wepener.
- English: My address is 8 van Riet Street, Wepener.
- You: Nomoro ya boitsebiso ke mang?
- English: What is your ID number?
- Client: Nomoro ya boitsebiso ke 7108042954056.
- English: My ID number is 7108042954056.
- You: Nomoro ya paseporoto ke mang?
- English: What is the passport number?
- Client: Nomoro ya paseporoto ke BL 0137285.
- English: The passport number is BL 0137285.
- You: O hlaetse kae?
- English: Where were you born?
- Client: Ke hlaetse Rouxville.
- English: I was born in Rouxville.
- You: O dilemo di kae?
- English: How old are you?
- Client: Ke dilemo tse mashome a mabedi le metso e tsheletseng.
- English: I am 26.
- You: O hlaetse neng?
- English: When were you born?
- Client: Ke hlaetse ka 1971/08/04.
- English: I was born in 1971/08/04.
- You: O nyetswe na?
- English: Are you married?
- Client: Ee, ke nyetswe.
- English: Yes, I am married.
- You: O na le bana?

- English: Do you have children?
 Client: Tjhee, ha ke na bana.
 English: No, I do not have children.
 You: Boahi ba hao ke mang?
 English: What is your nationality?
 Client: Boahi ba ka ke Leburu.
 English: My nationality is a South African.
 You: O ya kerekeng efe?
 English: Which church do you go to?
 Client: Ke ya kereke ya Fora.
 English: I go to the Dutch Reformed Church.
 You: O kenile sekolo sefe?
 English: To which school did you go?
 Client: Ke kenile sekolo sa Rouxville.
 English: I went to the school of Rouxville.
 You: O qetile sehlopha sefe sekolong?
 English: What standard did you pass at school?
 Client: Ke qetile sehlopha se leshome sekolong?
 English: I passed matric at school.
 You: O ile yunibesiting efe?
 English: To which university did you go?
 Client: Ke ile yunibesiti ya Foreisetata.
 English: I went to the university of the Free State.
 You: O qetile kgato efe yunibesiting?
 English: Which degree did you finish at University?
 Client: Ke qetile kgato ya thuta ya tshwaro ya dibuka.
 English: I finished my degree in bookkeeping.
 You: Mosebetsi wa hao ke eng?
 English: What is your profession?
 Client: Mosebetsi wa ka ke ho rupela.
 English: My profession is to teach.
 You: O sebetsa kae ?
 English: Where do you work?
 Client: Ke sebetsa ka sekolong.

English: I work at the school.
 You: O na le testamente ?
 English: Do you have a testament ?
 Client: Ee, ke na le tsetamente.
 English: Yes, I have a testament.
 You: Ke lebohile ka nako, monghadi.
 English: Thank you for the time, sir (mister).
 Client: Ke lebohile, ntate.
 English: Thank you, man.
 You: Ke tla o bona beke e tlang ka Labone.
 English: I will see you next week on Thursday.
 Client: Ho lokile, ntate.
 English: It's fine, man.
 You: Tsamaya ka kgotso, monghadi.
 English: Go in peace, sir (mister).
 Client: Sala ka kgotso, ntate.
 English: Stay in peace, man.

Task 2:

Sesotho:

Kajeno ke bua le motswalle. O mmotsa dipotso ka kgwebo. O mmotsa lebitso, nomoro ya lebokoso la poso, nomoro ya mohala, aterese le banka ya kwebo. Hape o mmotsa ka batho ya kgwebo, jwale ka motsamaisi, mohlalobi wa dibuka, morekisi wa dinshoreense, modula-setulo, basebetsi-mmoho le ramotlakase. Ka qetello o mmotsa ka feme, setsha, mohaho le lakansense la kgwebo.

English:

Today you talk to a friend. You ask him questions about the business. You ask him the name, postal address, telephone numbers, address, and the bank of the business. You ask him about the people of the business too, like the manager, accountant, insurance agent, chairperson, colleagues, and the electrician. The last



questions are about the factory, the piece of ground, the building, and the licence of the business.

Narrative:

There may be narrative in this dialogue with every question asked. With every question asked the person might feel the need to tell you a short story to answer your question, for example of something that might have happened, or why the person wants to do something in a particular way, or if the person might want to explain something to you, or you might find that you have common interests and discuss that etc.

Opinion exchange:

An opinion is usually a belief or feeling that might not be founded on empirical data or that others could plausibly take issue with. Opinions are difficult for students to deal with at the beginning levels of proficiency, but by the intermediate level, certain techniques can effectively include the exchange of various opinions.

One warning: You play an important and sensitive role when you ask students to discuss their beliefs. Some beliefs are deeply ingrained from childhood rearing or from religious training, among other factors. So, it is easy for a student to be offended by what another student says. In such exchanges, do everything you can to assure everyone in your class that, while there may be disagreement on issues, all opinions are to be valued, not scorned, to be respected, and not ridiculed.

These questions might lead to an answer based on belief / feeling:

- How are you? (O sa phela?)
- Please, may I ask you some questions today?
- What is the name of your business?
- Which bank is the bank of your business?
- Who is the manager of the business?
- Who is the electrician of the business?
- Who is the accountant of the business?

- Who is the insurance agent of the business?
- Who is the chairperson of the business?
- How many colleagues are there?
- Where is the piece of ground for the business?
- Is the building in a good condition?
- Do you have a trading license?
- Do you have any questions?

Information gap:

The term information gap covers a tremendous variety of techniques in which the objective is to convey or to request information. The two focal characteristics of information gap techniques are (1) their primary attention to information and not to language forms, and (2) the necessity of communicative interaction in order to reach the objective. The information that students must seek can range from very simple to complex.

Jigsaw:

Jigsaw techniques are a special form of information gap in which, each member of the group is given some specific information and the goal is to pool all information to achieve some objective.



Information gap and jigsaw can be represented in the following questions:

- What is the name of your business?
- Will you have a post box for your business?
- Will you have a telephone for your business?
- Where will you open your business – in which area?
- With which bank will you hold your bank account of the business?
- Who will be the manager / director of the business?
- Who will be the accountant of the business?
- Who will be the insurance agent of the business?
- Who will be the chairperson of the business?
- How many colleagues will there be?
- Who will the colleagues be in the business?

- Who will be the electrician of the business?
- Do you have a factory?
- Where will the factory be?
- Is the building in a good state or should you repair it or should you re-build it?
- Do you have a trading license?

Decision making:

This task has problem solving and decision making. Problem solving group techniques focus on the group's solution of a specified problem. Decision making techniques are simply one kind of problem where the ultimate goal is for students to make a decision.

The task will involve taking information from the written dialogue and coming to a decision or consensus about different questions.

- What is the name of your business?
- Will you have a post box for your business?
- Will you have a telephone for your business?
- Where will you open your business – in which area?
- With which bank will you hold your bank account of the business?
- Who will be the manager / director of the business?
- Who will be the accountant of the business?
- Who will be the insurance agent of the business?
- Who will be the chairperson of the business?
- How many colleagues will there be?
- Who will the colleagues be in the business?
- Who will be the electrician of the business?
- Do you have a factory?
- Where will the factory be?
- Is the building in a good state or should you repair it or should you re-build it?
- Do you have a trading license?

Several documents can be used as sources. This can be read before discussion begins.

Please take note that on some of these questions there is no single correct answer. This might have the implication that some discussions might take a while.

This technique, as well as, some of its' variations, give the learners considerably more control over what happens that is normally the case in the language classroom.

In order to help the discussion get going, the directions to the participants may contain a short list of factors that are to be considered, along with the instructions to 'justify your recommendations'.

Dialogue:

- You: Khotso, motswalle.
- English: Peace, friend.
- Client: Ee, khotso ntate.
- English: Yes, peace man.
- You: O sa phela?
- English: How are you?
- Client: Ke sa phela.
- English: I am fine.
- You: Ke kopa ho o botsa dipotso kajeno.
- English: Please, may I ask you some questions today?
- Client: Ho lokile.
- English: It's fine.
- You: Lebitso la kgwebo ke mang?
- English: What is the name of the business?
- Client: Lebitso la kgwebo ke Ramatheola.
- English: The name of the business is Ramatheola.
- You: Nomoro ya lebokoso la poso la kwebo ke mang?
- English: What is the postal address of the business?
- Client: Nomoro ya lebokoso la kgwebo ke 46, Dalsig, Stellenbosch, 7600.
- English: The postal address of the business is 46, Dalsig, Stellenbosch, 7600.
- You: Nomoro ya mohala ya kgwebo ke mang?
- English: What is the telephone number of the business?
- Client: Nomoro ya mohala ya kgwebo ke (021) 4473835.

- English: The telephone number of the business is (021) 4473835.
- You: Aterese ya kgwebo ke mang?
- English: What is the address of the business?
- Client: Aterese ya kgwebo ke 16 Serington seterata, Dalsig, Stellenbosch.
- English: The address of the business is 16 Serington street, Dalsig, Stellenbosch.
- You: Banka ya kgwebo ke efe?
- English: Which bank is the bank of the business?
- Client: Banka ya kgwebo ke Standard Banka.
- English: Standard Bank is the bank of the business.
- You: Motsomaisi ya kgwebo ke mang?
- English: Who is the manager (director) of the business?
- Client: Motsamaisi ya kgwebo ke Willie Lombaard.
- English: The manager of the business is Willie Lombaard.
- You: Mohlahlobi wa dibuka ya kgwebo ke mang?
- English: Who is the accountant of the business?
- Client: Mohlahlobi wa dibuka ya kgwebo ke Ranko Ntamae.
- English: The accountant of the business is Ranko Ntamae.
- You: Morekisi wa dinshoreense ya kgwebo ke mang?
- English: Who is the insurance agent of the business?
- Client: Morekisi wa dinshoreense ya kgwebo ke Motlalepula Mahlasi.
- English: The insurance agent of the business is Motlalepula Mahlasi.
- You: Modula-setulo ya kgwebo ke mang?
- English: Who is the chairperson of the business?
- Client: Modula-setulo ya kgwebo ke Phomolo Seboka.
- English: The chairperson of the business is Phomolo Seboka.
- You: Basebetsi-mmoho a boake?
- English: How many colleagues are there?
- Client: Ho na le basebetsi-mmoho a mararo.
- English: There are three colleagues.
- You: Ramotlakase ya kgwebo ke mang?
- English: Who is the electrician of the business?
- Client: Ramotlakase ya kgwebo ke Ntsokolo Faba.
- English: The electrician of the business is Ntsokolo Faba.

- You: O na le feme?
- English: Do you have a factory?
- Client: Ee, ke na le feme.
- English: Yes, I have a factory.
- You: Setsha sa kgwebo e kae?
- English: Where is the piece of ground of the business?
- Client: Setsha sa kgwebo e haufi le toropo.
- English: The piece of ground is near the town.
- You: Mohaho o motle na?
- English: Is the building nice?
- Client: Ee, mohaho o motle.
- English: Yes, the building is nice.
- You: O na le lakasense la ho hweba?
- English: Do you have a trading license?
- Client: Ee, ke rekile lakasense la ho hweba.
- English: Yes, I bought a trading license.
- You: Motswalle, na o na le dipotso na?
- English: Friend, do you have any questions?
- Client: Tjhee, ha ke na dipotso.
- English: No, I do not have questions.
- You: Ke hantle. Tsamaya hantle, motswalle.
- English: It's fine. Go well my friend.
- Client: Sala hantle, ntate.
- English: Stay well, man.

Task 3:

Sesotho:

Kajeno o bua le mohlankana ka batho ba sebetsang ka lebenkele. O botsa na ke mang a hirang batho le ka diforomo tsa kgiro. O bua ka konteraka le mangolo a lokelang. O botsa ka medikale edi le penshene. O bua ka mongodi le mosebetsi a hae le basebeletsi ba nkwana le mosebetso a bona. Le bua ka kopano le metsotso ya kopano le lenane tsamaiso.

English:

Today you talk to a young man about the people who are working in the shop. You ask who it is that employs the people and about the employment forms. You talk about the contract and the necessary qualifications. You ask about medical aid and pension. You talk about the secretary and her work and about the casual labourers and their work. You talk about a meeting, the minutes of the meeting and the agenda.

Narrative:

There may be narrative in this dialogue with every question asked. With every question asked the person might feel the need to tell you a short story to answer your question, for example of something that might have happened, or why the person wants to do something in a particular way, or if the person might want to explain something to you, or you might find that you have common interests and discuss that etc.

Opinion exchange:

An opinion is usually a belief or feeling that might not be founded on empirical data or that others could plausibly take issue with. Opinions are difficult for students to deal with at the beginning levels of proficiency, but by the intermediate level, certain techniques can effectively include the exchange of various opinions.

One warning: You play an important and sensitive role when you ask students to discuss their beliefs. Some beliefs are deeply ingrained from childhood rearing or from religious training, among other factors. So, it is easy for a student to be offended by what another student says. In such exchanges, do everything you can to assure everyone in your class that, while there may be disagreement on issues, all opinions are to be valued, not scorned, to be respected, and not ridiculed.

These questions might lead to an answer based on belief / feeling:

- How are you? (O phela hantle?)
- What is the problem?
- Who is it that employs people?

Information gap:

The term information gap covers a tremendous variety of techniques in which the objective is to convey or to request information. The two focal characteristics of information gap techniques are (1) their primary attention to information and not to language forms, and (2) the necessity of communicative interaction in order to reach the objective. The information that students must seek can range from very simple to complex.

Jigsaw:

Jigsaw techniques are a special form of information gap in which, each member of the group is given some specific information and the goal is to pool all information to achieve some objective.

Information gap and jigsaw can be represented in the following questions:

- Who is it that employs the people?
- Do you use employment forms?
- What is written in a contract?
- What is your view on qualifications needed for workers?
- What is your view on sick leave for workers?
- What is your view on overtime for workers?
- What is your view on leave for workers?
- What is your view on pension for workers?
- What is your view on medical aid for workers?
- What kind of workers do you want?
- Will you employ men and women?
- Who will work with the complaints?
- Who will work with the correspondence?
- What will be expected from the workers?
- How will you go about in your meetings?
- Who will take the minutes of the meeting?
- Who will write the agenda for the meeting?

Decision making:

This task has problem solving and decision making. Problem solving group techniques focus on the group's solution of a specified problem. Decision making techniques are simply one kind of problem where the ultimate goal is for students to make a decision.

The task will involve taking information from the written dialogue and coming to a decision or consensus about different questions.

- Who is it that employs the people?
- Do you use employment forms?
- What is written in a contract?
- What is your view on qualifications needed for workers?
- What is your view on sick leave for workers?
- What is your view on overtime for workers?
- What is your view on leave for workers?
- What is your view on pension for workers?
- What is your view on medical aid for workers?
- What kind of workers do you want?
- Will you employ men and women?
- Who will work with the complaints?
- Who will work with the correspondence?
- What will be expected from the workers?
- How will you go about in your meetings?
- Who will take the minutes of the meeting?
- Who will write the agenda for the meeting?

Several documents can be used as sources. This can be read before discussion begins.

Please take note that on some of these questions there is no single correct answer. This might have the implication that some discussions might take a while.

This technique, as well as, some of its' variations, give the learners considerably more control over what happens that is normally the case in the language classroom.

In order to help the discussion get going, the directions to the participants may contain a short list of factors that are to be considered, along with the instructions to 'justify your recommendations'.

Dialogue:

- You: Khotso, mohlankana.
- English: Peace, young man.
- Client: Ee, khotso ntate.
- English: Yes, peace man.
- You: O phela hantle?
- English: How are you?
- Client: Ee, ke phela hantle.
- English: Yes, I am fine.
- You: Molato ke eng?
- English: What is the problem?
- Client: Molato ke batho ba sebetsang ka lebenkele.
- English: The problem is the people who are working in the shop.
- You: Ke mang a hirang batho?
- English: Who is it that employs the people?
- Client: Ke nna.
- English: It is me.
- You: O sebedisa diforomo tsa kgiro?
- English: Do you use employment forms?
- Client: Ee, ba saena konteraka.
- English: Yes, they sign a contract.
- You: Ho ngotseng ka konteraka?
- English: What is written in a contract?
- Client: Ho ngotswe ka mangolo a lokelang, moputso a mosebeletsi, phomolo ya ho kula, letsatsi la phomolo ka beke le ho sebetsa ho feta nako e baletsweng.
- English: There are written about qualifications, the salary of the worker, sick leave, a day off in the week and to work overtime.
- You: Mosebeletsi o tla saena konteraka neng?

- English: When will the worker sign the contract?
- Client: O tla saena konteraka ka mora puisano.
- English: The worker will sign the contract after the interview.
- You: Mosebeletsi o na le mangolo a lokelang na?
- English: Does the worker have the necessary qualifications?
- Client: Tjhee, ke tla rupela basebeletsi.
- English: No, I will train the staff.
- You: Ba na le medikale edi le penshene?
- English: Do they have medical aid and pension?
- Client: Ee, ba na le tsona.
- English: Yes, they have it.
- You: O batla basebeletsi bafe?
- English: What kind of workers do you want?
- Client: Ke batla mongodi le basebeletsi ba nkwana.
- English: I want a secretary and casual labourers.
- You: Ke mang a tla sebestang ka dillo le ngollano?
- English: Who will work with complaints and correspondence?
- Client: Ke mongodi.
- English: It is the secretary.
- You: Basebeletsi ba nkwana ba tla etsang?
- English: What will the casual labourers do?
- Client: Ba tla thusa ka mesebetsi e sa tlwaelehang le ho bala thepa le ho nka ditaelo.
- English: They will help with odd jobs, to take stock and to take orders.
- You: O tseba hore kopano e sebetsa jwang?
- English: Do you know how a meeting works?
- Client: Ee, ke a tseba.
- English: Yes, I know.
- You: Ke mang a tla nka metsotso ya kopano?
- English: Who will take the minutes of the meeting?
- Client: Ke mongodi.
- English: It is the secretary.
- You: Ke mang a tla ngola lenane tsamaiso?
- English: Who will write the agenda?

Client: Ke nna hobane nna ke mong.
 English: I will because I am the boss.
 You: Ke o lakaletsa lehlohonolo.
 English: I wish you luck.
 Client: Ke lebohile haholo.
 English: Thank you very much.
 You: Tsela tshweu, mohlankana.
 English: Pleasant journey, young man.
 Client: Sala hantle, ntate.
 English: Stay well, man.

Task 4:

Sesotho:

Kajeno o bua le mme ka dithaba tsa banka. Le bua ka savings akhaonte le tjeheke akhaonte. O mo jwetsa hore o tla sebedisa akhaonte efe neneng. O mo hlaloseisa a etse eng ha a batla ho bula akhaonte, kentsha tjehelete kapa ntsha tjehelete. O mo jwetsa o tla fumana eng ha a boloka tjehelete bankeng. Hape o mo hlaloseisa hantle hore ba tla tseba jwang ke akhaonte sa hae.

English:

Today you speak to a woman on the things of the bank. You talk about a savings account and a cheque account. You tell her when to use which account. You explain to her what to do to open an account, put in some money and withdraw some money. You tell her what she will gain from putting her money in the bank. Then again you explain nicely to her how they will know which account is her account.

Narrative:

There may be narrative in this dialogue with every question asked. With every question asked the person might feel the need to tell you a short story to answer your question, for example of something that might have happened, or why the person wants to do something in a particular way, or if the person might want to explain

something to you, or you might find that you have common interests and discuss that etc.

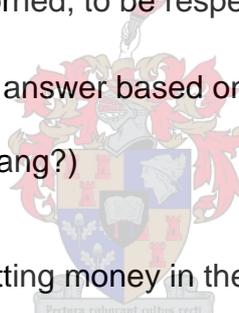
Opinion exchange:

An opinion is usually a belief or feeling that might not be founded on empirical data or that others could plausibly take issue with. Opinions are difficult for students to deal with at the beginning levels of proficiency, but by the intermediate level, certain techniques can effectively include the exchange of various opinions.

One warning: You play an important and sensitive role when you ask students to discuss their beliefs. Some beliefs are deeply ingrained from childhood rearing or from religious training, among other factors. So, it is easy for a student to be offended by what another student says. In such exchanges, do everything you can to assure everyone in your class that, while there may be disagreement on issues, all opinions are to be valued, not scorned, to be respected, and not ridiculed.

These questions might lead to an answer based on belief / feeling:

- How are you? (O phela jwang?)
- Can I help you?
- What will you gain from putting money in the bank?



Information gap:

The term information gap covers a tremendous variety of techniques in which the objective is to convey or to request information. The two focal characteristics of information gap techniques are (1) their primary attention to information and not to language forms, and (2) the necessity of communicative interaction in order to reach the objective. The information that students must seek can range from very simple to complex.

Jigsaw:

Jigsaw techniques are a special form of information gap in which, each member of the group is given some specific information and the goal is to pool all information to achieve some objective.

Information gap and jigsaw can be represented in the following questions:

- Should you open a bank account?
- Which kind of bank account will suit your business better?
- With which bank will you open the account?
- How much money should you deposit into your bank account with the opening of the account?
- What should you check for on the forms with the opening of your bank account?
- How much money should you deposit?
- How often should you deposit?
- Why should you deposit your money in a bank account?
- How much money should you withdraw?
- How often should you withdraw?
- Why should you withdraw your money from your bank account?

Decision making:

This task has problem solving and decision making. Problem solving group techniques focus on the group's solution of a specified problem. Decision making techniques are simply one kind of problem where the ultimate goal is for students to make a decision.

The task will involve taking information from the written dialogue and coming to a decision or consensus about different questions.

- Should you open a bank account?
- Which kind of bank account will suit your business better?
- With which bank will you open the account?
- How much money should you deposit into your bank account with the opening of the account?
- What should you check for on the forms with the opening of your bank account?
- How much money should you deposit?
- How often should you deposit?
- Why should you deposit your money in a bank account?

- How much money should you withdraw?
- How often should you withdraw?
- Why should you withdraw your money from your bank account?

Several documents can be used as sources. This can be read before discussion begins.

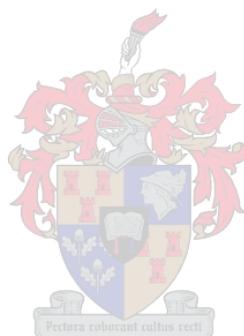
Please take note that on some of these questions there is no single correct answer. This might have the implication that some discussions might take a while.

This technique, as well as, some of its' variations, give the learners considerably more control over what happens that is normally the case in the language classroom.

In order to help the discussion get going, the directions to the participants may contain a short list of factors that are to be considered, along with the instructions to 'justify your recommendations'.

Dialogue:

- You: Dumela, mme.
- English: Good day, woman.
- Client: Ee, dumela ntate.
- English: Yes, good day man.
- You: O phela jwang, mme?
- English: How are you, woman?
- Client: Ke phela ha monate, ntate.
- English: I am fine, man.
- You: Nka o thusa, mme?
- English: Can I help you, woman?
- Client: Ee, ke kopa hore o ntlhalose akhoante tse fapaneneng bankeng.
- English: Yes, I ask that you explain the different bank accounts to me.
- You: Ho teng akhaonte tse ngata tse fapaneng tsa ho boloka tjhelete bankeng.
- English: There are many different accounts where you can put your money in the bank.
- Client: Akhaonte tse bohlokwa ke tse feng?



- English: Which are the main accounts?
- You: Akhaonte tse pedi tsa bohlokwa ke savings akhaonte le tjheke akhaonte.
- English: The two main accounts are savings accounts and cheque accounts.
- Client: Savings akhaonte ke eng?
- English: What is a savings account?
- You: Akhaonte ena ke ya bohlokwa haholo ha e le hore o batla ho kenya tjhelete le ho e ntsha kgafetsa.
- English: Savings accounts are good if you want to put in and take money out of the bank often.
- Client: Tjheke akhaonte ke eng?
- English: What is a cheque account?
- You: Akhaonte ena ke ya bohlokwa haholo ha e le hore o batla ho lefa melato ya hao ka tjheke. Tjheke ke sekotwana sa pampiri sa bankeng se sebediswang jwalo ka ha eka o sebedisa tjhelete ya hao e leng bankeng.
- English: This account is good if you want to pay all your accounts by cheque. A cheque is a piece of paper from the bank that is used as if it is your money in the bank.
- Client: Ke bula akhaonte jwang?
- English: How do I open an account?
- You: O hloka boitsebiso ya hao mmoho le tjhelete eo o tla bula akhaonte ka yona. Leba moo ho botswang teng mme o bolelle batho ba leng moo hore o batla ho bula akhaonte.
- English: You will need your ID and some money that you can open your account with. Go to enquiries and tell the people there that you want to open an account.
- Client: O ka mpoella ka diforomo tseo ke lokelang ho di tlatsa ha ke kenya tjhelete?
- English: Can you explain to me about all the forms I have to fill in if I want to put in money?
- You: Haeba o batla ho kenya tjhelete akhaonteng ya hao, hona ho bitswa ka hore ke deposit. O tla lokela ho tlatsa foromo ya deposit ha o etsa hona.

- English: If you wish to put money into your account, we call this a deposit. To do this you would need to fill in a deposit slip.
- Client: Haeba o batla ho ntsha tjhelete o tla etsa jwang?
- English: If you want to take out money, what will you do?
- You: Ha o ntsha tjhelete bankeng re re o etsa withdrawal. Ho teng foromo eo o lokelang ho e tlatsa bakeng sa hona.
- English: When you take money out of the bank, we say you make a withdrawal. There is a form you need to fill in for this.
- Client: Ke tla fumana eng ha ke boloka tjhelete bankeng?
- English: What will I gain from putting my money in the bank?
- You: Tjhelete ya hao ha e le sebakeng sena e bolokehile mme e ke ke ya utsuwa ha bonolo.
- English: If your money is kept here it is much safer and less likely to be stolen.
- Client: Hape?
- English: And?
- You: Taba ya bobedi ke hore ho teng tswala eo o tla e fumana ka tjhelete eo o e bolokang bankeng?
- English: Secondly, is the interest that you will get on the money in the bank.
- Client: O ka mpoella hore ba tla tseba jwang ke akhaonte sa ka?
- English: Can you please explain to me how they will know that it is my account?
- You: Akhaonte ya hao e na le lebitso la hao, nomoro mmoho le hore ke tjhelete e kae e teng ho yona.
- English: Your account has your name, a number and a record of money that you have in there.
- Client: Ke lebohile haholo, ntate.
- English: Thank you very much, man.
- You: Haeba ho ena le bothatha, jwalo ka ho tlatsa diforomo, o tla lokela ho ya moo ho botswang teng mme o kope hore ba othuse.
- English: If you have any problems, like filling in the forms, go to the enquiries desk and ask them to help you.
- Client: Ke lebohile ke ho nkgotsofalsa ka ho araba dipotso tsa ka hantle.
- English: Thank you for putting me at ease by answering all my questions clearly.
- You: Palama hantle, mme.
- English: Ride well, woman.

Client: Sala ka kgotso, ntate.

English: Stay in peace, man.

Task 5:

Sesotho:

Kajeno o bua le ntate ka ho kadima tjhelete. Le bua ka ho kadima tjhelete o lokela ho e kgutlisa e le ka hodima eo o e kadimmeng. O hlalosa phaello ho ntate. Hape o mo hlaloesa hore phaello e hodimo haholo kapa tlase haholo neng. Wena le ntate le bua ka tshireletso ke eng. O bua ka ntate hore a kadime tjhelete e boake mme o tla e lefa neng. O hlalosa ho ntate ba ngola eng diforomong mme a shebisiseng. Hape o re ho ntate a boloke khopi.

English:

Today you talk to the man about borrowing money. You talk about when you borrow money you must pay back more than you borrowed. You explain interest to the man. Then you explain to him when interest is very high or very low. You and the man talk about what security is. You talk to the man about how much money he must borrow and when he should pay it back. You explain to the man what they write on the forms and what he should check. You tell the man to keep a copy too.

Narrative:

There may be narrative in this dialogue with every question asked. With every question asked the person might feel the need to tell you a short story to answer your question, for example of something that might have happened, or why the person wants to do something in a particular way, or if the person might want to explain something to you, or you might find that you have common interests and discuss that etc.

Opinion exchange:

An opinion is usually a belief or feeling that might not be founded on empirical data or that others could plausibly take issue with. Opinions are difficult for students to

deal with at the beginning levels of proficiency, but by the intermediate level, certain techniques can effectively include the exchange of various opinions.

One warning: You play an important and sensitive role when you ask students to discuss their beliefs. Some beliefs are deeply ingrained from childhood rearing or from religious training, among other factors. So, it is easy for a student to be offended by what another student says. In such exchanges, do everything you can to assure everyone in your class that, while there may be disagreement on issues, all opinions are to be valued, not scorned, to be respected, and not ridiculed.

These questions might lead to an answer based on belief / feeling:

- How are you? (Ho jwang?)
- What can I do for you?
- Will you loan money to start a business?
- Will you give security for a loan?
- Will you pay interest on your loan?
- Will you sign an agreement for your loan?

Information gap:

The term information gap covers a tremendous variety of techniques in which the objective is to convey or to request information. The two focal characteristics of information gap techniques are (1) their primary attention to information and not to language forms, and (2) the necessity of communicative interaction in order to reach the objective. The information that students must seek can range from very simple to complex.

Jigsaw:

Jigsaw techniques are a special form of information gap in which, each member of the group is given some specific information and the goal is to pool all information to achieve some objective.

Information gap and jigsaw can be represented in the following questions:

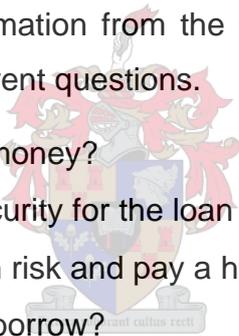
- Where will you loan some money?
- What will you put up for security for the loan of money not to be a high risk?

- Wouldl you rather be a high risk and pay a high interest?
- How much money will you borrow?
- When will you agree to pay back the money?
- If you do not understand what is written on the forms for the loan, who will you ask for help?
- Should you keep a copy of the loan agreement?
- For what should you check on the loan agreement?

Decision making:

This task has problem solving and decision making. Problem solving group techniques focus on the group's solution of a specified problem. Decision making techniques are simply one kind of problem where the ultimate goal is for students to make a decision.

The task will involve taking information from the written dialogue and coming to a decision or consensus about different questions.

- 
- Where will you loan some money?
 - What will you put up for security for the loan of money not to be a high risk?
 - Wouldl you rather be a high risk and pay a high interest?
 - How much money will you borrow?
 - When will you agree to pay back the money?
 - If you do not understand what is written on the forms for the loan, who will you ask for help?
 - Should you keep a copy of the loan agreement?
 - For what should you check on the loan agreement?

Several documents can be used as sources. This can be read before discussion begins.

Please take note that on some of these questions there is no single correct answer. This might have the implication that some discussions might take a while.

This technique, as well as, some of it's' variations, give the learners considerably more control over what happens that is normally the case in the language classroom.

In order to help the discussion get going, the directions to the participants may contain a short list of factors that are to be considered, along with the instructions to 'justify your recommendations'.

Dialogue:

You: Khotso, monna.

English: Peace, man.

Client: Ee, khotso ntate.

English: Yes, peace man.

You: Ho jwang?

English: How are you?

Client: Ho lokile.

English: I am fine.

You: Nka o etsetsang?

English: What can I do for you?

Client: Ke na le dipotso tse ngata.

English: I have a lot of questions.

You: O ka botsa dipotso.

English: You may ask questions.

Client: Hobaneng ho le theko e boima ho kadima tjhelete?

English: Why is it expensive to borrow money?

You: Ho kadima tjhelete o lokela ho e kgutlisa e le ka hodima eo o e kadimmeng.

English: When you borrow money you must pay back more than you borrowed.

Client: Hobaneng?

English: Why?

You: Lebaka ke hobane batho bao ba kadimana ka tjhelete bakeng sa ho etsa kgwebo.

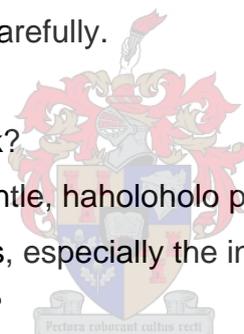
English: This is because those people who lend money do this as part of their business.

Client: Tjhelete e ka hodimo eo e o lefange bitswa eng?

English: What is the extra amount of money called?

- You: Tjhelete e ka hodimo eo o e lefang e bitswa phaello.
- English: The extra amount you pay is called interest.
- Client: Phaello ya bona e tla ba hodimo haholo neng?
- English: When will their interest be high?
- You: Dibaka tsena tse kadimanang ka tjhelete di ipeha kotsi hobane o ka nna wa hloleha ho kgutlisa tjhelete ena. Haeba ho le teng kotsi e jwalo, hangata phaello ya bona e ba hodimo haholo.
- English: Places which lend money take a risk that you may not pay back the money. If there is a big risk that you will not pay it back then the interest is high.
- Client: Phaello ya bona e tla ba tlase haholo neng?
- English: When will their interest be low?
- You: Haeba ba bona ba se kotsing, phaello ya bona e tla ba tlase haholo.
- English: If they see that their risk is low, their interest will be low.
- Client: Tshireletso ke eng?
- English: What is security?
- You: Tsena ke dintho tseo ba ka o amohang tsona ha o sa lefe.
- English: These are things that they can take if you do not pay up.
- Client: Ba tla batla ho tseba eng?
- English: What would they want to know?
- You: Ba tla batla ho tseba dintho tse kang moo o bolokang tjhelete ya hao teng bankeng kapa dintho tseo o nang le tsona, jwalo ka koloi kapa ntlo.
- English: They would want to know things like your savings in the bank or the things that you have, like a car or a house.
- Client: Ke tla kadima tjhelete e bokae?
- English: How much money will I borrow?
- You: Kadima tjhelete eo o tsebang hore o tla kgona ho e kgutlisa.
- English: Borrow an amount that you know you can pay back later.
- Client: Ke tla dumela ho lefa tjhelete neng?
- English: When will I agree to pay back the money?
- You: Dumela ho lefa tjhelete ka letsatsi leo o tla beng o na le yona, jwalo ka ha o qeta ho amohela.

- English: Agree to pay back the money on a day when you know you will have it, like your payday.
- Client: Ho ngotswe eng diforomong?
- English: What is written on the forms?
- You: Bala diforomo hantle. Etsa bonnete ba hore dintho tseo le dumellang ka tsona di ngotswe foromong. Ha ho le teng ho hong ho ngotsweng, bala seo hantle.
- English: Read the forms carefully. Make sure that the things you agreed on are written on the form. If there is extra writing, read it carefully.
- Client: Ha ke sa utlwise ho ngotsweng diforomong?
- English: If I do not understand what is written on the forms?
- You: Ha o sa utlwise, batla motho eo o mo tsebang ya tla o hlosetsa hantle.
- English: If you do not understand, look for a person that you know that will explain this to you carefully.
- Client: Ke shebisiseng?
- English: What should I check?
- You: Shebisisa dipalo hantle, haholoholo phaello eo le dumellang ka yona.
- English: Check all the figures, especially the interest you agreed to pay.
- Client: Ke ka boloka khopi?
- English: May I keep a copy?
- You: Ee, boloka khopi ya tumello ya lona e le hore o tle o e hlahise ha ho le teng mathatha.
- English: Yes, keep a copy of your agreement in case there is any problems later.
- Client: Ke lebohile ka thuso ya hao.
- English: Thank you for your help.
- You: Ho lokile.
- English: It is all right.
- Client: Sala hantle, ntate.
- English: Stay well, man.
- You: Phomola hantle, monna.
- English: Rest well, man.



Task 6:**Sesotho:**

Kajeno o bua le mofumahatsana ka thekiso. O bua ka mekgwa ho etsa thekiso. O bua ka ho phatlalatsa, ho bua le ho ngola dipampitshana. Hape o bua ka o ka nehelwa batho dipampitshana kae mme o ka di kentsha posong kapa o ka kenya dipampiri malapeng kaofela.

English:

Today you talk to a miss about marketing. You talk about the ways of marketing. You talk about advertising, to talk and to write flyers. You talk about where the flyers can be handed out and that you can post it or you can put it at every door.

Narrative:

There may be narrative in this dialogue with every question asked. With every question asked the person might feel the need to tell you a short story to answer your question, for example of something that might have happened, or why the person wants to do something in a particular way, or if the person might want to explain something to you, or you might find that you have common interests and discuss that etc.

Opinion exchange:

An opinion is usually a belief or feeling that might not be founded on empirical data or that others could plausibly take issue with. Opinions are difficult for students to deal with at the beginning levels of proficiency, but by the intermediate level, certain techniques can effectively include the exchange of various opinions.

One warning: You play an important and sensitive role when you ask students to discuss their beliefs. Some beliefs are deeply ingrained from childhood rearing or from religious training, among other factors. So, it is easy for a student to be offended by what another student says. In such exchanges, do everything you can to assure everyone in your class that, while there may be disagreement on issues, all opinions are to be valued, not scorned, to be respected, and not ridiculed.

These questions might lead to an answer based on belief / feeling:

- How are you? (O tsohile hantle?)
- Is marketing important for a business?
- Which different ways of marketing should one use?
- How much should one spend on marketing?
- What will you say if you use 'word of mouth' for advertising?
- Where will you hand out flyers to people?
- To whom should you post adverts? (mail-shots)
- Wher will you drop door to door adverts?

Information gap:

The term information gap covers a tremendous variety of techniques in which the objective is to convey or to request information. The two focal characteristics of information gap techniques are (1) their primary attention to information and not to language forms, and (2) the necessity of communicative interaction in order to reach the objective. The information that students must seek can range from very simple to complex.



Jigsaw:

Jigsaw techniques are a special form of information gap in which, each member of the group is given some specific information and the goal is to pool all information to achieve some objective.

Information gap and jigsaw can be represented in the following questions:

- What is marketing?
- How can you make marketing work for you?
- What is advertising?
- How many ways are there of advertising?
- Which way of advertising will work for you?
- How can word of mouth help?
- Whom should you tell about your business?
- What are flyers?
- Should it be printed, photocopied or handwritten?

- Where can you hand flyers out to people?
- Should adverts be posted?
- To whom should the mail-shots be posted?
- For what can mail-shots help?

Decision making:

This task has problem solving and decision making. Problem solving group techniques focus on the group's solution of a specified problem. Decision making techniques are simply one kind of problem where the ultimate goal is for students to make a decision.

The task will involve taking information from the written dialogue and coming to a decision or consensus about different questions.

- What is marketing?
- How can you make marketing work for you?
- What is advertising?
- How many ways are there of advertising?
- Which way of advertising will work for you?
- How can word of mouth help?
- Whom should you tell about your business?
- What are flyers?
- Should it be printed, photocopied or handwritten?
- Where can you hand flyers out to people?
- Should adverts be posted?
- To whom should the mail-shots be posted?
- For what can mail-shots help?

Several documents can be used as sources. This can be read before discussion begins.

Please take note that on some of these questions there is no single correct answer. This might have the implication that some discussions might take a while.

This technique, as well as, some of its' variations, give the learners considerably more control over what happens that is normally the case in the language classroom.

In order to help the discussion get going, the directions to the participants may contain a short list of factors that are to be considered, along with the instructions to 'justify your recommendations'.

Dialogue:

- You: Dumela, mofumahatsana.
- English: Good day, miss.
- Client: Ee, dumela ntate.
- English: Yes, good day man.
- You: O tsohile hantle?
- English: How are you? (Did you get up well?)
- Client: Ee, ke tsohile hantle.
- English: Yes, I am fine. (I got up well.)
- You: E se le kgale o eme?
- English: Have you been waiting long?
- Client: Tjhee. Ke kopa hore o ntlhalose thekiso.
- English: No. I ask if you please will explain marketing to me.
- You: Thekiso ke karolo ya bohlokwa kgwebong ka nngwe. Kgwebo e ka wa haeba ho se mereo e lokileng ya kgwebo. Bakeng sa hore kgwebo e tswellele, o tla tshwanela ho rekisa seo o se etsang.
- English: Marketing is a very important part of any business. A business can fail if there are no good marketing plans. For a business to be a success, you have to sell what you do.
- Client: Ho na le mekgwa e kae ya ho etsa thekiso?
- English: How many different ways are there to do marketing?
- You: Ho na le mekgwa e mengata ya ho etsa hona, jwalo ka ho phatlalatsa.
- English: There are many ways that you can do this, like advertising.
- Client: Ho phatlalatsa ke eng?
- English: What is advertising?
- You: Ho phatlalatsa ho tla o thusa ho rekisa kgwebo ya hao. Phatlalatso e hlileng e sebetsang haholo ke ya seyalemoya, TV kapa dikoranta.
- English: Advertising will help you sell your business. The most effective advertising is on the radio, TV and newspapers.

- Client: Hona ho bitsa tjhelete e bokae?
- English: How much will this cost?
- You: Hona o bitsa tjhelete e ngata. Haeba o na le tjhelete, leka ho batla motho ya tla akanya phatlalatso ya hao.
- English: This is very expensive. If you have the money, try to get someone to design an advert for you.
- Client: Ho bua ho tla thusa jwang?
- English: How will to speak (word of mouth) help?
- You: Bolella metswalle ya hao, ba lelapa kapa motho eo ho kopanang le yena feela ka kgwebo ya hao hore ba tsebe.
- English: Tell your friends, family or anyone you come across about your business, so that they can know.
- Client: Dipampitshana ke eng?
- English: What is flyers?
- You: Na o kile wa nehelwa pampirinyana e phatladitseng diphahlo tse rekiswang ka theolelo kapa tse buang ka ho bulwa ha lebenklele le letjha? Tsena di bitswa flyers.
- English: Have you ever been handed a pamphlet that advertises clothes for sale or those that tell you about the opening of a new shop? These are called flyers.
- Client: O ka di hatisa?
- English: Can they be printed?
- You: O ka di hatisa, o ka di etsa photocopy kapa hona ho di ngola ka letsoho.
- English: They can be printed, photocopied or even handwritten.
- Client: Di ka nehelwa batho kae?
- English: Where can it be handed out to people?
- You: Hore o tle o fumane tswelelo, o tshwanetse ho kgetha moo o tla isa flyers tsa aho teng. Hona ho tla ya ka kgwebo ya hao le mofuta wa bareki bao o ba batlang.
- English: To be most effective, you need to choose where to hand out your flyers. This will depend on your business and the type of customers you want to attract.
- Client: Ke ka romella diphatlalatso tsa ka ka poso?

- English: Can I post my adverts?
- You: Ee, ke ka lebaka lang ha o sa romele diphatlatso tsa hao ka poso ho ya dikopanong tse fapaneng le dikgwebong tse ka sebedisang tshebeletso ya hao ka moso.
- English: Yes, why not post adverts to the different organisations and businesses who could use your services in the future.
- Client: Hape mail-shots di thusa ka eng?
- English: For what can mail-shots help too?
- You: Hape mail shots di thusa haeba o batla ho phatlalatsa hara setjhaba se heno. Ho tsamaya o kenya dipampiri malapeng kaofela ho tla o neha setumo.
- English: Mail-shots are also useful if you want to advertise in your local community. Door to door advert drops will give you a lot of publicity.
- Client: Ke lebohile, ntate.
- English: Thank you, man.
- You: Phomola ka kgotso, mofumahatsana.
- English: Rest in peace, miss.
- Client: Sala ka kgotso.
- English: Stay in peace.



Task 7:

Sesotho:

Kajeno o bua le mofumahadi ka kgwebo e nyenyane e nang le katleho. Le bua ka phaello, ka moo thekiso ya thepa e tsamayang ka teng le ditshenyehelo. Hape le bua ka ho reka thepa e ngata ka nako e le nngwe le ho hlophisa ka tseo o batlang ho di reka le ho kgetha lebenkele.

English:

Today to you to a mrs about a succesful small business. You talk about profit margin, turn over and expenses. Then you talk about bulk buying and to plan your shopping and to choose a shop.

Narrative:

There may be narrative in this dialogue with every question asked. With every question asked the person might feel the need to tell you a short story to answer your question, for example of something that might have happened, or why the person wants to do something in a particular way, or if the person might want to explain something to you, or you might find that you have common interests and discuss that etc.

Opinion exchange:

An opinion is usually a belief or feeling that might not be founded on empirical data or that others could plausibly take issue with. Opinions are difficult for students to deal with at the beginning levels of proficiency, but by the intermediate level, certain techniques can effectively include the exchange of various opinions.

One warning: You play an important and sensitive role when you ask students to discuss their beliefs. Some beliefs are deeply ingrained from childhood rearing or from religious training, among other factors. So, it is easy for a student to be offended by what another student says. In such exchanges, do everything you can to assure everyone in your class that, while there may be disagreement on issues, all opinions are to be valued, not scorned, to be respected, and not ridiculed.

These questions might lead to an answer based on belief / feeling:

- How are you? (O tsohile?)
- What is a successful small business?
- How do you choose a shop?

Information gap:

The term information gap covers a tremendous variety of techniques in which the objective is to convey or to request information. The two focal characteristics of information gap techniques are (1) their primary attention to information and not to language forms, and (2) the necessity of communicative interaction in order to reach the objective. The information that students must seek can range from very simple to complex.

Jigsaw:

Jigsaw techniques are a special form of information gap in which, each member of the group is given some specific information and the goal is to pool all information to achieve some objective.

Information gap and jigsaw can be represented in the following questions:

- What is a profit margin?
- How to make a profit?
- What is turn over?
- How can turn over influence profit?
- What are expenses?
- How can expenses influence profit?
- How do you improve your profit?
- How can you cut down expenses?
- How do you choose where to shop?
- What is bulk buying?
- How do you plan your shopping?

*Decision making:*

This task has problem solving and decision making. Problem solving group techniques focus on the group's solution of a specified problem. Decision making techniques are simply one kind of problem where the ultimate goal is for students to make a decision.

The task will involve taking information from the written dialogue and coming to a decision or consensus about different questions.

- What is a profit margin?
- How to make a profit?
- What is turn over?
- How can turn over influence profit?
- What are expenses?
- How can expenses influence profit?

- How do you improve your profit?
- How can you cut down expenses?
- How do you choose where to shop?
- What is bulk buying?
- How do you plan your shopping?

Several documents can be used as sources. This can be read before discussion begins.

Please take note that on some of these questions there is no single correct answer. This might have the implication that some discussions might take a while.

This technique, as well as, some of its' variations, give the learners considerably more control over what happens that is normally the case in the language classroom.

In order to help the discussion get going, the directions to the participants may contain a short list of factors that are to be considered, along with the instructions to 'justify your recommendations'.



Dialogue:

You: Kgotso, mofumahadi.

English: Peace, Mrs.

Client: Ee, ke kgotso ntate.

English: Yes, peace man.

You: O tsohile?

English: How are you? (Did you get up?)

Client: Ee, ke tsohile.

English: Yes, I am fine. (I got up.)

You: Kajeno re bua ka kgwebo e nyenyane e nang le katleho.

English: Today we talk about a successful small business.

Client: Ee, batho bohle ha ba qala kgwebo, ba ebe ba batla ho fumana phaello.

English: Yes, anyone who starts a business wants to make a profit.

You: Ee, ho na le mantswe a mararo ao o hlolang o utlwileng borakgwebo ba bua ka ona ha ba bua ho etsa phaello.

- English: Yes, there are three words you always hear when business people talk about making profit.
- Client: Ona ke eng?
- English: What are these?
- You: Ona ke: phaello, ka moo thekiso ya thepa e tsamayang ka teng le ditshenyehelo.
- English: These are: profit margin, turn over and expenses.
- Client: Phaello ke eng?
- English: What is a profit margin?
- You: Ena ke tjhelete ya phaello eo o ka e fumanang ka mora ho rekisa. Haeba o rekile apole ya R1,00 mme o e rekisa R2,00, phaello ya hao e tla ba R1,00.
- English: This is the amount of profit you can make on the things you sell. If you buy an apple for R1,00 and sell it for R2,00, your profit margin is R1,00.
- Client: Ka moo, thekiso ya thepa e tsamayang ka teng ke eng?
- English: What is turn over?
- You: Ke palo ya thepa eo o e rekisitseng ka mora nako e itseng. Ha o rekisitse thepa e ngata, ho pepenene hore o tla fumana phaello e ngata. Palo ya thepa e rekisitsweng e ntse e batlile e tshwana le phaello.
- English: This is the amount of stock you sell over a period. If you sell a lot of stock, your profit will be high. Turnover of goods is directly linked to profits.
- Client: Ditshenyehelo ke eng?
- English: What are expenses?
- You: Ditshenyelo ke e nngwe ya dintho tse among kgwebo haholo. Ke ditshwenyehelo tsa hao: tse kang ho lefella sepalangwang le ho lefella rente ya lebenkele.
- English: Expenses is another thing that affects your profit. Expenses are your costs, like for transport and rent of your shop.
- Client: Ha ditshenyehelo di ngata ho etsahalang?
- English: What happens if the expenses are high?
- You: Ha o ena le ditshenyehelo tse ngata ha se hangata o kgonang ho etsa phaello e ntle.

- English: The bigger your expenses the less likely you are to make a good profit.
- Client: Hobaneng?
- English: Why?
- You: Hona ho etswa ke hobane phaello o e bala tjena, Phaello = Tjhelete eo o e fumaneng thekising – ditshenyehelo.
- English: This is because Profit = Income – Expenses.
- Client: Ho ntlafatsa phaello ya hao ho etsuang?
- English: What do you do to improve your profit?
- You: Ho na le dintho tse pedi tseo o lokelang ho di etsa bakeng sa ho ntlafatsa phaello ya hao. O ka nna wa nyolla phaello ya hao, ka ho nyolla tjhelete ya hao ya thekiso.
- English: To improve your profit you need to do two things. You can increase your profit margin by increasing your selling price.
- Client: Ntho ya bobedi?
- English: The second thing?
- You: Theola ditshenyehelo. O ka nna wa fokotsa ditshenyehelo ka ho reka thepa e ngata ka nako e le nngwe.
- English: Reduce your expenses. You can reduce costs by buying in bulk.
- Client: Hobaneng?
- English: Why?
- You: Ho reka thepa e ngata ka nako e le nngwe ho bolela ho boloka tjhelete.
- English: Bulk buying means the more you buy the less you pay.
- Client: O ka fokotsa dintshenyehelo jwang?
- English: How can you cut down on expenses?
- You: O ka boela wa fokotsa dintshenyehelo ka ho tlohela ho reka dintho tseo o hlokomelang hore ha di thuse kgwebo hore e hole.
- English: You can continue to cut down on cost expenses by not spending money on things that do not help build the business.
- Client: O fokotsa dintshenyehelo jwang hape?
- English: How can you cut down more expenses?
- You: Ho hlophisa ka tseo o batlang ho di reka.
- English: To plan your shopping.
- Client: Jwang?
- English: How?

- You: Ho hlopisa hona ho bolela hore o ngole le thathama la dintho tseo o tluaetseng ho di reka. O ntano shebisisa dintho tseo o felang o di hloka e le ka nnete.
- English: Planning your shopping means writing out a list you would normally buy. From this list, you then need to work out which of the things you really need.
- Client: O kgetha lebenkele jwang?
- English: How do you choose a shop?
- You: Ho molemo ho etela mabenkele a mangata mme o shebisisa mananeo a ona a thekiso. O tla makatswa ke phapang e kgolo eo o tla e bona.
- English: It helps to go to different shops and compare prices. You will be surprised to find out how these can differ between places.
- Client: Ke lebohile haholo ka thuso, ntate.
- English: Thank you very much for your help, man.
- You: Palama ka kgotso, mofumahadi.
- English: Ride in peace, mrs.
- Client: Sala hantle.
- English: Stay well



CHAPTER FIVE

CONCLUSION

This chapter presented a review of some central issues in language testing, especially drawing from Bachman and Palmer's work. We would like to use scores from language tests to make inferences about language ability or to make decisions about individuals. In order to do that we must be able to demonstrate how test performance corresponds to non-test language use. According to Bachman and Palmer, this can be done by developing a framework of language use that enables us to consider test performance as a specific instance of language use, a test taker as a language user in the context of language test, and a language test as a specific language use setting.

Bachman and Palmer stated that for language testing purposes we must consider language ability within an interactional framework of language use. A view is presented of language use that focuses on the interactions among areas of language ability, topical knowledge, and affective schemata, on the one hand, and how these interact with characteristics of the language use setting, or test task, on the other. This view is presented not as a working model of language processing, but as a basis for understanding how to design and develop language tests and how to use their results appropriately.

According to Bachman and Palmer, the characteristics of individuals – language users or test takers – include personal characteristics, topical knowledge, affective schemata, and language ability. Language ability consists of language knowledge and strategic competence, or metacognitive strategies. Language knowledge, which is information specific to language use that is stored in memory, includes both organizational knowledge and pragmatic knowledge. Organizational knowledge, which includes grammatical knowledge and textual knowledge, enables language users to create and interpret utterances or sentences that are grammatically accurate, and to combine these to form texts, either oral or written, that are cohesive and rhetorically or conversationally organized. Pragmatic knowledge, which includes functional knowledge and sociolinguistic knowledge, enables language users to

relate words, utterances, and texts to concepts, communicative goals, and the features of language use setting.

Strategic competence consists of metacognitive strategies, which are executive processes that enable language users to engage in goal setting, assessment and planning. According to Bachman and Palmer, goal setting involves:

Identifying the language use tasks or test tasks,

Choosing, where given a choice, one or more tasks from a set of possible tasks, and

Deciding whether or not to attempt to complete the task(s).

Assessment involves three aspects:

Assessing the characteristics of the language use task to determine the desirability and feasibility of accomplishing it,

Assessing the elements of topical and language knowledge to determine whether the necessary elements are available, and

Monitoring and evaluating the correctness and appropriateness of response – utterances or interpretation – in accomplishing the chosen task

Planning involves the formulation of one or more plans for implementation as a response to the task. Plans are implemented through the performance of language use tasks, involving interpreting and producing utterances or sentences in discourse. This model of language ability can be used both in the design of new language tests and in the selection of an existing test for a particular language testing situation.

According to Bachman and Palmer, the notion of language ‘skills’ as abstract modalities in which language is realized is felt not to be useful for language testing since attempts to characterize the four language skills (listening, speaking, reading, and writing) in terms of channel (audio, visual) and mode (receptive, productive) are inadequate on two counts. First, these features fail to capture important differences among language use activities that are within the same ‘skill’ (for instance, engaging in an oral conversation and listening to a radio newscast). Second, this approach to

distinguish the four skills treats them as abstract of language ability, ignoring the fact that language use is realized in specific situated language use tasks. Thus, rather than attempting to distinguish among four abstract skills, we find it more useful to identify specific language use tasks and to describe these in terms of their task characteristics and the areas of language ability they engage. If we are to find the concept 'skill' useful in the design, development, and use of language tests, then we must define specific 'skills' in terms of their task characteristics and the areas of language ability and topical knowledge these tasks engage.

Bachman and Palmer have looked at the extent to which writers in the field have experienced problems in conceptualizing the performance dimension of language performance tests. Progress in understanding the nature of second language performance testing and conducting the necessary research in its validity requires a commitment to rigorous analysis of what is involved in a language performance test. Only in such a context can empirical validation of scores from such assessments be carried out. The search for an adequate conceptualization of second language performance is not easy, but the need to broaden the discussion of the issues involved is pressing.

Skehan (1998: 290) concludes that until recently, performance testing has functioned at a fairly superficial level. Communicative tests have been seen as those which meet a set of performance conditions, but there has been little attempt to conceptualize the dimensions of performance in any convincing way. As a result, abilities-based approaches 'clothed' by communicative tests have been regarded as the most effective way to assess. Performance assessment is the result of a number of interacting factors. These include the nature of the raters of the performance, as well as the rating scales that they use. They also include the nature of interactional conditions which generate performance. According to Bachman and Palmer, most important for present purposes, they include (a) the tasks and task conditions which elicit language and (b) the dual-mode system that the candidate draws upon to actually use language. It is only meaningful to assess performance when each of these influences is operative in ways which can be described systematically. This analysis has a major impact on how the competence-performance relationship is conceptualized, since it demonstrates that performance itself has to be considered

independently, and is not there simply as a convenient method of accessing underlying competence. But the analysis also has fundamental importance for the nature of assessment and testing.

In order to say anything worthwhile about the nature of a second language learner's proficiency, Bachman and Palmer state that some clear view is necessary about how dual-mode systems are engaged, and how language is elicited through the tasks which are chosen, and how those tasks are elicited. The method of assessment, in other words, is not a technical problem at the level of format effects whose intrusive influences needs to be eliminated – it is *the* central problem in testing, since it is how languages are elicited, and how performance is engaged that is the crux of the matter and the basis for any generalization on language ability that is made. To that end, Bachman and Palmer conclude, studies of task characteristics and of task conditions such as planning or the introduction of surprise elements are essentially studies into testing as much as they are studies of communicative performance and development. Only in the context of such research can defensible decisions be made about the composition of assessment batteries. Without such research the development of 'communicative' approaches to testing will only be a veneer, based on intuition and hunch, rather than founded on principle.

The stages in the development of work sample performance tests have been outlined, and the carrying out of such a development project has been described in some detail. An example has been constructed and it is useful for the reader to have a clear idea of its structure and purpose.

Bachman & Palmer (1996: 38) note that the most important consideration in designing a language test is its usefulness, and this can be defined in terms of six test qualities: reliability, validity, authenticity, interactiveness, impact, and practicality. These six test qualities all contribute to test usefulness. The relative importance of these different qualities will vary from one testing situation to another. The most important consideration to keep in mind is not to ignore any one quality at the expense of the others. Rather, we need to strive to achieve an appropriate balance, given the purpose of the test, the characteristics of the TLU domain and the test takers, and the way we have defined the construct to be measured. Reliability can

be defined as consistency of measurement; inconsistency is variation in test scores that is due to factors other than the construct we want to measure.

Construct validity pertains to the meaningfulness and appropriateness of the interpretations that we make on the basis of test scores. Authenticity is an important quality for language tests for two reasons:

It provides a link between test performance and the TLU tasks and domain to which we want to generalize, and

The way test takers perceive the relative authenticity of test tasks can, potentially, facilitate their test performance.

Interactiveness is an important test quality because it pertains to the degree to which the constructs we want to assess are critically involved in accomplishing the test task.

Impact can be defined broadly in terms of the various ways in which test use affects society, an education system, and the individuals within these. Impact operates at two levels: macro and micro level.

Unlike the other five qualities, which pertain to the uses that are made of test scores, practicality pertains to the ways in which the test will be implemented in a given situation, or whether the test will be used at all.

Several types of resources can be identified.

In designing and developing a test, we try to achieve the optimum balance among the qualities of reliability, construct validity, authenticity, interactiveness, and impact, for our particular testing situation. In addition, we must determine the resources required to achieve this balance, in relationship to the resources that are available.

Douglas (1995: 176) concludes that with regard to testing theory, an emphasis has been placed on the further understanding of the construct of language proficiency. In the area of test methods, there has been an increasing interest in performance testing and the promotion of professional standards of test development and use. In test analysis, emphasis has been placed on the use of item response and

generalizability theories as well as understanding the multiple sources of variance in test performance.

Characteristics of the tasks

Characteristics that can have an important influence on both language use and language test performance like, characteristics of an individual, personal characteristics, topical knowledge that test takers bring to the language testing situation, their affective schemata and their language ability are considered in the design of these tests.

These tests are designed so that these characteristics can rather facilitate than impede the test takers' performance.

In the design of these tests attention is focused on the qualities of execution in the performance and of the evidence it provides about the candidate's control of the underlying linguistic system.

These tests are designed to be practicable, which make it good tests.

In the designed of these tests self esteem are helped to be built. The learner can start with the real easy questions and can 'qualify' for the next level – feeling good about themselves for mastering the first level of questions. Only when the learner feels confident enough he might move onto the next level.

The inhibitions, which we place between others and ourselves, are important factors contributing to second language success. These tests are characterised by the creation of context for meaningful classroom communication, such that the interpersonal ego barriers are lowered to pave the way for free, unfettered communication.

These test tasks promote risk-taking, which is an important characteristic of successful learning of a second language. The learner can do these test tasks by himself, alone, which removes the threat of being foolish in front of the other learners.

Facilitative anxiety is some concern – some apprehensions – over a task to be accomplished, which one will find in these test tasks. There are positive effects of competitiveness by means of the construct of facilitative anxiety.

In teaching languages we need to be sensitive to cultural norms, to a student's willingness to speak out in class, and to optimal points between extreme extroversion and introversion, that may vary from student to student. With these different test tasks, there has been taken precautions to facilitate these different elements.

Motivation is a characteristic of these tests. Intrinsically motivation in these tests will be found when a person is learning for their own self-perceived needs and goals. Extrinsically motivation will be found with those who learn only to receive an external reward from someone else. These test tasks are also designed to motivate learners to move from one level to the next.

To elicit test takers' best performance, is to design the characteristics of the test task to promote feelings of comfort or safety in test takers that will in turn facilitate flexibility of response on their part. Test takers might feel comfortable and safe doing these test tasks since they can be exercised in the comfort of one's own home and safe surroundings without any other person around.

Language ability is involving two components: language competence and strategic competence, which we will describe as a set of metacognitive strategies. It is this combination of language knowledge and metacognitive strategies that provide language users with the ability to create and interpret discourse, either in responding to tasks on language tests or in non-test language use. In these language tests language knowledge and metacognitive strategies are tested in different ways. These language tests are informed by a broad view of language ability.

Language knowledge includes two broad categories: organisational knowledge and pragmatic knowledge. We recognise that many of the language tests we develop will focus on only one or a few areas of language knowledge. There are two areas of organisational knowledge: grammatical knowledge and textual knowledge. Grammatical knowledge, which includes the knowledge of vocabulary, syntax, phonology and graphology, tested in producing formally accurate utterances or sentences, are well tested in these tests.

Textual knowledge is involved in producing or comprehending texts, which are units of language –spoken or written- that consists of two or more utterances or sentences. There are two areas of textual knowledge: knowledge of cohesions and knowledge of rhetorical or conversational organisation.

Knowledge of cohesion is involved in producing or comprehending the explicitly marked relationships among sentences in written texts among utterances in conversations.

Knowledge of rhetorical or conversational organisation is involved in producing and comprehending organisational development in written texts or in conversations.

Textual knowledge plays a very important role and is tested in different ways in these tests.

Pragmatic knowledge enables us to create or interpret discourse by relating utterances or sentences and texts to their meanings, to the language use setting. Pragmatic knowledge can be tested in a lot of different ways, which is also done in these tests as well. The result is that in the end language use will be of a higher and higher standard, the more pragmatic knowledge is developed.

Bachman and Palmer (1996: 69) suggest that there are two areas of pragmatic knowledge: functional knowledge and socio-linguistic knowledge.

Functional knowledge, or what Bachman (1990) calls "illocutionary competence", enables us to interpret relationships between utterances or sentences and texts and the intentions of language users. Functional knowledge includes knowledge of four categories of language functions: ideational, manipulative, instrumental, and imaginative.

Knowledge of ideational functions enables us to express or interpret meaning in terms of our experience of the real world. These functions include the use of language to express or exchange information about ideas, knowledge, or feelings. Descriptions, classifications, explanations, and expressions of sorrow or anger are examples of utterances that perform ideational functions.

Knowledge of manipulative functions enables us to use language to affect the world around us. This includes knowledge of the following:

- 1) Instrumental functions, which are performed to get other people to do things for us (examples include requests, suggestions, commands, and warnings);
- 2) Regulatory functions, which are used to control what other people do (examples include rules, regulations, and laws); and
- 3) Interpersonal functions, which are used to establish, maintain, and change interpersonal relationships (examples include greetings and leave-takings, compliments, insults, and apologies).

Knowledge of heuristic functions enables us to use language to extend our knowledge of the world around us, such as when we use language for teaching and learning, for problem solving and for the retention of information.

Knowledge of imaginative functions enables us to use language to create an imaginary world or extend the world around us for humorous or aesthetic purposes; examples include jokes and the use of figurative language and poetry.

Socio-linguistic knowledge enables us to create or interpret language that is appropriate to a particular language use setting. This includes knowledge of the conventions that determine the appropriate use of dialects or varieties, registers, natural or idiomatic, expressions, cultural references, and figures of speech.

Strategic competence is a set of metacognitive components, or strategies, which can be thought of as higher order executive processes that provide a cognitive management function in a language use. With respect to language testing, this conceptualisation of strategic competence as metacognitive components provides an essential basis both for designing and developing potentially interactive test tasks and for evaluating the interactiveness of the test task we use. We identify three areas: goal setting, assessment, and planning.

Goal settings involves identifying the language use tasks, choosing, where given a choice, one or more tasks from a set of possible tasks, and deciding whether to attempt to complete the tasks.

Since the purpose of a language test is to elicit a specific sample of language use, tests typically present the test taker with a limited range of tasks, as in the tests that are designed in the small business' program. The test takers' flexibility in setting goals for performance on test tasks is generally not as great as that enjoyed by language users in non-test language use.

Assessment provides a means by which the individual relates her topical knowledge and language knowledge to the language use setting and tasks or to the testing situation and tasks, which is carefully considered in the designing of these tests.

Planning is a very strong characteristic in the design of these tasks. Planning involves deciding how to utilise language knowledge, topical knowledge and affective schemata to complete the test task successfully. Planning involves three aspects:

- 1) Selecting a set of specific elements from topical knowledge and language knowledge (for example, concepts, words, structures, function) that will be used in a plan.
- 2) Formulating one or more plans whose realism will be a response (interpretation, utterance) to the task, and
- 3) Selecting one plan for implementation as a response to the task.

Language performance is divided in two forms of productive performance, oral and written, and two forms of receptive performance, aural and reading. Production and reception are quite simple two sides of the same coin. The interrelationship between written and spoken language is an intrinsically motivating reflection of language and culture and society. In these tasks we invited all of the four skills that are relevant.

A communicative test has to meet some rather stringent criteria. These tests are testing for grammatical, discourse, socio-linguistic, and illocutionary competence as well as strategic competence. It has to be pragmatic in that it requires the learner to use language naturally for genuine communication and to relate to thoughts and feelings, in short, to put authentic language to use within a context. It should be direct (as opposed to indirect tests that may lose validity as they lose content validity). In addition, it should test the learner in a variety of language functions. Bachman offers four distinguishing characteristics: First, such tests create an

“information gap”, requiring test takers to process complementary information through the use of multiple sources of input. A second characteristic is that of task dependency, with tasks in one section of the test building upon the content of earlier sections, including the test taker’s answers to those sections. Third, communicative tests can be characterised by their integration of test tasks and content within a given domain of discourse. Finally, communicative tests attempt to measure a much broader range of language abilities – including knowledge of cohesion, functions, and socio-linguistic appropriateness – than did earlier tests which tended to focus on the formal aspects of language – grammar, vocabulary, and pronunciation.

These tests have motivating, interesting, and substantive content, and it is integrated and interactive. By eliciting opinions and the expression of one’s own ideas, and interactive elements occurs in these tests.

These tests do everything possible to elicit the very best performance from students.

In these tests a needs analysis (as in task 1 to task 7) has been done to specify the context of the second language use, the type of interaction foreseen; the roles, discourse types, and language functions to be performed; and the basis on which successful fulfilment of the second language tasks is to be judged. It is with respect to these needs that the performance tests were designed, tests and tasks were selected, and evaluation criteria were determined. These were then translated into appropriate test objectives, and tasks, and later into actual test design and scoring.

These tests are attempting to measure real-world tasks and have standards that are authentic and are clear to students.

These tests are concerned to measure students’ abilities to respond to real-life language tasks. A more communicative picture of student knowledge and abilities than traditional (multiple-choice, and so forth) assessment are offered by these tests.

Critical thought, creativity and self-reflection are encouraged by the tests.

These tests are showing students’ strengths and weaknesses in real-world terms.

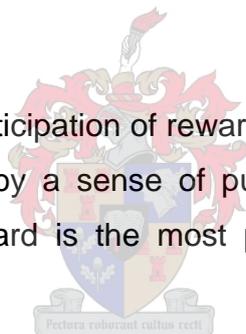
Achievement information by the levels that students work through is supplied by these tests. The students’ abilities in future, real-world situations are predicted.

These tests are designed well in the sense that all 4 levels of proficiency are represented in these tests.

Automaticity is a characteristic of these tasks – it is better to acquire language subconsciously – without overtly analyzing the forms of language. Overanalyzing language, thinking too much about its forms, and consciously lingering on rules of automaticity all tend to impede this graduation to automaticity.

Meaningful learning is represented in these tasks. The power of meaningful learning is capitalized by appealing to students' interests. The pitfalls of rote learning are avoided: too much grammar explanations, too many abstract principles and theories, activities whose purpose are not clear, activities that do not contribute to accomplishing the goals of lesson or unit or course, and techniques that are so mechanical its tricky that it get centred on the mechanics instead of the language or meanings. Meaningful learning will lead to better long-term retention than rote learning.

In these tasks you will find the anticipation of reward. There is virtually nothing we do that is not inspired and driven by a sense of purpose or goal, and according to Skinner, the anticipation of reward is the most powerful factor in directing one's behaviour.



The most powerful rewards are those that are intrinsically motivated within the learner.

Learner-centred teaching is represented in these tasks. Techniques are used that allow for student creativity and innovation, techniques that enhance a student's sense of competence and self-worth.

Interactive learning – doing a significant amount of pair work and group work, receiving authentic language input in real-world contexts, producing language of genuine meaningful communication, performing classroom tasks that prepare them for actual language use 'out there', practising oral communication through the give and take and spontaneity of actual conversations, writing to and for audiences, are well incorporated in these test tasks.

1. Pictionary
2. Snakes & Ladders
3. Sound Snap
4. Crossword Puzzle
5. Role-play
6. Games
7. Simulation
8. Drama
9. Projects
10. Interview
11. Brainstorming
12. Jigsaw

1. Pictionary

Devide the class into two groups. Everyone in the group gets a chance. There are 15 different words each written on a piece of paper which in thrown in a hat. Each group, taking turns, has a person – taking turns, drawing a piece of paper from the hat. There is a Sesotho word written on the piece of paper. That person must draw the picture of the Sesotho word to the rest of his group and they must guess what he is trying to draw in Sesotho (without knowing what is written on the piece of paper).

Higher points are scored the quicker the group can guess, in Sesotho, what the person is trying to draw. Bonus points will be given for correct pronunciation. The person who is drawing is not allowed to utter a word untill the group has decided correctly what he is trying to draw.

Option 2 is to have the groups guessing what the learner is drawing, giving the correct English translation of the answer. Higher points are scored the quicker the group can guess what the person is trying to draw.

Option 3 is to have the groups guessing what the learner is drawing, giving the correct Sesotho word together with the correct English translation of the answer. Higher points are scored the quicker the group can guess, in Sesotho together with

the correct English translation, what the person is trying to draw. Bonus points will be given for correct pronunciation.

Option 4 is to have the groups writing down the Sesotho word correctly after guessing what the person was trying to draw. Points will be lost for each spelling mistake.

Option 5 is to have the groups writing down the English translation correctly after guessing what the person was trying to draw. Points will be lost for each mistake and each spelling mistake.

Option 6 is to have the groups writing down the Sesotho word together with the English translation correctly after guessing what the person was trying to draw. Points will be lost for each mistake and each spelling mistake.

Option 7 is to use the pictures attached and to give all the different Sesotho words to the groups as well. The groups then has to shuffle all the pictures and Sesotho words to fit. The group to have the correct Sesotho word put together with the correct pictures first is the group to win.

Option 8 is to use the pictures attached and to give all the different English words together and to ask the group to put the picture and the English word together. The group who is finished first with all the words and pictures correctly is the group to win.

Option 9 is to use the pictures attached and to give all the different Sesotho and English words together and to ask the group to put the picture, the Sesotho word and the English word together. The group who is finished first with all the words and pictures correctly is the group to win.

Option 10 is to give the pictures to the group and to ask them to write down the correct Sesotho word for each picture. The group who is finished first will be the group to win. Points will be lost for any wrong words or spelling mistakes.

Option 11 is to give the pictures to the group and to ask them to write down the correct English translation for each picture. The group who is finished first will be the group to win. Points will be lost for any wrong translations or spelling mistakes.

Option 12 is to give the pictures to the group and to ask them to write down the correct Sesotho word along with the correct English translation for each picture. The group who is finished first will win. Points will be lost for any wrong words or spelling mistakes.

This test-item will contribute to the pronunciation of words (oral performance), the writing of words (written performance), listening to Sesotho (aural performance) and to reading Sesotho (reading performance). Thus productive and receptive performance will be tested.

This test-item has information gap as it is requesting information. It also has problem solving as the group is looking for a solution of a specified problem. The test-item does not really have decision making since the group is not involved in a decision about what to *do*. The test-item has a low level of opinion exchange as to what the picture is. Narrative would not really be found in this test-item as it is not presenting the opportunity for a short or long story.

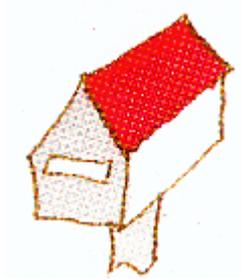
Notes:

- Make sure that you help the learners to get accustomed to the test format.
- This test can be divided into 4 levels of proficiency. Roughly represented as follows: level 1 (20%), level 2 (25%), level 3 (30%) and level 4 (25%).



Pictionary
word.

The pictures with the English translation and the Sesotho



1. Post box
Lebokoso la poso



2. School
Sekolo



3. Child
Ngwana



4. Secretary
Mongodi



5. Telephone
Mohala



6. Electrician
Ramotlakase



7. House
Ntlo



8. Sir
Monghadi



9. Money
Tjhelete



10. Meeting
Kopano



11. Car
Koloi



12. Mrs
Mofumahadi



13. Radio
Seyalemoya



14. Church
Kereke



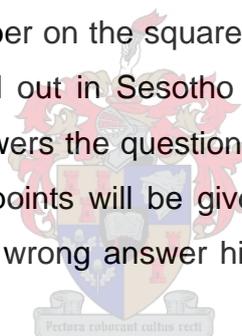
15. Newspaper
Koranta

2. Snakes & Ladders

Option 1: Each person in the group, taking turns, gets a chance to throw the dice. The one who throws the first 6 starts. The learners, taking turns, must first throw a 6 before he / she can start. Once the learner has thrown a 6, he / she can throw the dice again. The number on the dice is the number of squares that the learner must move forward with his “piece”.

The question of the number must be read out in Sesotho to the learner by another learner – taking turns to read. The learner must answer the question in Sesotho correctly before he can throw the dice again. Higher points are scored the quicker the learner can give the correct answer.

If the question is answered correctly, the learner can throw the dice again straight away. The number on the dice is the amount of blocks that the learner can move with his “piece”. Again, the number on the square that the piece has landed on, that number’s question must be read out in Sesotho to the learner by another learner (taking turns). If the learner answers the question correctly in Sesotho he can throw the dice straight away. Bonus points will be given for correct pronunciation. The moment that the learner gives a wrong answer his turn is lost and the next learner can take his turn.



If you land on a block with a ladder on it, after answering the question correctly, the learner can climb the ladder with his “piece”. If the learner answered the question incorrectly, he / she lost the chance to make use of the ladder at all. When the learner reaches the top of the ladder, that block’s question must be asked and answered again. If the question is answered correctly, the learner can throw the dice again straight away. The moment that the learner gives a wrong answer his turn is lost and the next learner can take his turn. If the learner lands on a block where a snake’s head is, he can “save” himself by answering the question correctly. If the learner answered the question incorrectly he is bitten by the snake and has to move his “piece” all the way down to the end of the snake’s tail.

Option 2 is that the question of the number must be read out in Sesotho to the learner by another learner – taking turns to read. The learner must answer the

question, giving the correct English translation of the answer, before they can throw the dice again. Higher points are scored the quicker the learner can give the correct answer.

Option 3 is that the question of the number must be read out in Sesotho to the learner by another learner – taking turns to read. The learner must answer the question, giving the Sesotho word together with the correct English translation of the answer, before they can throw the dice again. Higher points are scored the quicker the learner can give the correct answer. Bonus points will be given for correct pronunciation.

Option 4 is to write down the correct answer in Sesotho on a piece of paper after the question was read to the learner in Sesotho by another learner. If there are any spelling mistakes, the learner has lost his chance for that round. Higher points are scored the quicker the learner can give the correct answer.

Option 5 is to write down the correct English translation of the answer after the question was read to the learner in Sesotho by another learner. If there are any mistakes or spelling mistakes, the learner has lost his chance for that round. Higher points are scored the quicker the learner can give the correct answer.

Option 6 is to write down the correct answer in Sesotho together with the correct English translation of the answer after the question was read to the learner in Sesotho by another learner. If there are any mistakes or spelling mistakes, the learner has lost his chance. Higher points are scored the quicker the learner can give the correct answer.

Option 7 is to write down the question in Sesotho without any spelling mistakes after the question was read to the learner in Sesotho by another learner. If there are any spelling mistakes, the learner has lost his chance. Higher points are scored the quicker the learner can give the correct answer.

Option 8 is to write down the correct English translation of the question after the question was read to the learner in Sesotho by another learner. If there are any mistakes or spelling mistakes, the learner has lost his chance. Higher points are scored the quicker the learner can give the correct answer.

Option 9 is to write down the question in Sesotho together with the correct English translation of the question after the question was read to the learner in Sesotho by another learner. If there are any mistakes or spelling mistakes, the learner has lost his chance. Higher points are scored the quicker the learner can give the correct answer.

Option 10 is that the question of the number must be read out in Sesotho to the learner by another learner – taking turns to read. All the cards with the different Sesotho answers to the questions are given to the learner. The learner then has to choose from the cards which is the correct answer to the question. Higher points are scored the quicker the learner can give the correct answer.

Option 11 is that the question of the number must be read out in Sesotho to the learner by another learner – taking turns to read. All the cards with the different English translations of the answers are given to the learner. The learner then has to choose from the cards which is the correct translation to the answer. Higher points are scored the quicker the learner can give the correct answer.

Option 12 is that the question of the number must be read out in Sesotho to the learner by another learner – taking turns to read. All the cards with the different Sesotho answers to the questions as well as the cards with the different English translations are given to the learner. The learner then has to choose from the cards which is the correct answer to the question. Higher points are scored the quicker the learner can give the correct answer.

Option 13 is that the question of the number must be read out in Sesotho by the learner himself. The learner must answer the question in Sesotho correctly before they can throw the dice again. Higher points are scored the quicker the learner can give the correct answer. Bonus points will be given for correct pronunciation.

Option 14 is that the question of the number must be read out in Sesotho by the learner himself. The learner must answer the question, giving the correct English translation of the answer, before they can throw the dice again. Higher points are scored the quicker the learner can give the correct answer.

Option 15 is that the question of the number must be read out in Sesotho by the learner himself. The learner must answer the question, giving the Sesotho word together with the correct English translation of the answer, before they can throw the dice again. Higher points are scored the quicker the learner can give the correct answer. Bonus points will be given for correct pronunciation.

Option 16 is to write down the correct answer in Sesotho on a piece of paper after the question was read by the learner himself. If there are any spelling mistakes, the learner has lost his chance for that round. Higher points are scored the quicker the learner can give the correct answer.

Option 17 is to write down the correct English translation of the answer after the question was read by the learner himself. If there are any mistakes or spelling mistakes, the learner has lost his chance for that round. Higher points are scored the quicker the learner can give the correct answer.

Option 18 is to write down the correct answer in Sesotho together with the correct English translation of the answer after the question was read by the learner himself. If there are any mistakes or spelling mistakes, the learner has lost his chance. Higher points are scored the quicker the learner can give the correct answer.

Option 19 is to write down the correct English translation of the question after the question was read by the learner himself. If there are any mistakes or spelling mistakes, the learner has lost his chance. Higher points are scored the quicker the learner can give the correct answer.

Option 20 is that the question of the number must be read out in Sesotho by the learner himself. All the cards with the different Sesotho answers to the questions are given to the learner. The learner then has to choose from the cards which is the correct answer to the question. Higher points are scored the quicker the learner can give the correct answer.

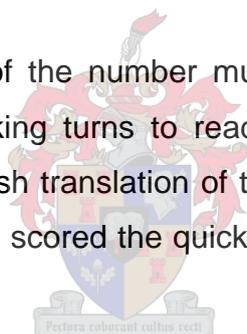
Option 21 is that the question of the number must be read out in Sesotho by the learner himself. All the cards with the different English translations of the answers are given to the learner. The learner then has to choose from the cards which is the

correct translation to the answer. Higher points are scored the quicker the learner can give the correct answer.

Option 22 is that the question of the number must be read out in Sesotho by the learner himself. All the cards with the different Sesotho answers to the questions as well as the cards with the different English translations are given to the learner. The learner then has to choose from the cards which is the correct answer to the question. Higher points are scored the quicker the learner can give the correct answer.

Option 23 is that the question of the number must be read out in English to the learner by another learner – taking turns. The learner must answer the question in Sesotho correctly before they can throw the dice again. Higher points are scored the quicker the learner can give the correct answer. Bonus points will be given for correct pronunciation.

Option 24 is that the question of the number must be read out in English to the learner by another learner – taking turns to read. The learner must answer the question, giving the correct English translation of the answer, before they can throw the dice again. Higher points are scored the quicker the learner can give the correct answer.



Option 25 is that the question of the number must be read out in English to the learner by another learner – taking turns to read. The learner must answer the question, giving the Sesotho word together with the correct English translation of the answer, before they can throw the dice again. Higher points are scored the quicker the learner can give the correct answer. Bonus points will be given for correct pronunciation.

Option 26 is to write down the correct answer in Sesotho on a piece of paper after the question was read to the learner in English by another learner. If there are any spelling mistakes, the learner has lost his chance for that round. Higher points are scored the quicker the learner can give the correct answer.

Option 27 is to write down the correct English translation of the answer after the question was read to the learner in English by another learner. If there are any

mistakes or spelling mistakes, the learner has lost his chance for that round. Higher points are scored the quicker the learner can give the correct answer.

Option 28 is to write down the correct answer in Sesotho together with the correct English translation of the answer after the question was read to the learner in English by another learner. If there are any mistakes or spelling mistakes, the learner has lost his chance. Higher points are scored the quicker the learner can give the correct answer.

Option 29 is to write down the question in Sesotho without any spelling mistakes after the question was read to the learner in English by another learner. If there are any spelling mistakes, the learner has lost his chance. Higher points are scored the quicker the learner can give the correct answer.

Option 30 is to write down the correct English translation of the question after the question was read to the learner in English by another learner. If there are any mistakes or spelling mistakes, the learner has lost his chance. Higher points are scored the quicker the learner can give the correct answer.

Option 31 is to write down the question in Sesotho together with the correct English translation of the question after the question was read to the learner in English by another learner. If there are any mistakes or spelling mistakes, the learner has lost his chance. Higher points are scored the quicker the learner can give the correct answer.

Option 32 is that the question of the number must be read out in English to the learner by another learner – taking turns to read. All the cards with the different Sesotho answers to the questions are given to the learner. The learner then has to choose from the cards which is the correct answer to the question. Higher points are scored the quicker the learner can give the correct answer.

Option 33 is that the question of the number must be read out in English to the learner by another learner – taking turns to read. All the cards with the different English translations of the answers are given to the learner. The learner then has to choose from the cards which is the correct translation to the answer. Higher points are scored the quicker the learner can give the correct answer.

Option 34 is that the question of the number must be read out in English to the learner by another learner – taking turns to read. All the cards with the different Sesotho answers to the questions as well as the cards with the different English translations are given to the learner. The learner then has to choose from the cards which is the correct answer to the question. Higher points are scored the quicker the learner can give the correct answer.

Option 35 is that the question of the number must be read out in English by the learner himself. The learner must answer the question in Sesotho correctly before they can throw the dice again. Higher points are scored the quicker the learner can give the correct answer. Bonus points will be given for correct pronunciation.

Option 36 is that the question of the number must be read out in English by the learner himself. The learner must answer the question, giving the correct English translation of the answer, before they can throw the dice again. Higher points are scored the quicker the learner can give the correct answer.

Option 37 is that the question of the number must be read out in English by the learner himself. The learner must answer the question, giving the Sesotho word together with the correct English translation of the answer, before they can throw the dice again. Higher points are scored the quicker the learner can give the correct answer. Bonus points will be given for correct pronunciation.

Option 38 is to write down the correct answer in Sesotho on a piece of paper after the question was read by the learner himself in English. If there are any spelling mistakes, the learner has lost his chance for that round. Higher points are scored the quicker the learner can give the correct answer.

Option 39 is to write down the correct English translation of the answer after the question was read by the learner himself in English. If there are any mistakes or spelling mistakes, the learner has lost his chance for that round. Higher points are scored the quicker the learner can give the correct answer.

Option 40 is to write down the correct answer in Sesotho together with the correct English translation of the answer after the question was read by the learner himself in English. If there are any mistakes or spelling mistakes, the learner has lost his

chance. Higher points are scored the quicker the learner can give the correct answer.

Option 41 is to write down the question in Sesotho without any spelling mistakes after the question was read by the learner himself in English. If there are any spelling mistakes, the learner has lost his chance. Higher points are scored the quicker the learner can give the correct answer.

Option 42 is that the question of the number must be read out in English by the learner himself. All the cards with the different Sesotho answers to the questions are given to the learner. The learner then has to choose from the cards which is the correct answer to the question. Higher points are scored the quicker the learner can give the correct answer.

Option 43 is that the question of the number must be read out in English by the learner himself. All the cards with the different English translations of the answers are given to the learner. The learner then has to choose from the cards which is the correct translation to the answer. Higher points are scored the quicker the learner can give the correct answer.

Option 44 is that the question of the number must be read out in English by the learner himself. All the cards with the different Sesotho answers to the questions as well as the cards with the English translations are given to the learner. The learner then has to choose from the cards which is the correct answer to the question. Higher points are scored the quicker the learner can give the correct answer.

Option 45 is that the answer of the number must be read out in Sesotho to the learner by another learner – taking turns to read. The learner must give the question to the answer in Sesotho correctly before they can throw the dice again. Higher points are scored the quicker the learner can give the correct answer. Bonus points will be given for correct pronunciation.

Option 46 is that the answer of the number must be read out in Sesotho to the learner by another learner – taking turns to read. The learner must give the question to the answer, giving the correct English translation of the question, before they can

throw the dice again. Higher points are scored the quicker the learner can give the correct answer.

Option 47 is that the answer of the number must be read out in Sesotho to the learner by another learner – taking turns to read. The learner must give the question to the answer, giving the Sesotho question together with the correct English translation of the question, before they can throw the dice again. Higher points are scored the quicker the learner can give the correct answer. Bonus points will be given for correct pronunciation.

Option 48 is to write down the correct question in Sesotho on a piece of paper after the answer was read to the learner in Sesotho by another learner. If there are any spelling mistakes, the learner has lost his chance for that round. Higher points are scored the quicker the learner can give the correct answer.

Option 49 is to write down the correct English translation of the question after the answer was read to the learner in Sesotho by another learner. If there are any mistakes or spelling mistakes, the learner has lost his chance for that round. Higher points are scored the quicker the learner can give the correct answer.

Option 50 is to write down the correct question in Sesotho together with the correct English translation of the question after the answer was read to the learner in Sesotho by another learner. If there are any mistakes or spelling mistakes, the learner has lost his chance. Higher points are scored the quicker the learner can give the correct answer.

Option 51 is to write down the answer in Sesotho without any spelling mistakes after the answer was read to the learner in Sesotho by another learner. If there are any spelling mistakes, the learner has lost his chance. Higher points are scored the quicker the learner can give the correct answer.

Option 52 is to write down the correct English translation of the answer after the answer was read to the learner in Sesotho by another learner. If there are any mistakes or spelling mistakes, the learner has lost his chance. Higher points are scored the quicker the learner can give the correct answer.

Option 53 is to write down the answer in Sesotho together with the correct English translation of the answer after the answer was read to the learner in Sesotho by another learner. If there are any mistakes or spelling mistakes, the learner has lost his chance. Higher points are scored the quicker the learner can give the correct answer.

Option 54 is that the answer of the number must be read out in Sesotho to the learner by another learner – taking turns to read. All the cards with the different Sesotho questions to the answers are given to the learner. The learner then has to choose from the cards which is the correct question to the answer. Higher points are scored the quicker the learner can give the correct answer.

Option 55 is that the answer of the number must be read out in Sesotho to the learner by another learner – taking turns to read. All the cards with the different English translations of the questions are given to the learner. The learner then has to choose from the cards which is the correct translation to the question. Higher points are scored the quicker the learner can give the correct answer.

Option 56 is that the answer of the number must be read out in Sesotho to the learner by another learner – taking turns to read. All the cards with the different Sesotho questions to the answers as well as the cards with the different English translations are given to the learner. The learner then has to choose from the cards which is the correct question to the answer. Higher points are scored the quicker the learner can give the correct answer.

Option 57 is that the answer of the number must be read out in Sesotho by the learner himself. The learner must give the question to the answer in Sesotho correctly before they can throw the dice again. Higher points are scored the quicker the learner can give the correct answer. Bonus points will be given for correct pronunciation.

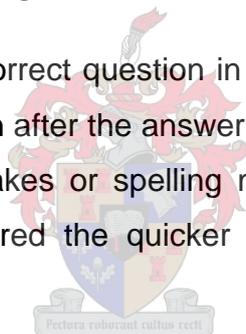
Option 58 is that the answer of the number must be read out in Sesotho by the learner himself. The learner must give the question to the answer, giving the correct English translation of the question, before they can throw the dice again. Higher points are scored the quicker the learner can give the correct answer.

Option 59 is that the answer of the number must be read out in Sesotho by the learner himself. The learner must give the question to the answer, giving the Sesotho question together with the correct English translation of the question, before they can throw the dice again. Higher points are scored the quicker the learner can give the correct answer. Bonus points will be given for correct pronunciation.

Option 60 is to write down the correct question in Sesotho on a piece of paper after the answer was read by the learner himself in Sesotho. If there are any spelling mistakes, the learner has lost his chance for that round. Higher points are scored the quicker the learner can give the correct answer.

Option 61 is to write down the correct English translation of the question after the answer was read by the learner himself in Sesotho. If there are any mistakes or spelling mistakes, the learner has lost his chance for that round. Higher points are scored the quicker the learner can give the correct answer.

Option 62 is to write down the correct question in Sesotho together with the correct English translation of the question after the answer was read by the learner himself in Sesotho. If there are any mistakes or spelling mistakes, the learner has lost his chance. Higher points are scored the quicker the learner can give the correct answer.



Option 63 is to write down the correct English translation of the answer after the answer was read out in Sesotho by the learner himself. If there are any mistakes or spelling mistakes, the learner has lost his chance. Higher points are scored the quicker the learner can give the correct answer.

Option 64 is that the answer of the number must be read out in Sesotho by the learner himself. All the cards with the different Sesotho questions to the answers are given to the learner. The learner then has to choose from the cards which is the correct question to the answer. Higher points are scored the quicker the learner can give the correct answer.

Option 65 is that the answer of the number must be read out in Sesotho by the learner himself. All the cards with the different English translations of the questions are given to the learner. The learner then has to choose from the cards which is the

correct translation to the question. Higher points are scored the quicker the learner can give the correct answer.

Option 66 is that the answer of the number must be read out in Sesotho by the learner himself. All the cards with the different Sesotho questions to the answers as well as the cards with the different English translations are given to answer. Higher points are scored the quicker the learner can give the correct answer.

Option 67 is that the answer of the number must be read out in English to the learner by another learner – taking turns. The learner must give the question to the answer in Sesotho correctly before they can throw the dice again. Higher points are scored the quicker the learner can give the correct answer. Bonus points will be given for correct pronunciation.

Option 68 is that the answer of the number must be read out in English to the learner by another learner – taking turns to read. The learner must give the question to the answer, giving the correct English translation of the question, before they can throw the dice again. Higher points are scored the quicker the learner can give the correct answer.

Option 69 is that the answer of the number must be read out in English to the learner by another learner – taking turns to read. The learner must give the question to the answer, giving the Sesotho word together with the correct English translation of the question, before they can throw the dice again. Higher points are scored the quicker the learner can give the correct answer. Bonus points will be given for correct pronunciation.

Option 70 is to write down the correct question in Sesotho on a piece of paper after the answer was read to the learner in English by another learner. If there are any spelling mistakes, the learner has lost his chance for that round. Higher points are scored the quicker the learner can give the correct answer.

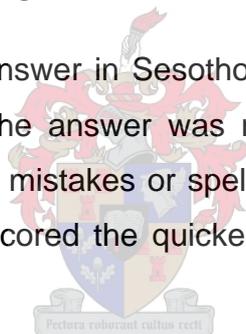
Option 71 is to write down the correct English translation of the question after the answer was read to the learner in English by another learner. If there are any mistakes or spelling mistakes, the learner has lost his chance for that round. Higher points are scored the quicker the learner can give the correct answer.

Option 72 is to write down the correct question in Sesotho together with the correct English translation of the question after the answer was read to the learner in English by another learner. If there are any mistakes or spelling mistakes, the learner has lost his chance. Higher points are scored the quicker the learner can give the correct answer.

Option 73 is to write down the answer in Sesotho without any spelling mistakes after the answer was read to the learner in English by another learner. If there are any spelling mistakes, the learner has lost his chance. Higher points are scored the quicker the learner can give the correct answer.

Option 74 is to write down the correct English translation of the answer after the answer was read to the learner in English by another learner. If there are any mistakes or spelling mistakes, the learner has lost his chance. Higher points are scored the quicker the learner can give the correct answer.

Option 75 is to write down the answer in Sesotho together with the correct English translation of the answer after the answer was read to the learner in English by another learner. If there are any mistakes or spelling mistakes, the learner has lost his chance. Higher points are scored the quicker the learner can give the correct answer.



Option 76 is that the answer of the number must be read out in English to the learner by another learner – taking turns to read. All the cards with the different Sesotho questions to the answers are given to the learner. The learner then has to choose from the cards which is the correct question to the answer. Higher points are scored the quicker the learner can give the correct answer.

Option 77 is that the answer of the number must be read out in English to the learner by another learner – taking turns to read. All the cards with the different English translations of the questions are given to the learner. The learner then has to choose from the cards which is the correct translation to the question. Higher points are scored the quicker the learner can give the correct answer.

Option 78 is that the answer of the number must be read out in English to the learner by another learner – taking turns to read. All the cards with the different Sesotho

questions to the answers as well as the cards with the different English translations are given to the learner. The learner then has to choose from the cards which is the correct answer to the question. Higher points are scored the quicker the learner can give the correct answer.

Option 79 is that the answer of the number must be read out in English by the learner himself. The learner must give the question to the answer in Sesotho correctly before they can throw the dice again. Higher points are scored the quicker the learner can give the correct answer. Bonus points will be given for correct pronunciation.

Option 80 is that the answer of the number must be read out in English by the learner himself. The learner must give the question to the answer, giving the correct English translation of the question, before they can throw the dice again. Higher points are scored the quicker the learner can give the correct answer.

Option 81 is that the answer of the number must be read out in English by the learner himself. The learner must give the question to the answer, giving the Sesotho question together with the correct English translation of the question, before they can throw the dice again. Higher points are scored the quicker the learner can give the correct answer. Bonus points will be given for correct pronunciation.

Option 82 is to write down the correct question in Sesotho on a piece of paper after the question was read by the learner himself in English. If there are any spelling mistakes, the learner has lost his chance for that round. Higher points are scored the quicker the learner can give the correct answer.

Option 83 is to write down the correct English translation of the question after the answer was read by the learner himself in English. If there are any mistakes or spelling mistakes, the learner has lost his chance for that round. Higher points are scored the quicker the learner can give the correct answer.

Option 84 is to write down the correct question in Sesotho together with the correct English translation of the question after the answer was read by the learner himself in English. If there are any mistakes or spelling mistakes, the learner has lost his

chance. Higher points are scored the quicker the learner can give the correct answer.

Option 85 is to write down the answer in Sesotho without any spelling mistakes after the question was read by the learner himself in English. If there are any spelling mistakes, the learner has lost his chance. Higher points are scored the quicker the learner can give the correct answer.

Option 86 is that the answer of the number must be read out in English by the learner himself. All the cards with the different Sesotho questions to the answers are given to the learner. The learner then has to choose from the cards which is the correct answer to the question. Higher points are scored the quicker the learner can give the correct answer.

Option 87 is that the answer of the number must be read out in English by the learner himself. All the cards with the different English translations of the questions are given to the learner. The learner then has to choose from the cards which is the correct translation to the question. Higher points are scored the quicker the learner can give the correct answer.

Option 88 is that the answer of the number must be read out in English by the learner himself. All the cards with the different Sesotho questions to the answers as well as the cards with the English translations of the questions are given to the learner. The learner then has to choose from the cards which is the correct question to the answer. Higher points are scored the quicker the learner can give the correct answer.

This test-item will contribute to the pronunciation of words (oral performance), the writing of words (written performance), listening to Sesotho (aural performance) and to reading Sesotho (reading performance). Thus productive and receptive performance will be tested.

This test-item has information gap as it is requesting information. It also has a low level of problem solving as the learner is looking for a solution of a specified problem with the cards. The test-item does not really have decision making since the learner is not involved in a decision about what to do. The test-item has no opinion

exchange. Narrative would not really be found in this test-item as it is not presenting the opportunity for a short or long story.

Notes:

- Make sure that you help the learners to get accustomed to the test format.
- This test can be divided into 4 levels of proficiency. Roughly represented as follows: level 1 (23%), level 2 (23%), level 3 (28%) and level 4 (26%).

Snakes & Ladders

	Question	Answer
1.	Dumela, monghadi. <i>Good day, sir (mister).</i>	Ee, dumela ntate. <i>Yes, good day man.</i>
2.	O tswa kae? <i>Where do you come from?</i>	Ke tswa Foreisetata. <i>I come from the Free State.</i>
3.	Nomoro ya lebokoso la poso la kgwebo ke mang? <i>What is the postal address of the business?</i>	Nomoro ya lebokoso la kgwebo ke 46, Dalsig, Stellenbosch, 7600. <i>The postal address of the business is 46, Dalsig, Stellenbosch, 7600.</i>
4.	Ke tla fumana eng ha ke boloka tjelele bankeng? <i>What will I gain from putting my money in the bank?</i>	Tjhelete ya hao ha e le sebakeng sena e bolokehile mme e ke ke ya utsuwa ha bonolo. Taba ya bobedi ke hore ho teng tswala eo o tla e fumana ka tjelele eo o e bolokang bankeng. <i>If your money is kept here it is much safer and less likely to be stolen. Secondly, is the interest that you will get on the money in the bank.</i>
5.	O tsohile? <i>How are you? (Did you get up?)</i>	Ee, ke tsohile. <i>Yes, I am fine. (I got up.)</i>
6.	Nomoro ya lebokoso la poso ke mang? <i>What is your postal address?</i>	Nomoro ya lebokoso la poso ke 63, Wepener, 9944 <i>My postal address is 63, Wepener, 9944.</i>
7.	Mosebetsi wa hao ke eng? <i>What is your profession?</i>	Mosebetsi wa ka ke ho rupela. <i>My profession is to teach.</i>
8.	Mosebeletsi o na le mangolo a lokelang na? <i>Does the worker have the necessary qualifications?</i>	Tjhee, ke tla rupela basebeletsi. <i>No, I will train the staff.</i>
9.	Tshireletso ke eng? <i>What is security?</i>	Tsena ke dintho tseo ba ka o amohang tsona ha o sa lefe. <i>These are things that they can take if you do not pay up.</i>
10.	O tsohile hantle? <i>How are you? (Did you get up well?)</i>	Ee, ke tsohile hantle. <i>Yes, I am fine. (I got up well.)</i>

11.	Khotso, mohlankana. <i>Peace, young man.</i>	Ee, khotso ntate. <i>Yes, peace man.</i>
12.	O ka fokotsa dintshenyehelo jwang? <i>How can you cut down on expenses?</i>	O ka boela wa fokotsa dintshenyehelo ka ho tlohela ho reka dintho tseo o hlokomelang hore ha di thuse kgwebo hore e hole. <i>You can continue to cut down on cost expenses by not spending money on things that do not help build the business.</i>
13.	Nomoro ya mohala ke mang? <i>What is the telephone number ?</i>	Nomoro ya mohala ke (051) 5831484. <i>The telephone number is (051) 5831484.</i>
14.	O ile yunibesiting efe? <i>To which university did you go ?</i>	Ke ile yunibesiti ya Foreisetata. <i>I went to the university of the Free Stat</i>
15.	O ka mpoella hore ba tla tseba jwang ke akhaonte sa ka? <i>Can you please explain to me how they will know that it is my account?</i>	Akhaonte ya hao e na le lebitso la hao, nomoro mmoho le hore ke tjelete e kae e teng ho yona. <i>Your account has your name, a number and a record of money that you have in there.</i>
16.	Ke mang a hirang batho? <i>Who is it that employs the people?</i>	Ke nna a hiring batho. <i>It is me who employs the people .</i>
17.	O na le testamente? <i>Do you have a testament ?</i>	Ee, ke na le testamente. <i>Yes, I have a testament.</i>
18.	E se le kgale o eme? <i>Have you been waiting long?</i>	Tjhee. <i>No.</i>
19.	Ee, ho na le mantswe a mararo ao o hlohang o utlwileng borakgwebo ba bua ka ona ha ba bua ho etsa phaello. Ona ke eng? <i>Yes, there are three words you always hear when business people talk about making profit. What are these?</i>	Ona ke: phaello, ka moo thekiso ya thepa e tsamayang ka teng le ditshenyehelo. <i>These are: profit margin, turn over and expenses.</i>
20.	Nomoro ya boitsebiso ke mang? <i>What is your ID number ?</i>	Nomoro ya boitsebiso ke 7108042954056. <i>My ID number is 7108042954056.</i>
21.	Basebetsi-mmoho a bokae? <i>How many colleagues are there?</i>	Ho na le basebetsi-mmoho a mararo. <i>There are three colleagues.</i>
22.	Ke mang a tla sebetsang ka dillo le ngollano? <i>Who will work with complaints and correspondence?</i>	Ké mongodi. <i>It is the secretary.</i>
23.	O kgetha lebenkele jwang? <i>How do you choose a shop?</i>	Ho molemo ho etela mabenkele a mangata mme o shebise mananeo a ona a thekiso. O tla makatswa ke phapang e kgolo eo o tla e bona. <i>It helps to go to different shops and compare prices. You will be surprised to find out how these can differ between places.</i>

24.	O dilemo di kae? <i>How old are you ?</i>	Ke dilemo tse mashome a mabedi le metso e tshelletseng. <i>I am 26.</i>
25.	Lebitso la hao ke mang, monghadi? <i>What is your name, sir (mister) ?</i>	Lebitso la ka ke Willie. <i>My name is Willie.</i>
26.	Ke kopa hore o ntlhalose thekiso. <i>I ask if you please will explain marketing to me.</i>	Thekiso ke karolo ya bohlokwa kgwebong ka nngwe. Kgwebo e ka wa haeba ho se mereo e lokileng ya kgwebo. Bakeng sa hore kgwebo e tswelelle, o tla tshwanela ho rekisa seo o se etsang. <i>Marketing is a very important part of any business. A business can fail if there are no good marketing plans. For a business to be a success, you have to sell what you do.</i>
27.	O hlaetse neng? <i>When were you born ?</i>	Ke hlaetse ka 1971-08-04. <i>I was born in 1971/08/04.</i>
28.	Mosebeletsi o tla saena konteraka neng? <i>When will the worker sign the contract?</i>	O tla saena konteraka ka mora puisano. <i>The worker will sign the contract after the interview.</i>
29.	Ke tla kadima tjehele e bokae? <i>How much money will I borrow?</i>	Kadima tjehele eo o tsebang hore o tla kgona ho e kgutlisa. <i>Borrow an amount that you know you can pay back later.</i>
30.	Tsamaya ka kgotso, monghadi <i>Go in peace, sir (mister).</i>	Sala ka kgotso, ntate. <i>Stay in peace, man.</i>
31.	Palama hantle, mme. <i>Ride well, woman.</i>	Sala ka kgotso, ntate. <i>Stay in peace, man.</i>
32.	O kenile sekolo sefe? <i>To which school did you go ?</i>	Ke kenile sekolo sa Rouxville. <i>Ke kenile sekolo sa Rouxville.</i>
33.	O na le bana? <i>Do you have children ?</i>	Tjhee, ha ke na bana. <i>No, I do not have children.</i>
34.	Ba na le medikale edi le penshene? <i>Do the have medical aid and pension?</i>	Ee, ba na le tsona. <i>Yes, they have it.</i>
35.	O phela jwang, mme? <i>How are you, woman?</i>	Ke phela ha monate, ntate. <i>I am fine, man.</i>
36.	Ke tla dumela ho lefa tjehele neng? <i>When will I agree to pay back the money?</i>	Dumela ho lefa tjehele ka letsatsi leo o tla beng o na le yona, jwalo ka ha o qeta ho amohela. <i>Agree to pay back the money on a day when you know you will have it, like your payday.</i>
37.	Ho na le mekgwa e kae ya ho etsa thekiso? <i>How many different ways are there to do marketing?</i>	Ho na le mekgwa e mengata ya ho etsa hona, jwalo ka ho phatlalatsa. <i>There are many ways that you can do this, like advertising.</i>

38.	Palama ka kgotso, mofumahadi. <i>Ride in peace, mrs.</i>	Sala hantle. <i>Stay well.</i>
39.	Boahi ba hao ke mang? <i>What is your nationality ?</i>	Boahi ba ka ke Leburu. <i>My nationality is a South African.</i>
40.	Phaello ke eng? <i>What is a profit margin?</i>	Ena ke tjelete ya phaello eo o ka e fumanang ka mora ho rekisa. Haeba o rekile apole ya R1,00 mme o e rekisa R2,00, phaello ya hao e tla ba R1,00. <i>This is the amount of profit you can make on the things you sell. If you buy an apple for R1,00 and sell it for R2,00, your profit margin is R1,00.</i>
41.	O batla basebeletsi bafe? <i>What kind of workers do you want?</i>	Ke batla mongodi le basebeletsi ba nkwana. <i>I want a secretary and casual labourers.</i>
42.	O sa phela? <i>How are you?</i>	Ke sa phela. <i>I am fine.</i>
43.	O ya kerekeng efe? <i>Which church do you go to ?</i>	Ke ya kerekeng ya Fora. <i>I go to the Dutch Reformed Church.</i>
44.	Ho phatlalatsa ke eng? <i>What is advertising?</i>	Ho phatlalatsa ho tla o thusa ho rekisa kgwebo ya hao. Phatlalatsa e hlileng e sebetsang haholo ke ya seyalemoya, TV kapa dikoranta. <i>Advertising will help you sell your business. The most effective advertising is on the radio, TV and newspapers.</i>
45.	Fane ya hao ke mang? <i>What is your surname ?</i>	Fane ya ka ke Lombaard. <i>My surname is Lombaard.</i>
46.	O ka mpoella ka diforomo tseo ke lokelang ho di tlatsa ha ke kenya tjelete? <i>Can you explain to me about all the forms I have to fill in if I want to put in money?</i>	Haeba o batla ho kenya tjelete akhaonteng ya hao, hona ho bitswa ka hore ke deposit. O tla lokela ho tlatsa foromo ya deposit ha o etsa hona. <i>If you wish to put money into your account, we call this a deposit. To do this you would need to fill in a deposit slip.</i>
47.	Dipampitshana ke eng? <i>What is flyers?</i>	Na o kile wa nehelwa pampirinyana e phatladitseng diphahlo tse rekiswang ka theolelo kapa tse buang ka ho bulwa ha lebenklele le letjha? Tsena di bitswa flyers. <i>- Have you ever been handed a pamphlet that advertises clothes for sale or those that tell you about the opening of a new shop? These are called flyers.</i>
48.	O qetile sehlopha sefe sekolong? <i>What standard did you pass at school ?</i>	Ke qetile sehlopha se leshome sekolong. <i>I passed matric at school.</i>
49.	Modula-setulo ya kgwebo ke mang? <i>Who is the chairperson of the business?</i>	Modula-setulo ya kgwebo ke Phomolo Seboka. <i>The chairperson of the business is Phomolo Seboka.</i>

50.	Lebitso la kgwebo ke mang? <i>What is the name of the business?</i>	Lebitso la kgwebo ke Ramatheola. <i>The name of the business is Ramatheola.</i>
51.	Basebeletsi ba nkwana ba tla etsang? <i>What will the casual labourers do?</i>	Ba tla thusa ka mesebetsi e sa tlwaelehang le ho bala thepa le ho nka ditaelo. <i>They will help with odd jobs, to take stock and to take orders.</i>
52.	Ke kopa hore o ntlhalose akhoante tse fapaneng bankeng. <i>I ask that you explain the different bank accounts to me.</i>	Ho teng akhaonte tse ngata tse fapaneng tsa ho boloka tjehelete bankeng. <i>There are many different accounts where you can put your money in the bank.</i>
53.	Ho ngotswe eng diforomong ho kadima tjehelete? <i>What is written on the forms to borrow money?</i>	Bala diforomo hantle. Etsa bonnete ba hore dintho tseo le dumellang ka tsona di ngotswe foromong. Ha ho le teng ho hong ho ngotsweng, bala seo hantle. <i>Read the forms carefully. Make sure that the things you agreed on are written on the form. If there is extra writing, read it carefully.</i>
54.	O qetile kgato efe yunibesiting? <i>Which degree did you finish at University?</i>	Ke qetile kgato ya thuta ya tshwaro ya dibuka. <i>I finished my degree in bookkeeping.</i>
55.	O nyetswe na? <i>Are you married ?</i>	Ee, ke nyetswe. <i>Yes, I am married.</i>
56.	O tseba hore kopano e sebetsa jwang? <i>Do you know how a meeting works?</i>	Ee, ke a tseba. <i>Yes, I know.</i>
57.	O ka hatisa dipampitshana? <i>Can flyers be printed?</i>	O ka di hatisa, o ka di etsa photocopy kapa hona ho di ngola ka letsoho. <i>They can be printed, photocopied or even handwritten.</i>
58.	Ka moo, thekiso ya thepa e tsamayang ka teng ke eng? <i>What is turn over?</i>	Ke palo ya thepa eo o e rekisitseng ka mora nako e itseng. Ha o rekisitse thepa e ngata, ho pepenene hore o tla fumana phaello e ngata. Palo ya thepa e rekisitsweng e ntse e batlile e tshwana le phaello. <i>This is the amount of stock you sell over a period. If you sell a lot of stock, your profit will be high. Turnover of goods is directly linked to profits.</i>
59.	Mohlalobi wa dibuka ya kgwebo ke mang? <i>Who is the accountant of the business?</i>	Mohlalobi wa dibuka ya kgwebo ke Ranko Ntamae. <i>The accountant of the business is Ranko Ntamae.</i>
60.	Ho jwang? <i>How are you?</i>	Ho lokile. <i>I am fine.</i>
61.	O sebetsa kae? <i>Where do you work?</i>	Ke sebetsa ka sekolong. <i>I work at the school.</i>

62.	Nomoro ya mohala ya kgwebo ke mang? <i>What is the telephone number of the business?</i>	Nomoro ya mohala ya kgwebo ke (021) 4473835. <i>The telephone number of the business is (021) 4473835.</i>
63.	Ha ke sa utlwisise ho ngotsweng diforomong ? <i>If I do not understand what is written on the forms?</i>	Ha o sa utlwisise, batla motho eo o mo tsebang ya tla o hlosetsa hantle. <i>If you do not understand, look for a person that you know that will explain this to you carefully.</i>
64.	Dipampitshana di ka nehelwa batho kae? <i>Where can flyers be handed out to people?</i>	Hore o tle o fumane tswelelo, o tshwanetse ho kgetha moo o tla isa flyers tsa aho teng. Hona ho tla ya ka kgwebo ya hao le mofuta wa bareki bao o ba batlang. <i>To be most effective, you need to choose where to hand out your flyers. This will depend on your business and the type of customers you want to attract.</i>
65.	Ke mang a tla nka metsotso ya kopano? <i>Who will take the minutes of the meeting?</i>	Ké monghodi. <i>It is the secretary.</i>
66.	Ditshenyehelo ke eng? <i>What are expenses?</i>	Ditshenyelo ke e nngwe ya dintho tse among kgwebo haholo. Ke ditshwenyehelo tsa hao: tse kang ho lefella sepalangwang le ho lefella rente ya lebenkele. <i>Expenses is another thing that affects your profit. Expenses are your costs, like for transport and rent of your shop.</i>
67.	Diinisheale tsa hao ke mang? <i>What are your initials ?</i>	Diinisheale tsa ka ke W. S. <i>My initials are W. S.</i>
68.	Savings akhaonte ke eng? <i>What is a savings account?</i>	Akhaonte ena ke ya bohlokwa haholo ha e la hore o batla ho kenya tjelete le ho e ntsha kgafetsa. <i>Savings accounts are good if you want to put in and take money out of the bank often.</i>
69.	Ke mang a tla ngola lenane tsamaiso? <i>Who will write the agenda?</i>	Ke nna hobane nna ke mong. <i>I will because I am the boss.</i>
70.	Hobaneng ho le theko e boima ho kadima tjelete? <i>Why is it expensive to borrow money?</i>	Ho kadima tjelete o lokela ho e kgutlisa e le ka hodima eo o e kadimmeng. <i>When you borrow money you must pay back more than you borrowed.</i>
71.	Banka ya kgwebo ke efe? <i>Which bank is the bank of the business?</i>	Banka ya kgwebo ke Standard Banka. <i>Standard Bank is the bank of the business.</i>
72.	O hlaetse kae? <i>Where were you born ?</i>	Ke hlaetse Rouxville. <i>I was born in Rouxville.</i>

73.	Ke ka romella diphatlatso tsa ka ka poso? <i>Can I post my adverts?</i>	Ee, ke ka lebaka lang ha o sa romele diphatlatso tsa hao ka poso ho ya dikopanong tse fapaneng le dikgwebong tse ka sebedisang tshebeletso ya hao ka moso. <i>Yes, why not post adverts to the different organisations and businesses who could use your services in the future.</i>
74.	O na le feme? <i>Do you have a factory?</i>	Ee, ke na le feme. <i>Yes, I have a factory.</i>
75.	Motsamaisi ya kgwebo ke mang? <i>Who is the manager (director) of the business?</i>	Motsamaisi ya kgwebo ke Willie Lombaard. <i>The manager of the business is Willie Lombaard.</i>
76.	Ke ka boloka khopi? <i>May I keep a copy?</i>	Ee, boloka khopi ya tumello ya lona e le hore o tle o e hlahise ha ho le teng mathatha. <i>Yes, keep a copy of your agreement in case there is any problems later.</i>
77.	Morekisi wa dinshoreense ya kgwebo ke mang? <i>Who is the insurance agent of the business?</i>	Morekisi wa dinshoreense ya kgwebo ke Motlalepula Mahlasi. <i>The insurance agent of the business is Motlalepula Mahlasi.</i>
78.	Ho ntlafatsa phaello ya hao ho etsuang? <i>What do you do to improve your profit?</i>	Ho na le dintho tse pedi tseo o lokelang ho di etsa bakeng sa ho ntlafatsa phaello ya hao. O ka nna wa nyolla phaello ya hao, ka ho nyolla tijelete ya hao ya thekiso. Theola ditshenyehelo. O ka nna wa fokotsa ditshenyehelo ka ho reka thepa e ngata ka nako e le nngwe. <i>To improve your profit you need to do two things. You can increase your profit margin by increasing your selling price. Reduce your expenses. You can reduce costs by buying in bulk.</i>
79.	Tsela tshweu, mohlankana. <i>Pleasant journey, young man.</i>	Sala hantle, ntate. <i>Stay well, man.</i>
80.	Hape mail-shots di thusa ka eng? <i>For what can mail-shots help too?</i>	Hape mail shots di thusa haeba o batla ho phatlatso hara setjhaba se heno. Ho tsamaya o kenya dipampiri malapeng kaofela ho tla o neha setumo. <i>Mail-shots are also useful if you want to advertise in your local community. Door to door advert drops will give you a lot of publicity</i>
81.	Mohaho o motle na? <i>Is the building nice?</i>	Ee, mohaho o motle. <i>Yes, the building is nice.</i>
82.	O dula kae? <i>Where do you live?</i>	Ke dula ka Wepener. <i>I live in Wepener.</i>

83.	Tjheke akhaonte ke eng? <i>What is a cheque account?</i>	Akhaonte ena ke ya bohlokwa haholo ha e le hore o batla ho lefa melato ya hao ka tjheke. Tjheke ke sekotwana sa pampiri sa bankeng se sebediswang jwalo ka ha eka o sebedisa tjhelete ya hao e leng bankeng. <i>This account is good if you want to pay all your accounts by cheque. A cheque is a piece of paper from the bank that is used as if it is your money in the bank.</i>
84.	Ke o lakaletsa lehlohonolo. <i>I wish you luck.</i>	Ke lebohile haholo. <i>Thank you very much.</i>
85.	Nomoro ya paseporoto ke mang? <i>What is the passport number ?</i>	Nomoro ya paseporoto ke BL0137285. <i>The passport number is BL 0137285.</i>
86.	Tjhelete e ka hodimo eo e o lefang e bitswa eng? <i>What is the extra amount of money called?</i>	Tjhelete e ka hodimo eo o e lefang e bitswa phaello. <i>The extra amount you pay is called interest.</i>
87.	Phomola ka kgotso, mofumahatsana. <i>Rest in peace, miss.</i>	Sala ka kgotso. <i>Stay in peace.</i>
88.	Tsamaya hantle, motswalle. <i>Go well my friend.</i>	Sala hantle, ntate. <i>Stay well, man.</i>
89.	O na le lakasense la ho hweba? <i>Do you have a trading license?</i>	Ee, ke rekile lakansense la ho hweba. <i>Yes, I bought a trading license.</i>
90.	Ke lebohile ka thuso ya hao. <i>Thank you for your help.</i>	Ho lokile. <i>It is all right.</i>
91.	O phela hantle? <i>How are you?</i>	Ee, ke phela hantle. <i>Yes, I am fine.</i>
92.	Ramotlakase ya kgwebo ke mang? <i>Who is the electrician of the business?</i>	Ramotlakase ya kgwebo ke Ntsokolo Faba. <i>The electrician of the business is Ntsokolo Faba.</i>
93.	Molato ke eng? <i>What is the problem?</i>	Molato ke batho ba sebetsang ka lebenkele. <i>The problem is the people who are working in the shop.</i>
94.	Ke bula akhaonte jwang? <i>How do I open an account?</i>	O hloka boitsebiso ya hao mmoho le tjhelete eo o tla bula akhaonte ka yona. Leba moo ho botswang teng mme o bolelle batho ba leng moo hore o batla ho bula akhaonte. <i>You will need your ID and some money that you can open your account with. Go to enquiries and tell the people there that you want to open an account.</i>
95.	Aterese ya hao ke mang? <i>What is your address ?</i>	Aterese y a ka ke 8 van Riet seterata, Wepener. <i>Aterese ya ka ke 8 van Riet seterata, Wepener.</i>
96.	Dumela, mme. <i>Good day, woman.</i>	Ee, dumela ntate. <i>Yes, good day man.</i>

97.	Phaello ya bona e tla ba hodimo haholo neng? <i>When will their interest be high?</i>	- Dibaka tse na tse kadimanang ka tjelele di ipeha kotsi hobane o ka nna wa hloleha ho kgutlisa tjelele ena. Haeba ho le teng kotsi e jwalo, hangata phaello ya bona e ba hodimo haholo. <i>Places which lend money take a risk that you may not pay back the money. If there is a big risk that you will not pay it back then the interest is high.</i>
98.	O sebedisa diforomo tsa kgiro? <i>Do you use employment forms?</i>	Ee, ba tla saena konteraka. <i>Yes, they sign a contract.</i>
99.	Setsha sa kgweo e kae? <i>Where is the piece of ground of the business?</i>	Setsha sa kgwebo e haufi le toropo. <i>The piece of ground is near the town.</i>
100.	Haeba o batla ho ntsha tjelele o tla etsa jwang? <i>If you want to take out money, what will you do?</i>	Ha o ntsha tjelele bankeng re re o etsa withdrawal. Ho teng foromo eo o lokelang ho e tlatsa bakeng sa hona. <i>When you take money out of the bank, we say you make a withdrawal. There is a form you need to fill in for this.</i>

3. Sound Snap

The learners have cards – on one set of cards the Sesotho word is written and on the other set of the cards the English word is written. The cards are shuffled before the learners can start to play.



Option 1 is to deal the set of cards with the Sesotho words to all the learners with the Sesotho words facing the bottom. The learners are not allowed to look at their cards at all, during the whole game. The learners are now, taking turns, putting the cards on one pile with the Sesotho words facing the top. If there are two cards been put down on the pile right after each other of which the sound of pronunciation sounds the same, the first learner to put his hand on the deck of cards and yell “snap”, is the person to take the pile of cards to himself. If the learner was mistaken and the pronunciation of the words are not similar, the deck of cards are divided between the other learners. The winner of the game is to have all the cards to himself. If one learner is left with no cards at all, he is out of the game.

Option 2 is to deal the set of cards with the English words to all the learners with the English words facing the bottom. The learners are now, taking turns, putting the cards on one pile with the English words facing the top. If there are two cards been

put down on the pile right after each other of which the sound of pronunciation of the English word translated into Sesotho, sounds the same, the first learner to put his hand on the pile of cards and yell “snap”, is the person to take the pile of cards to himself. If the learner was mistaken and the Sesotho pronunciation of the words are not similar, the deck of cards are divided between the other players.

Option 3 is to deal the Sesotho cards together with the English cards to all the learners (shuffled). The learners are now, taking turns, putting the cards on one pile with the English or Sesotho words facing the top. If there are two cards been put down on the pile right after each other of which the sound of pronunciation of the English word translated into Sesotho or the pronunciation of the Sesotho word sounds the same, the first learner to put his hand on the pile of cards and yell “snap”, is the person to take the pile of cards to himself. If the learner was mistaken and the Sesotho pronunciation of the words are not similar, the deck of cards are divided between the other learners.

Option 4 is to give the set of cards of the Sesotho words to one learner. The learner must shuffle the Sesotho words, cards facing the top, and put all the words that sounds the same when pronounced, together. The learner to put the cards together correctly in the shortest time will be the winner.

Option 5 is to give the set of cards of the English words to one learner. The learner must shuffle the English words, cards facing the top, and put all the English words that sounds the same when translated and pronounced in Sesotho, together. The learner to put the cards together correctly in the shortest time will be the winner.

Option 6 is to give both sets of cards, Sesotho and English words to one learner. The learner must shuffle the Sesotho and English words, cards facing the top, and put all the English words that sounds the same when translated and pronounced in Sesotho, and all the Sesotho words that sounds the same when pronounced, together. The learner to put the cards together correctly in the shortest time will be the winner.

Option 7 is to have a learner reading the set of cards with the Sesotho words to all the learners – taking turns. The learners are now listening if there two cards (words) read right after each other, that sounds the same. The first learner to clap his hands

together, is the person to take the pile of cards to himself. If the learner was mistaken and the pronunciation of the words are not similar, the deck of cards are divided between the other learners. The winner of the game is to have all the cards to himself. If one learner is left with no cards at all, he is out of the game.

Option 8 is to have a learner reading the set of cards with the English words to all the learners – taking turns. The learners are now listening if there are two cards (words) read right after each other, of which the sound of pronunciation of the English word translated into Sesotho, sounds the same. The first learner to clap his hands together, is the person to take the pile of cards to himself. If the learner was mistaken and the pronunciation of the words are not similar, the deck of cards are divided between the other learners. The winner of the game is to have all the cards to himself. If one learner is left with no cards at all, he is out of the game.

Option 9 is to have a learner reading both sets of cards, shuffled, with the Sesotho and English words, to all the learners – taking turns. The learners are now listening if there are two cards (words) read right after each other, that sounds the same if pronounced in Sesotho or which the sound of pronunciation of the English word translated into Sesotho, sounds the same. The first learner to clap his hands together, is the person to take the pile of cards to himself. If the learner was mistaken and the pronunciation of the words are not similar, the deck of cards are divided between the other learners. The winner of the game is to have all the cards to himself. If one learner is left with no cards at all, he is out of the game.

Option 10 is to have a learner read out the set of cards of the Sesotho words to the other learners – taking turns. The learners must write down all the words and then shuffle the Sesotho words, putting all the words that sounds the same when pronounced, together. The learner to put the words together correctly in the shortest time will be the winner.

Option 11 is to have a learner read out the set of cards of the English words to the other learners – taking turns. The learners must write down all the words and then shuffle the English words, putting all the English words that sounds the same when translated and pronounced in Sesotho, together. The learner to put the words together correctly in the shortest time will be the winner.

Option 12 is to have a learner read out both sets of cards, Sesotho and English words to the other learners – taking turns. The learners must write down all the words and then shuffle the Sesotho and English words, putting all the Sesotho and English words that sounds the same when translated and pronounced in Sesotho, and all the Sesotho words that sounds the same when pronounced, together. The learner to put the words together correctly in the shortest time will be the winner.

Notes:

- Make sure that you help the learners to get accustomed to the test format.
- This test can be divided into 4 levels of proficiency. Roughly represented as follows: level 1 (25%), level 2 (25%), level 3 (35%) and level 4 (15%).

Sound Snap

1.	Monghadi	(Sir)	-	Mongodi	(Secretary)
2.	Phela	(Live)	-	Tshela	(Road)
3.	Dula	(Sit)	-	Kula	(Sick)
			-	Bula	(Open)
4.	Etsa	(Do)	-	Sebetsa	(Work)
5.	Toropo	(Town)	-	Nomoro	(Number)
6.	Kgato	(Degree)	-	Molato	(Problems)
7.	Kgafetsa	(Often)	-	Lakaletsa	(Wish)
8.	Tjheke	(Cheque)	-	Kereke	(Church)
9.	Tswa	(Come from)	-	Bitswa	(Called)
10.	Dumela	(Good day)	-	Rupela	(Teach)
			-	Lokela	(Need)
			-	Amohela	(Pay)
11.	Mangolo	(Papers)	-	Bonolo	(Easily)
12.	Mohala	(Phone)	-	Tswala	(Interest)
			-	Sala	(Stay)
			-	Bala	(Read)
13.	Bana	(Children)	-	Fumana	(Get)
14.	Araba	(Answer)	-	Taba	(Story, subject)
15.	Hobaneng	(Why)	-	Fapaneng	(Different)
16.	Kadima	(Borrow)	-	Hodima	(More / above)

17.	Mang	(Who)	-	Jwang	(How)
18.	Tlase	(Low)	-	Ramotlakase	(Electrician)
19.	Poso	(Post)	-	Lebokoso	(Box)
20.	Kopano	(Meeting)	-	Puisano	(Interview)
21.	Bona	(See)	-	Kgona	(Able)
22.	Nyetswe	(Married)	-	Ngotswe	(Written)
23.	Motswalle	(Friend)	-	Kgale	(Long time)
24.	Boitsebiso	(ID)	-	Thekiso	(Marketing)
25.	Phatlalatsa	(Advertising)	-	Tlatsa	(Fill in)
26.	Ditaelo	(Orders)	-	Theollelo	(Sale)
			-	Phaello	(Profit, Interest)
27.	Rekisa	(Sell)	-	Hatisa	(Printed)
28.	Thuta	(Study)	-	Mofuta	(Type)
29.	Jwalo	(Like)	-	Palo	(Amount)
30.	Mong	(Boss)	-	Among	(Affects)
31.	Botsa	(Ask)	-	Fokotsa	(Cut down)

4. Crossword Puzzle

The puzzle can be given to the students without anything written on it. The directions can be given, for example: A5 – A12 is the Sesotho word for an account. The whole crossword puzzle can be filled in by using this way.

Option 1 is to give the directions (A5 – A12) along with the English word written on paper. The Sesotho word must be completed in the crossword puzzle. The whole crossword puzzle is to be completed in this way. The learner who completes his crossword puzzle first wins.

Option 2 is to give the directions (A5 – A12) along with the English word written on paper. The Sesotho word is to be completed in the crossword puzzle. Half of the crossword puzzle is already completed; the other half is to be completed in this way. Here, the half of the crossword puzzle that is already completed, can differ every time. The level of difficulty can be changed through the words that are already given. One might give all the difficult words already filled in for the lower level of proficiency. From there, onwards the options are immensely. The learner who completes his crossword puzzle first wins.

The third option is to give the crossword puzzle filled in and to ask the student to identify all the words that he recognises. The learner who is finished first, wins.

Option 4 is to give the crossword puzzle filled in and to ask the student to identify 101 words. The learner who is finished first, wins.

Option 5 is to give the crossword puzzle filled in and to ask the student to identify all the words that he recognises. The student then has to write down all the English translations of the words identified. The learner who is finished first, wins.

Option 6 is to give the crossword puzzle filled in and to ask the student to identify 101 words. The student then has to write down all the English translations of all the 101 words.

Option 7 is to give the learner the directions (A5 – A12) along with the English word, read by another learner. The Sesotho word must be completed in the crossword puzzle. The whole crossword puzzle is to be completed in this way. The team who completes their crossword puzzle first wins.

Option 8 is to give the learner the directions (A5 – A12) along with the English word, read by another learner. The Sesotho word is to be completed in the crossword puzzle. Half of the crossword puzzle is already completed; the other half is to be completed in this way. Here, the half of the crossword puzzle that is already completed can differ every time. The level of difficulty can be changed through the words that are already given. One might give all the difficult words already filled in for the lower level of proficiency. From there, onwards the options are immensely. The team who completes their crossword puzzle first wins.

Option 9 is to give the learner all the English words, read to him by another learner. The crossword puzzle is fully filled in. The learner must identify all the Sesotho words (101) on the crossword puzzle.

Option 10 is to give the directions (A5 – A12) along with the English word written on paper to a learner. The learner then has to explain in English to the other learner what the Sesotho word is without using the English word on the piece of paper, giving the student the directions (A5 – A12) of the crossword puzzle. The Sesotho word

must be completed in the crossword puzzle. The whole crossword puzzle is to be completed in this way. The team who completes their crossword puzzle first, wins.

Option 11 is to give the directions (A5 – A12) along with the English word written on paper to a learner. The learner then has to explain in English to the other learner what the Sesotho word is without using the English word on the piece of paper, giving the student the directions (A5 – A12) of the crossword puzzle. The Sesotho word is to be completed in the crossword puzzle. Half of the crossword puzzle is already completed; the other half is to be completed in this way. Here, the half of the crossword puzzle that is already completed, can differ every time. The level of difficulty can be changed through the words that are already given. One might give all the difficult words already filled in for the lower level of proficiency. From there, onwards the options are immensely. The team who completes their crossword puzzle first, wins.

Option 12 is to give the directions (A5 – A12) along with the Sesotho word written on paper to a learner. The learner then has to explain in Sesotho to the other learner what the Sesotho word is without using that specific Sesotho word on the piece of paper, giving the student the directions (A5 – A12) of the crossword puzzle. The Sesotho word must be completed in the crossword puzzle. The whole crossword puzzle is to be completed in this way. The team who completes their crossword puzzle first, wins.

Option 13 is to give the directions (A5 – A12) along with the Sesotho word written on paper to a learner. The learner then has to explain in Sesotho to the other learner what the Sesotho word is without using that specific Sesotho word on the piece of paper, giving the student the directions (A5 – A12) of the crossword puzzle. The Sesotho word must be completed in the crossword puzzle. The whole crossword puzzle is to be completed in this way. The team who completes their crossword puzzle first, wins.

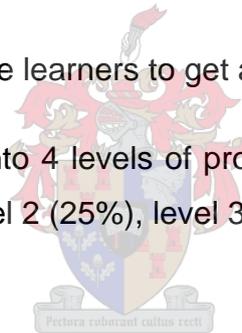
Option 14 is to give the directions (A5 – A12) along with the Sesotho word written on paper to a learner. The learner then has to explain in Sesotho to the other learner what the Sesotho word is without using that specific Sesotho word on the piece of paper, giving the student the directions (A5 – A12) of the crossword puzzle. The

Sesotho word is to be completed in the crossword puzzle. Half of the crossword puzzle is already completed; the other half is to be completed in this way. Here, the half of the crossword puzzle that is already completed, can differ every time. The level of difficulty can be changed through the words that are already given. One might give all the difficult words already filled in for the lower level of proficiency. From there, onwards the options are immensely. The team who completes their crossword puzzle first, wins.

Option 15 is to give the crossword puzzle filled in to a learner. The other learner has all the Sesotho words written on paper. The learner then has to explain each word in Sesotho to the other learner what the Sesotho words are without using that specific Sesotho words on the paper. The learner has to identify 101 words. The team who is finished first, wins.

Notes:

- Make sure that you help the learners to get accustomed to the test format.
- This test can be divided into 4 levels of proficiency. Roughly represented as follows: level 1 (30%), level 2 (25%), level 3 (20%) and level 4 (25%).



Crossword Puzzle

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
1	T	T	S		H	T	B	P	A	M	P	I	R	I	N	F	A	P	A	N	E	N	G
2	H	S	E	P	A	L	A	N	G	W	A	N	G	P	E	K	M	A	M	A	N	G	L
3	U	H	B	H	N	A	L	E	F	A	S		L	E	N	G	O	L	O	K	I	L	E
4	S	I	E	O	T	T	A	B	A	L	E	F	E	N	G	I	R	A	N	T	S	H	A
5	A	R	D	M	L	S	T	D	D	I	P	O	T	S	O	R	E	M	G	K	O	P	A
6	K	E	I	O	E	A	H	I	M	P	O	K	S	H	L	O	K	A	R	A	B	A	
7	H	L	S	L	H	W	E	B	A	O	R	O	O	E	A	M	I	K	E	I	L	E	D
8	A	E	A	O	M	O	P	U	T	S	O	T	H	N	K	A	S	A	K	O	T	S	I
9	O	T	S	W	A	H	A	K	H	O	T	S	O	E	A	D	I	T	A	E	L	O	I
10	N	S	T	S	E	L	A	A	A	B	O	A	H	I	S	Y	A	S	T	E	N	G	N
11	T	O	H	E	L	M	R	E	T	O	K	G	A	L	E	K	G	W	E	B	O	D	I
12	E	K	E	T	E	O	S	E	H	L	O	P	H	A	N	G	B	A	R	E	K	I	S
13	N	K	K	U	B	H	N	K	A	O	R		L	B	S	O	A	L	E	K	A	F	H
14	A	G	O	M	O	A	L	A	N	K	A	N	A	O	E	N	T	A	S	E	J	O	E
15	K	E	B	O	H	H	O	K	W	A	N	K	E	N	Y	A	H	A	E	R	E	R	A
16	O	T	U		I	O	N	N	A	N	T	A	T	E	N	G	O	K	K	E	N	O	L
17	P	H	A	T	L	A	L	A	T	S	A		S	E	T	S	H	A	O	K	O	M	E
18	K	A	E	T	E	S	T	A	M	E	N	T	E	E	T	S	W	E	L	E	L	O	H
19	Q	E	T	I	L	E	B	A	S	E	B	E	T	S	I	-	M	M	O	H	O		O

Crossword Puzzle



	Sesotho word		English translation		Placing
1.	Akhaonte	-	Account	-	A5 – A12
2.	Among	-	Affects	-	S1 – S5
3.	Araba	-	Answer	-	R6 – V6
4.	Aterese	-	Address	-	S9 – S5
5.	Bala	-	Read, Count	-	G1 – G4
6.	Banka	-	Bank	-	E13 – I13
7.	Bareki	-	Customers	-	Q12 – V12
8.	Basebetsi-mmoho	-	Colleagues	-	G19 – U19
9.	Batho	-	People	-	Q12 – Q16
10.	Beke	-	Week	-	T11 – T14
11.	Boahi	-	Nationality	-	J10 – N10
12.	Bohlokwa	-	Good	-	C15 – J15
13.	Boloka	-	Put	-	J10 – J15

14.	Bua	-	Speak	-	C15 – C17
15.	Dibuka	-	Books	-	H5 – H10
16.	Diforomo	-	Forms	-	V11 – V18
17.	Diinisheale	-	Initials	-	W7 – W17
18.	Dipotso	-	Questions	-	I5 – O5
19.	Ditaelo	-	Orders	-	P9 – V9
20.	Ee	-	Yes	-	M18 – N18
21.	Eng	-	What	-	N16 – P16
22.	Fapaneng	-	Different	-	P1 – W1
23.	Fokotsa	-	Cut down	-	L4 – L10
24.	Ha	-	If	-	F9 – G9
25.	Hantle	-	Fine, well	-	E1 – E6
26.	Hae	-	His / Hers	-	Q15 – S15
27.	Hao	-	Your	-	Q17 – S17
28.	Hlaetse	-	Born	-	M12 – M18
29.	Hloka	-	Need	-	N6 – R6
30.	Ho	-	To	-	W18 – W19
31.	Hweba	-	Trade	-	E7 – I7
32.	Ile	-	Go	-	T7 – V7
33.	Ka	-	In	-	R16 – R17
34.	Kae	-	Where	-	A18 – C18
35.	Kajeno	-	Today	-	U12 – U17
36.	Ke	-	I	-	R7 – S7, B12 – C12
37.	Kenya	-	Put in	-	L15 – P15
38.	Kereke	-	Church	-	T13 – T18
39.	Kgale	-	Long time	-	K11 – O11
40.	Kgetha	-	Choose	-	B13 – B18
41.	Kgiro	-	Employment	-	P2 – P6
42.	Kgona	-	Can, Be able	-	P11 – P15
43.	Kgwebo	-	Business	-	P11 – U11
44.	Khotso	-	Peace	-	H9 – M9
45.	Kopa	-	Please	-	T5 – W5
46.	Koranta	-	Newspaper	-	K11 – K17
47.	Kotsi	-	Risk	-	S8 – W8
48.	Labone	-	Thursday	-	N11 – N16
49.	Lakasense	-	License	-	O6 – O14
50.	Le	-	You	-	W2 – W3

51.	Lebohile	-	Thank you	-	E11 – E18
52.	Lefa	-	Pay	-	G3 – J3
53.	Lefeng	-	Pay	-	J4 – O4
54.	Leka	-	Try	-	R13 – U13
55.	Lengolo	-	Letter	-	M3 – S3
56.	Letsoho	-	Hand	-	M3 – M9
57.	Lokile	-	Fine	-	R3 – W3
58.	Makatswa	-	Surprised	-	R5 – R12
59.	Mang	-	What	-	S2 – V2
60.	Mathatha	-	Problems	-	I6 – I13
61.	Mohaho	-	Building	-	F11 – F16
62.	Mohlankana	-	Young man	-	D14 – M14
63.	Moputso	-	Salary	-	E8 – K8
64.	Morekisi	-	Sales person, agent	-	Q2 – Q9
65.	Nako	-	Time	-	A13 – A16
66.	Neng	-	When	-	O1 – O4
67.	Ngola	-	Write	-	O3 – O7
68.	Nka	-	Take	-	N8 – P8
69.	Nna	-	I	-	G16 – I16
70.	Ntate	-	Man	-	J16 – N16
71.	Ntsha	-	Take out	-	S4 – W4
72.	Palama	-	Ride	-	R1 – R6
73.	Pampiri	-	Paper	-	H1 – N1
74.	Paseporoto	-	Passport	-	K1 – K10
75.	Penshene	-	Pension	-	N2 – N9
76.	Phatlalatsa	-	Advertise	-	A17 – K17
77.	Phomolo	-	Leave	-	D2 – D8
78.	Poso	-	Postal, post	-	J6 – J9
79.	Qetile	-	Finish, pass	-	A19 – F19
80.	Re	-	We	-	G11 – H11
81.	Reka	-	Buy	-	S6 – S9
82.	Roma	-	Send	-	P5 – P8
83.	Sebedisa	-	Use	-	C1 – C8
84.	Sehlopha	-	Standard	-	G12 – N12
85.	Sekolo	-	School	-	S14 – S19
86.	Sepalangwang	-	Transport	-	B2 – M2
87.	Setsha	-	Piece of ground	-	M17 – R17

88.	Setumo	-	Publicity	-	D10 – D15
89.	Taba	-	Story, subject	-	F4 – I4
90.	Teng	-	Here, Fine	-	S10 – V10
91.	Testamente	-	Testament	-	D18 – M18
92.	Theko	-	Expensive	-	C10 – C14
93.	Thepa	-	Stock	-	G5 – G9
94.	Thusa	-	Help	-	A1 – A5
95.	Tlatsa	-	Fill in	-	F1 – F6
96.	Tsela	-	Road	-	C10 – G10
97.	Tshireletso	-	Security	-	B1 – B11
98.	Tswa	-	Come	-	B9 – E9
99.	Tswala	-	Interest	-	R9 – R14
100.	Tswelelo	-	Effective	-	O18 – V18
101.	Ya	-	Go	-	P10 – Q10

5. Role – play

Today you speak to a woman on the things of the bank. You talk about a savings account and a cheque account. You tell her when to use which account. You explain to her what to do to open an account, put in some money and withdraw some money. You tell her what she will gain from putting her money in the bank. Then again you explain nicely to her how they will know which account is her account.

This is a task for group work.

The others in the group must watch for certain grammatical or discourse elements as the roles are acted out. Notes must be made by them during the role–play and compared afterwards. Feedback will be given from the class, each learner will have a chance, which will be accompanied by a discussion by the group of learners of what was seen and heard.

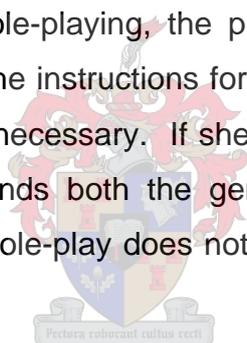
The role–play will be acted out again, where the learners will be trying to accomplish the points of discussion. This might be repeated as many times as the group has new ideas and new feedback.

According to Brown (1994), role-play minimally involves (a) giving a role to one or two members of the group and (b) assigning an objective or purpose that participants must accomplish.

Bygate, Skehan and Swain (2001) asserts that the learner is asked to take on a particular role and to imagine himself in that role in a particular situation. He has to converse with the interviewer in a way that is appropriate to the role and the situation given.

The learner is given a set of instructions, just before the test, that explain in simple language exactly what he is supposed to do. For more confident learners, these instructions may be expressed in terms of the general situation. Alternatively, the instructions may be made more specific, to give the learner more direction and to elicit more comparable language from each learner.

If learners are unfamiliar with role-playing, the procedure and purpose should be explained well beforehand, and the instructions for each role-play should be given in writing, in the native language if necessary. If she is in any doubt, the tester should check that the learner understands both the general procedure and the specific instructions. It is important that role-play does not become a test of comprehension of instructions.



Bygate, Skehan and Swain (2001) states that a possible problem is personal reluctance to participate. Role-playing by definition implies pretending even in a small way, to be someone other than you really are. Some people can do this more easily than others can. In some cultures, role-playing may be seen as unusual and therefore unsettling behaviour, especially in an educational context. Some individuals are personally unhappy about pretending to be someone else, and any attempt at persuasion may only increase this unease. Therefore, there are both cultural and individual differences between people in their ability to role-play which do not reflect differences in language proficiency.

Notes:

- Make sure that you help the learners to get accustomed to the test format.
- This test can be divided into 4 levels of proficiency. Roughly represented as follows: level 1 (20%), level 2 (25%), level 3 (30%) and level 4 (25%).

6. Games

Each member in the group gets a turn to be a specific employee at the small business corporation. The member decides secretly who he / she will be. All the other members of the group, taking turns, asks a yes/no question to that member to try to find out who that member is as an employee. Each member has 5 yes/no questions that he / she can ask the “employee”. Each member must ask 5 questions, taking turns. All the questions must be answered by the “employee” by either saying yes or no. No question may be repeated otherwise you lose 5 points immediately. Each member must write down who he / she decided the employee is. For a correct answer you score 5 points. For a correct grammatical or discourse mistake that the class could point out when you asked your question, you lose 1 point.

According to Brown (1994) a game could be any activity that formalizes a technique into units that can be scored in some way.

Guessing games are more common language classroom activities.

7. Simulation

Each person has been assigned an occupation as an employee in the small business corporation. One person in the group must be fired. The group must decide who will be the person to leave the small business. Each person gets a chance to explain why he / she must not be the person who gets fired. Each member in the group writes down on a piece of paper who is to get fired, along with a reason. All the pieces of paper is thrown in a hat. Each member of the group, taking turns, picks a piece of paper out of the hat and reads it to the class. The employee with the most votes is the employee to get fired.

According to Brown (1994) simulations usually involve a more complex structure and often larger groups where the entire group is working through an imaginary situation as a social unit, the object of which is to solve some specific problem.

8. Drama

Subject:

Today you talk to a miss about marketing. You talk about the ways of marketing. You talk about advertising, to talk and to write flyers. You talk about where the flyers can be handed out and that you can post it or you can put it at every door.

The group of learners are divided into small groups. Each group has to prepare their own dramatization of some event using the subject above, writing the script and rehearsing the scene as a group.

According to Brown (1994) drama is a more formalized form of role-play or simulations with a preplanned story line and script. Sometimes small groups may prepare their own short dramatization of some event, writing the script and rehearsing the scene as a group. This may be more commonly referred to as a 'skit'. Longer, more involved dramatic performances have been shown to have positive effects on language learning, but they are time consuming and rarely can form part of a typical school curriculum.

9. Projects

Divide the group of learners into 3 groups. Each group is assigned 1 of the 3 subjects below.

Group A: Profit margin.

What is profit margin, how can you improve profit margin and why.

Group B: Turn over.

What is turn over, how can you improve turn over and why.

Group C: Expenses.

What are expenses, how can you reduce expenses and why.

According to Brown (1994) projects are for all learner groups and perhaps learners can greatly benefit from hands-on approaches to language. He also notes that certain projects can be rewarding indeed.

10. Interview

Today you talk to a sir and you ask him many questions. You start by asking his name, surname, initials, where he stays and where he comes from. You ask him his postal address, his telephone number, his address, his ID number and his passport number too. You ask where he was born, how old he is, when he was born and if he is married and has children. You ask him about his nationality, church, university, degree, work and testament.

According to Brown (1994) interview is a popular activity for pair work, but also appropriate for group work. Interviews are useful at all levels of proficiency. At the lower levels, interviews can be structured, in terms of both the information that is sought and the grammatical difficulty and variety. At the higher levels, interviews can probe facts that are more complex, opinions, ideas and feelings.

According to Underhill (1987) the interview is the most common of all oral tests; for many people, it is the only kind of oral test. Its direct, face-to-face exchange between learner and interviewer. It follows a pre-determined structure, but still allows both people a degree of freedom to say what they genuinely think. It therefore falls between two other techniques described: Discussion / conversation and Question and answer. The distinctions between these three are frequently blurred in practice; but they are worth making clear in principle.

Underhill (1987) asserts that compared with discussion / conversation, an interview is structured. The interviewer sets out to find out certain things about the learner and to get answers to certain questions. She maintains firm control, and keeps the initiative as well; whatever the learner says is more or less direct response to her questions or statements. However, the learner still has the freedom to answer as he likes, or to

develop his comments and opinions. When he has finished his answer or his comment, it is then up to the interviewer to make the next move; to develop the topic further or raise a new one.

The questions and topics raised by the interviewer are chosen for their success in eliciting a representative sample of the learner's speech. Interviewers usually have a prepared list of written or memorised questions to ask, or topics to bring up. This mental or written list will contain quite a wide variety of questions and topics in order to avoid constant repetition and possible compromise. The final choice of topics and questions used will be left to the interviewer to decide during the course of the interview.

Compared with question and answer, however, an interview is more authentic; it has consistency and a relevance that stretches over more than one question or comment. There may be several topics raised in an interview, but each is explored in enough detail, with follow-up questions and prompting, to allow the learner to develop it and to show his proficiency, rather than just giving a straight answer to a straight question.

Underhill (1987) states that oral interviews often fail to discriminate effectively at higher levels. The oral interview technique is well suited for testing learners at the intermediate level and below.

11. Brainstorming

Divide the group of learners into small groups. Ask the small groups to brainstorm the characteristics of a successful small business. The groups' task would be to make a composite list of everything they can think of, without evaluating it.

The ideas of each group will be shared with the rest of the class.

After this, new groups will be assigned and they will discuss and evaluate the points made.

According to Brown (1994) brainstorming is a technique whose purpose is to initiate some sort of thinking process. It gets students' "creative juices" flowing without

necessarily focusing on specific problems or decisions or values. In brainstorming, no discussion of the relative merits of a thought takes place; everything and anything goes. This way, all ideas are legitimate, and students are released to soar the heights and plum the depths, as it were, with no obligation to defend a concept. In whatever follow-up brainstorming you plan, at that point evaluation and discussion can take place.

12. Jigsaw

Each member of a group is given an application form, on each form different information is provided. As the students ask each other questions (without showing anyone their own application form), they must eventually complete all the information on the form. The students must go from student to student to student asking questions.

The other option is to take task 2 (questions about the business) and cut it up into little strips. These strips of information must be shuffled. Each student is then given a strip. The goal is for the students to determine where each of their strips of information belongs in the whole context. Once their position is determined they must stand in their position so that the information is reconstructed and makes sense.

Jigsaw techniques are a special form of information gap.

13. Form-filling

Underhill (1987) states that the learner and interviewer work together to fill in a form. The questions usually concern the learner's personal details, professional situation or language needs. Either the interviewer or the learner may actually write in the answers; if the learner does it, the test will take longer and the learner will obviously feel his writing as well as his speaking skills is being tested.

This activity is authentic. Filling in forms is something we have to do quite often in everyday life. Sometimes we fill in the form ourselves, sometimes the official person we are dealing with fills in the form, asking questions that may need to be discussed before deciding the best way to phrase the answer.

Form-filling is also communicative in that it elicits previously unknown information for an apparent purpose. It allows the interviewer, as well as acting as recording clerk, to exploit the information for discussion purposes, to elicit narration or description, or to find areas of common interest. By directing attention to the form, the interviewer makes the test an opportunity for co-operation rather than competition; learner and interviewer work together to complete the form instead of the interviewer trying to catch the learner out with trick questions, and the learner trying to avoid them.

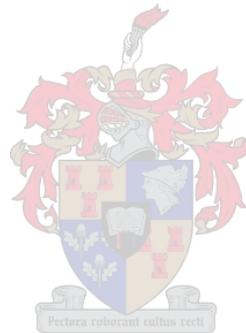
Several notes on specific purpose Sesotho for personnel in small business corporation

This program is:

- Based on a needs analysis in terms of criteria, content and contexts
- As authentic as possible with the goal of measuring real-world activities
- Has collaborative elements that stimulate communicative interactions
- Contextualised and complex
- Integrates skills with content
- Can be appropriate in terms of number, timing, and frequency of assessment
- Can be generally non-intrusive, that is, be aligned with the daily actions in the language classroom
- The rating scale can be based on appropriate categories of language learning and development
- The rating scale is based on appropriate breadth of information regarding learner performance abilities
- The rating scale is based on appropriate standard that are both authentic and clear to students
- To enhance the reliability and validity of decisions as well as accountability, performance assessment is combined with other methods of gathering information (for instance, self-assessments, portfolio's, conferences, classrooms behaviours, and so forth)
- The uses to which the resulting scores/information will be put is defined
- The criteria by which a given performance will be judged is defined

- Designed in terms of type of performances that will be required
- Designed in terms of knowledge, skill, and content focus, and how they will be elicited
- Designed in terms of types of discourse, subject matter, and authenticity of tasks
- Type of score needed is described
- People who will rate the performance can be described
- Performance criteria that will be applied is described
- Score recording method to be used is described
- Directions are clearly given for each performance to each student
- Physical conditions (in terms of space, light, noise, temperature, etc.) can be the same for each student's performance
- Administration conditions (in terms of timing, equipment, realia, language stimulus, etc.) can be the same for each student's performance
- The results can be recorded
- The results can be reported to students and other interested parties in a manner that is thorough, easily understood, and conducive to productive change (e.g., in terms of emphasis on an individual's future studies)
- The validity of the performance assessment procedures as well as the resulting decisions are considered
- The results can be summarised and interpreted in view of the types of decision involved
- Compensating for negative washback (i.e., any negative effects of testing on teaching, curriculum development, or educational policies)
- Documenting and encouraging critical thought, creativity and self-reflection
- Aligning classroom assessment and instructional activities with authentic, real-life activities
- Showing students' strengths and weaknesses in detailed and real-world terms
- Measure abilities to respond to real-life language tasks

- Create more accurate assessments of student's knowledge and ability that traditional multiple-choice tests
- Predict students' abilities in future, real-world situations
- Be more valid than traditional tests in terms of predicting students' abilities to use language in the future real-life situations



BIBLIOGRAPHY

- ALDERSON, CHARLES J., CLAPHAM, CAROLINE, WALL, DIANNE. 1995. **Language Test Construction and Evaluation**. Cambridge University Press.
- BACHMAN LYLE F., DAVIDSON, FRED., MILANOVIC, MICHAEL. 1996. **The use of test method characteristics in the content analysis and design of EFL proficiency tests**. Arnold.
- BACHMAN, LYLE F., PALMER, ADRIAN S. 1996. **Language Testing in Practice: Designing and Developing Useful Language Tests**. Oxford University Press.
- BACHMAN, LYLE F. 1990. **Fundamental Considerations in Language Testing**. Oxford University Press.
- BROWN, H. DOUGLAS. 1994. **Principles of Language Learning and Teaching**. San Francisco State University.
- CHALHOUB-DEVILLE, MICHELINE. 1997. **Theoretical models, assessment frameworks and test construction**. University of Minnesota.
- CROOKES, GRAHAM; GASS, SUSAN M. 1993. **Tasks and Language Learning Integrating Theory and Practice**. Multilingual Matters Ltd.
- DAVIES, ALAN. 1990. **Principles of Language Testing**. Basil Blackwell.
- DOUGLAS, DAN. 1995. **Developments in Language Testing**. Cambridge University Press.
- FULCHER, GLENN. 1996. **Testing tasks: issues in task design and the group oral**. University of Surrey.
- LITTLEWOOD, WILLIAM. 1992. **Teaching oral communication: A methodological framework**. Blackwell Publishers, Oxford.
- MCDONOUGH, JO AND CHRISTOPER SHAW. 1993. **Materials and methods in ELT: A teacher's guide**. Blackwell, Oxford.

- McNAMARA, T.F. 1996. **Measuring Second Language Performance**. Longman, London and New York
- NORRIS, J.M., BROWN, J.D., HUDSON, T & YOSHIOKA, J. 1998. **Designing Second Language Performance Assessments**. Second Language Teaching & Curriculum Center, University of Hawaii at Manoa
- ROBINSON, P. & ROSS, S. 1996. **The Development of Task-Based Assessment in English for Academic Purposes Programs**. Oxford University Press.
- SHOHAMY, ELANA. 1995. **Performance assessment in language testing**. Cambridge University Press
- SKEHAN, PETER. 1996. **A Framework for the Implementation of Task-based Instruction**. Oxford University Press. (Applied Linguistics, Vol. 17, No. 1)
- SKEHAN, PETER. 1998. **A Cognitive Approach to Language Learning**. Oxford University Press
- SKEHAN, PETER; BYGATE, MARTIN; SWAIN, MERRILL 2001. **Researching Pedagogic Tasks**. Pearson Education Limited
- UNDERHILL, NIC. 1987. Testing **spoken language: A handbook of oral testing techniques**. Cambridge University Press
- WEIR, C.J. 1993. **Understanding and Developing Language Tests**. Prentice Hall Europe