DIPLOMARBEIT

Titel der Diplomarbeit

"Objective test formats for Language Testing – Importance, Implementation & Improvement"

Verfasserin

Christina Carina Moser

angestrebter akademischer Grad

Magistra der Philosophie (Mag.phil.)

Wien, 2012

Acknowledgements

Writing a thesis is a challenging project in a student's life, which cannot be successfully completed without the help, support and encouragement of a number of important people. Therefore, I would like to offer a few words of thanks to those who have been especially helpful to me in writing this thesis.

First of all, I would like to thank Prof. Christiane Dalton- Puffer, who encouraged me to deal with the fascinating field of language testing. I want to thank her for her help during the writing of this paper. Her comments and suggestions guided me in the process of writing and I am grateful for her support.

Of course, I must not forget my dear friends and family, who were always there for me and, who successfully managed to talk me into having fun and relaxing breaks. Especially I would like to thank my parents, grandparents and my sister Selina, who always believed in me. Above all I am particularly indebted to my friend Tina not only for her friendship but also for engaging in conversations about language testing with me until late. She managed to continuously encourage me and was proud even of the smallest steps I took in the completion of this thesis.

Declaration of Authenticity

I confirm to have conceived and written this paper in English all by myself. Quotations from other authors and any ideas borrowed and/or passages paraphrased from the works of other authors are all clearly marked within the text and acknowledged in the bibliographical references.

Vienna, September 2012 _____

Zusammenfassung

Die Forschungsarbeit behandelt die verschiedenen Arten von objektiven Testformaten und deren Anwendungsformen innerhalb des Schulunterrichts und in formalen Testsituationen. Nach einer kurzen Einleitung werden die Anfänge der Sprachtests beschrieben und anschließend die wichtigsten Begriffe im Zusammenhang mit Sprachtests im Allgemeinen definiert. Diese Definitionen umschließen Begriffe wie der „Washback Effekt", Reliabilität, Validität und Authentizität. Des Weiteren werden die unterschiedlichen Formen von Tests und deren Zwecke detailliert erklärt. Im Anschluss daran werden die Grundlagen zur späteren Analyse von ausgewählten Testbeispielen gelegt. Dazu gehört unter anderem die Definition und Charakterisierung der vier Hauptkompetenzen der Sprache, Lesen, Schreiben, Hören und Sprechen, mit besonderem Augenmerk auf die Lesekompetenz, da diese besonders bei den verschiedenen Arten der analysierten Testbeispielen eine wichtige Rolle spielt.

Die Erkenntnisse aus dieser Grundlagenforschung dienen als Basis für die anschließende Forschung im Bereich der Sprachtests sowie deren Anwendung und Auswirkungen. Dabei wurde festgestellt, dass verschiedene objektive Testformate, wie Mehrfachwahlaufgaben, Richtig/falsch/ nicht im Text -Aufgaben, und Zuordnungsaufgaben, in Testsituationen viele Vorteile hinsichtlich des Einsatzes im Schulunterricht mitbringen. Dennoch haben sich auch einige Beschränkungen gezeigt. Durch eine detaillierte Analyse von Testbeispielen aus einem österreichischen Schulbuch und offiziellen Aufgaben der neuen Matura wurde aufgezeigt, dass sich Variablen wie Alter, Geschlecht, Wortschatz und Textgenre auf Testresultate auswirken können.

Abstract

This research paper is concerned with the various types of objective testing formats and their fields of application within the language classroom as well as in formal testing situations. After a short introduction and a description of the roots of language testing as such follows a definition of the most important terms in connection to language testing in general. These definitions include terms like washback effect, reliability, validity and authenticity. Furthermore the various purposes of language testing have been explained. Afterwards the basis for the following analysis of test items will be developed. This includes the definition and characterisation of the four skills in language, reading, listening, writing and speaking, with a focus on the reading competence, because this is especially important in connection with reading tests and the examples that have been analysed.

The knowledge gained from the first part of the thesis serves as the basis for the following research in the field of language testing and its' application and impact. It has been observed that multiple choice items, including classical multiple choice items, true/false/not given and matching, bring a variety of advantages to the language classroom. However, also some limitations have been identified. Moreover, through a detailed analysis of sample items from textbooks and the New Matura, it has been shown that numerous variables like age, sex, vocabulary and text genre can have an influence on the results of a test.

Table of contents

1. Introduction

Reading is a very important skill not only in first but also in foreign language situations. Furthermore, reading is a very effective exposure to the second language and a crucial part of every foreign language classroom and usually students are taught various reading techniques and strategies. This thesis intends to demonstrate the importance of various objective testing formats and how they are used not only in the language classroom but also in the New Matura. I chose this focus mainly because, as a future second language teacher, I am particularly interested in language testing, which forms a fascinating field within applied linguistics. Because reading is a receptive skill it does not require any productive skills from the test takers. Accordingly, it is advisable to assess receptive skills by methods, which do not require production on part of the students, as this could influence the learner's reading ability negatively, which may lead to unreliable assessment (see Alderson 2000: 30). This view is supported by Brown and Hudson (2002) who claim that selected-response formats are "most appropriate" (59) for testing the receptive skill of reading. Although there is "no one best method" (Alderson 2000: 203) some methods seem to be more adequate than others. Given these reasons it is important for teachers and people involved in the testing of language to assess students' reading comprehension abilities.

At this stage it is essential to make a clear distinction between the terms *testing* and *assessment*. The term *assessment* is considered a rather broad conception because it is part of the entire progression of learning and teaching. "It refers to the general process of monitoring or keeping track of the learners' progress" (Hedge 2000: 376). The process of assessment involves methods like "gathering of language data, including test data, for the purpose of evaluation and making use of such instruments as interview, case study, questionnaire, observation techniques" (Davies et al. 1999: 11). On the contrary the term *testing* is more limited by definition as being only one type of assessment and is "typically used at the end of a stage of instruction to measure student achievement" (Hedge: 376). Bearing these definitions in mind it is obvious that this broad concept of assessment has various purposes. Firstly, assessment can have a formative purpose. According to Tricia Hedge, formative assessment is always motivated by a pedagogical aspect. The classroom assessment with a formative approach provides information for the

teacher or evaluator that helps to improve the teachers' classroom work (see Hedge: 376). Henning (1987) provides a definition and describes it as "an evaluation, which is ongoing and iterative during an instructional sequence" (192). Furthermore Hedge provides several unique characteristics of formative assessment. It is prepared and carried out by the teacher as a part of the teaching process and relates to the content of the lessons that have been taught. The information gained from formative assessment can be used to identify the learner's needs as well as their weaknesses and strengths and how the teacher could help the learner to further improve (see Hedge: 376).

The second purpose for assessment is called summative assessment and refers to a method of evaluating the learner's success and accomplishment. In comparison with formative assessment, summative assessment is not essentially carried out by the teacher but can also be imposed externally. Moreover summative assessment is not focusing on the process of learning but on the result of learning and does therefore not have to relate to what has been taught by the teacher (see Hedge: 377). Summative assessment is often used by an institution to measure the abilities of a learner in a foreign language, which leads to the third purpose of assessment, formal certification. Furthermore Henning (1987: 197) describes it as an "[e]valuation that comes at the conclusion of an educational program or instructional sequence".

Finally, formal certification forms the third purpose of assessment. It is typically executed by an institution other than the regular school environment of a learner. In a formal certification a variety of the conventional paper-pencil test is used. There is a great number of language testing organisations, for example the Cambridge Proficiency in English (CPE) or the Test of English as a Foreign Language (TOEFL).

Besides those three purposes of assessment there is another method of assessing a student. It is generally referred to as informal assessment and includes various methods. One type of informal assessment is referred to as self-assessment or self-evaluation and "involves learners in making judgements about their own level and/or progress" (Davies et al. 1999: 177). Techniques that are commonly used in self-assessment include standardised tests that are marked by the learners themselves, questionnaires and different reflective and descriptive tasks like reading

aloud, the usage of checklists, keeping diaries or portfolios and other self-report techniques. Informal assessment methods also involve the teacher in the way that he or she has to observe the learners individually and in groups while reading and also pay attention to their non-verbal behaviour i.e. gestures and facial expressions (see Alderson 2000: 259). One way of carrying out observation-driven assessment of learners is to go "around the class while students are doing a reading activity and observing which students seem to be understanding it and which are having difficulties." (Harris and McCann 1994: 17) However techniques of informal assessment are regarded as being more time-consuming and complex than formal assessment methods and are therefore not used regularly by teachers. But they offer the teacher the possibility to develop insight into the progress learners make (see Alderson 2000: 265) and enables the teacher to diagnose strengths and weaknesses of learners in terms of their reading abilities.

After having defined the difference between assessment and testing in general, I want to describe the main focus of this thesis in greater detail. This thesis is divided into four main sections. The first part intends to give an overview of the roots of language testing and explains the various purposes for language testing in general. Furthermore definitions of the most important terms in connection with language testing are included.

The second part describes the four skills, listening, writing, speaking and reading in detail and explains the difference between skills and strategies. Moreover, a focus is put on the skill of reading because it is commonly tested with the help of objective test formats for various reasons, which will be elaborated later in this thesis.

The third part of this thesis engages in the nature of objective test formats and describes the characteristics of items and distractors. In the centre of this section however are the various selected-response formats. The three selected-response formats under investigation are multiple choice, true/false/not given and matching. According to Popham (2002) multiple choice, true/false/not given and matching items are "the most common kinds of selected-response test items" (145). These formats will be characterised by means of their definition, composition as well as their advantages and disadvantages. Other interesting aspects that I want to clarify include various factors that can influence a test taker's performance in a reading test as well as the question if female and male test takers can be influenced differently

by the format of a test. On the basis of part three, the last part of the thesis is concerned with the discussion of my research questions.

The main research question guiding this paper is about the implementation of objective testing formats in the language classroom. More specifically this part of my thesis is concerned with examples of multiple choice, true/false/not given and matching tasks from current textbooks used in the Austrian classroom with examples from the New Matura, that I have been provided with by the Bundesinstitut bifie. I want to analyse some examples used in the New Matura and find out if these examples are comparable to instances of selected response formats that are used in schoolbooks. Furthermore I am interested in ways of preparing students for this testing format. Finally I will try to come up with suggestions for improving the analysed samples and describe their advantages and disadvantages. The following list provides an overview of the research questions that will be addressed:

1. What different ways of preparing learners for objective test formats are there?

2. Does the New Matura use examples that are comparable to the textbook examples?

3. What could be improved in objective testing formats?

2. Overview & Definitions

This chapter serves as a starting point for the following analysis of various objective formats of testing reading. In order to be able to identify characteristics of each testing format it is essential to be familiar with the history and origin of language testing as such. Furthermore I want to identify different purposes of language tests and give definitions of some important terms connected to language testing in general.

2.1. Roots of Language Testing

In the dictionary of language testing Davies et al. (1999) defined language tests as an instrument to measure language ability or aptitude that consist of specified tasks, which elicit language abilities (see 107). "What language testing does is to compel attention to the meaning of ideas in linguistics and applied linguistics" (Davies 1990: 2).

The methods of testing a learner's language proficiency before the Second World War cannot be compared to the ones that modern teaching and testing is dealing with. "If a learner had to be assessed for second language proficiency this was done using the same means as had been used to teach it: composition, translation, dictation, etc" (Baker 1989: 29). The roots of modern language testing lie in developments in applied linguistics as well as educational measurement (see Hedge 2000: 378). Concerning applied linguistics, language testing was mainly influenced by a model developed by Robert Lado and J. B. Carroll in 1961. Their structuralist approach assumed that "'knowing' a language was identified with knowing the 'form' or 'structure' as well as the other linguistic elements of 'lexis' and 'phonology'" (ibid).

2.2. Purposes of Language Tests

The main question that is guiding this section is why language tests are conducted in the first place. Alan Davies (1990: 20) lists four main purposes of language testing, namely selection, feedback, evaluation and research. Grant Henning (1987) also identifies numerous purposes of language tests. For Henning these are

diagnosis and feedback, screening and selection, placement, program evaluation, providing research criteria, assessment of attitudes and socio-psychological differences. Those purposes will be discussed briefly in the following chapter in order to distinguish the various characteristics of each purpose. The purposes provided by Alan Davies (1990) will be paired with the purposes suggested by Grant Henning (1987) with the intention of arriving at a general picture of the various purposes of language testing.

2.2.1. Selection and Screening

When a test is used for the purpose of selection or screening there are several factors that need to be considered. The main purpose of this kind of test is to decide on "who should be allowed to participate in a particular program of instruction" (Henning 1987: 2). It is necessary to determine who will benefit the most from the language program. According to David Baker (1989) selection tests can be classified as entrance, progress and exit tests. The biggest difference between those three types of selection tests is at which point in time the learner is confronted with them.

The entrance selection test is meant to fulfil the purpose of forming a homogenous group of learners in order to guarantee an effective learning environment (see Baker 1989: 101). The main purpose of a progress test is to judge the previously used teaching methods and materials. At this point of a teaching program it would be possible to adjust the "course content, learner placement and future course design" (Baker 1989: 103).

The exit test is comparable to the progress test. It can measure what knowledge a learner has gained. The results of the exit test may decide whether a certificate is awarded or not, or in some cases, if the learner may proceed to another course with a higher level (see Baker 1989: 104). Critics of selection tests question in how far such tests can be objective, i.e. do not contain any culture specific questions. This is why the factors validity and reliability are of such great importance.

2.2.2. Feedback, Diagnosis and Placement

Probably the most prominent purpose of a language test in the context of a language classroom is providing learners or students with feedback on their abilities in a language. It gives learners as well as teachers the possibility to determine strengths and weaknesses in the learned language (Henning 1987: 1). As the term diagnosis implies, this purpose of a language test can be most commonly found in diagnostic test. This term will be explained in 2.3. Language Uses. Placement is connected to diagnosis as well as to selection and refers to a case when "tests are used to identify a particular performance level of the student and to place him or her at an appropriate level of instruction" (Henning 1987: 2).

2.2.3. Evaluation and Program Evaluation

The purpose of evaluation is basically concerned with the evaluation of the materials and the methods used in language teaching (see Davies 1990: 20). This is a very important instrument of language testing, which is not used continuously. This can be ascribed to the nature of the tests. Because they are all individually constructed it is hard to find a starting point for the evaluation. In terms of the evaluation of a program it is their effectiveness that is in question. Therefore the focus lies on group scores at the beginning and the end of a program. By comparing both scores a specific testing instrument or method may be qualified as effective or can be adapted for future uses.

2.2.4. Research

Language tests do not only present knowledge about an individual learner but they can "provide a standard of judgement in a variety of other research contexts" (Henning 1987: 3). When a test score is analysed it may lead to information about the test as such and lead the way to an improved version of the test. This comparison can affect the methods of instruction as well as textbooks and other aids in teaching a language (ibid).

2.3. Test Types

The previously described purposes of language tests can be used in various types of tests. Defining types of a language test we can differentiate four types. As Alan Davies (1990: 20) suggests these are: achievement, diagnostic, proficiency and aptitude tests. These four test types will be defined briefly so as to notice the individual features of each use.

Achievement tests are only concerned with the previous learning of a student. Therefore they are generally used at the end of a teaching or learning period and serve as the basis for decision-making processes. It is "an instrument designed to measure what a person has learned within or up to a given time" (Davies et al. 1999: 2). Kitao et al. (1999) mention a specific type of achievement test that is called progress test. A progress test is given periodically to measure if a learner is learning the content that is covered in the course (see 7).

The second test type is called diagnostic test. Diagnostic tests are used "to identify test takers' strengths and weaknesses, by testing what they know or do not know in a language, or what skills they have or do not have" (Davies et al. 1999: 43). Information that is gained with a diagnostic test is used for the purpose of placement of learners into appropriate courses, to select and to plan a course. It "is of value in that it provides critical information to the student, teacher, and administrator that should make the learning process more efficient" (Henning 1987: 2) But because diagnostic tests are very "difficult and time-consuming to construct" (Davies et al. 1999: 43) they are not used in the classroom situation and are replaced with achievement or proficiency tests, even if Davies et al. (ibid) say that they only provide a more general picture of the abilities of a learner.

Proficiency tests are, similar to achievement test, interested in what has been learned but in "a much more vague way" because instead of exhibiting control over previous learning it "establishes generalizations on the basis of a typical syllabus" (Davies 1990: 20). Proficiency tests are supposed to "measure how much of a language has been learned" (Davies et al. 1999: 154). Proficiency tests can be classified as general and specific purpose tests. The former deal with language in general and the latter cover language that is needed for specific purposes (see Kitao et al. 1999: 7). The most commonly known use of proficiency tests is the

standardised version that can be found in official language tests like the TOEFL test or the IELTS test.

The fourth type is called aptitude test and its purpose is to "measure the extent to which an individual possesses specific language learning ability" (Davies et al. 1999: 10). Aptitude tests are most commonly used "for the selection of students for foreign language training" because they can help to "decide how long it will take a student to achieve sufficient mastery of a foreign language to study in the country where the language is spoken" (Lado 1972: 370). Skills that can be tested with aptitudes test include memory skills, intelligence or deduction of rules (see Kitao et al. 1999: 8). The area where aptitude tests are generally used is selection, diagnosis and the prediction of language learning success.

2.4. Washback Effect

The effect that learning, instructing and testing processes have on individuals, educational systems and society is referred to as backwash or washback effect and can either be positive or negative. Hughes (1989: 1) defines washback as "the effect of testing on teaching and learning". A language test can affect test takers in three different phases during the procedure of testing. Firstly, the experience of preparing for a test and actually taking it, secondly the feedback phase and thirdly the situation where decisions are based on test results (see Bachman 2004: 31).

Generally, positive washback has a beneficial effect on the learning because the "skills that students will be developing in order to take the test will also be skills that are useful to them in writing" and this means that "if tests are designed wisely, they can motivate students to do the kinds of things that they ought to be doing in order to learn communicate effectively in a language" (Kitao et al. 1999: 3).

Negative washback on the other hand happens "when test items are based on an outdated view of language" (Davies et al. 1999: 225) and do not relate closely to the curriculum. In the case of multiple choice testing this would mean that a language learner who is confronted with multiple choice tasks only is under considerable strain of learning such items instead of practising the skill that is to be tested in other ways. This may lead to a lack in language proficiency and can have a harmful effect on learning.

2.5. Validity

In language testing two fundamental concepts, which are strongly related to each other are validity and reliability. They play an important role because the results of language tests can have an impact on people's lives.

Concerning validity, which "attempts to provide a theoretical framework which gives reassurance to the test" (Davies 1990: 6), the central question a test constructor has to ask him or herself is if the constructed test is actually testing what the constructor thinks it is testing. "Validity is the extent to which a test measures the ability or knowledge that is purported to measure" (Henning 1987: 198). Davies (1990), Morrow (1981) as well as Davies et al. (1999) name five types of validity: face validity, construct validity, content validity, predictive validity and concurrent validity. Kitao et al. (1999: 12-13) make a distinction between internal and external validity, and classify face and content validity as being part of internal validity and names concurrent and predictive validity as being part of external validity. Kitao et al. (ibid) do not mention construct validity.

A definition of internal validity comes from Kitao et al., who say that it "is validity in terms of the test itself – whether the content of the test and the way the test is carried out allows it to test what it is intended to test." (1999: 12) Face validity, as a type of internal validity, is defined by Davies et al. (1999) as "the degree to which a test appears to measure the knowledge or abilities it claims to measure, as judged by an untrained observer (such as for instance the candidate taking the test [...])" (59). Furthermore Kitao et al. (1999: 12) explain that "it [face validity] is the impression a test makes on general readers as to what skills it requires". Even if face validity is considered as being important, "if there is conflict between it and one of the other validities, then face validity must be the first one to go" (Davies et al. 1999: 23).

In any language test construct validity is "an indication of how representative it is of an underlying theory of language learning […] and involves an investigation of the qualities that a test measures […]" (Davies et al. 1999: 33). This means determining if the test measures samples of the language skills representatively (see Hughes 2003: 26). Construct validity can be tested statistically by carrying out a factor analysis or multi-trait multi-method analysis. For a detailed explanation of those two statistical methods see Davies et al. (1999).

Content validity is also referred to as "conceptual or non-statistical validity based on a systematic analysis of the test content to determine whether it includes an adequate sample of the target domain to be measured" (Davies et al. 1999: 34). In other words it clarifies whether a test measures the content it is intended to measure involving the examination of the specifications and the syllabus the test may be based on. A test can be considered having content validity if the content of the test constitutes a representative example of the language skill or structure that it is meant to be concerned with (see Hughes 2003: 26). Furthermore Hughes stresses the importance of the execution of content validation during the development of a test to assure the reflection of the test specifications in the test content (see Hughes 2003: 27).

Validity that "has to do with the relationship between the test and other measures" (Kitao et al. 1999: 12) is called external validity. Validity types that fall into this category are predictive and concurrent validity.

Predictive validity provides information about the ability of a test to predict future outcomes in terms of performance, i.e. "the test accurately predicts performance in some subsequent situation" (Alderson et al. 1981: 13). But there is a clear limit "to how much of subsequent performance can be predicted by even the most valid test" (Kitao et al. 1999: 13).

Another type of external validity is concurrent validity, which by definition is

> concerned with the relationship between what is measured by a test […] and another existing criterion measure, which may be a well-established standardised test, a set of judgements or some other quantifiable variable. If the two measures function similarly they are considered to have concurrent validity (Davies et al. 1999: 30).

This means that the test under examination "gives similar results to existing tests which have already been validated" (Alderson et al. 1981: 13). The main difficulty with concurrent validity is to decide how much emphasis should be put on the previously existing criterion measure (see Davies et al. 1999: 30). Furthermore it is important that "the measure must be valid for the same purpose as the test whose validity is being considered" (Kitao et al. 1999: 13).

2.6.   Reliability

Reliability, which can be regarded as being one component of validity (see Davies 1990: 61), forms a second concept that is vital for language testing. It is concerned with "the consistency of test judgements and results" (Davies 1990: 21) and is based on objectivity. Robert Lado (1972) defines reliability by asking the question if "the test yield[s] dependable scores in the sense that they will not fluctuate very much so that we may know that the score obtained by a student is pretty close to the score he would obtain if we gave the test again" (31). 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" (Davies 1990: 5). In contrast to subjective tests, in objective tests it is not necessary that the test takers produce language themselves because they are rather asked to select from a variety of possible answers or alternatives. Accordingly, the production of language itself plays a minor role in objective tests. Furthermore "an objective test can reveal only differences and similarities between the language norms of the examiner and the candidate; it can tell us nothing of the norms which the candidate himself would apply in a use situation" (Alderson et al. 1981: 13). Concluding Davies (1990) states that

> we cannot allow reliability to dominate our thinking and we are compelled to agree that validity deserves its central place, but at the same time we must not go the opposite way and claim reliability – because it says nothing about truth – is unimportant and we can therefore afford to let test reliability slip (23).

2.7.   Authenticity

An interesting point that is raised during the discussions in Lyle Bachman's "Fundamental considerations in language testing" (1990) and other publications is the aspect of authenticity. The term authentic has been used to characterise "the extent to which the tasks required on a given test are similar to 'normal', or 'real-life' language use" (Bachman 1990: 9). Another definition comes from McNamara (2000) who describes authenticity as "[t]he degree to which test materials and test conditions succeed in replicating those in the target use situation" (131). Basically what authenticity refers to is the way in which test tasks are able to correspond to language tasks outside of the test itself. Therefore authenticity is an important characteristic of a language test "because it relates the test task to the domain of

generalization to which we want our score interpretations to generalize" (Bachman et al. 2004: 23-24). When considering authenticity in test construction it is essential to be aware of the fact that a variety of perspectives have to be taken into account. Those perspectives include the test taker, the test task and the target language use. Therefore I can conclude that authenticity is to be taken relative, there is no task that can be described as being completely authentic or completely inauthentic. In addition authenticity can be defined as an essential part of language testing because firstly, it links test performance with target language use and the domain to which it is intended to generalize and secondly, the perception of authenticity by the test takers has the potential to improve their test performance (see Bachman et al. 2004: 39).

3. Testing Reading

Language ability traditionally consists of four skills: listening, writing, speaking, and reading (see Bachmann 1996: 75). These four skills can further be distinguished between productive skills, i.e. speaking and writing, and receptive skills, i.e. reading and listening. The knowledge about these four skills goes back to models of language ability, established by Carroll and Lado in 1961 (Bachman 1996: 75) with their stucturalist approach. Lado claims that the skills are "elements of language" and they "can be profitably studied and described-and tested-as separate universes, yet they never occur separately in language. They are integrated in the total skill of speaking, listening, reading and writing" (Lado 1972: 25).

Even though there is basic agreement on the division of the four skills amongst applied linguists there is some opposition. Kitao et al. (1999) for instance state that "in real life the different skills are not often used entirely in isolation" (88) and brings in the argument of authenticity, which has already been discussed. Also Bachmann (1996) argues against a strict division of the four skills and characterises the four skills as "widely divergent language use tasks and abilities together under a single 'skill'" (75). In the following section the four skills will be defined separately, but a focus will be put on the reading skill because it plays an important role in objective testing. Moreover this section illuminates the difference between what is referred to as skill and strategy.

3.1. Listening

Some methods commonly used in English language teaching "have assumed that listening ability will develop automatically through exposure to the language and trough practice of grammar, vocabulary and pronunciation" (Hedge 2000: 227-228). But because listening is such an important skill for learners of the English language in order to be able to communicate successfully, it seems inevitable to consider listening as a new focus in the language classroom. When we talk about the skill of listening we have to distinguish between participatory and non-participatory listening. The former refers to face-to-face interaction as in everyday conversation where the listener is supposed to react in a certain way whereas the latter describes a rather passive listening as in listening to radio programmes or watching television.

Another important differentiation has to be made in terms of the listening process between bottom-up and top-down processes. In the bottom-up process a listener uses "knowledge of the language and [the] ability to process acoustic signals to make sense of the sounds that speech presents to [the listener]" (Hedge 2000: 230). Opposed to the bottom-up processing stand the top-down processing, where the knowledge that a listener already has is more important than the knowledge that can be gained from a speech act itself. "Top-down listening, […], infers from *contextual clues* and from making links between the spoken message and various types of *prior knowledge* which listeners hold inside their heads" (Hedge 2000: 232).

3.2. Writing

 "To write is to put down the graphic symbols that represent a language one understands, so that others can read these graphic symbols if they know the language and the graphic representation" (Lado 1964: 143). But there is more that forms the skill of writing. There are several strategies that writers use when they compose a written text. It is important to be aware of those strategies in order to help learners when they have to generate their own ideas and formulate them into a written text.

> Writing is the result of employing strategies to manage the composing process, which is one of gradually developing a text. It involves a number of activities: setting goals, generating ideas, organizing information, selecting appropriate language, making a draft, reading and reviewing it, then revising and editing (Hedge 2000: 302).

A well accepted study was published by Zamel (1983), that has been supported by many studies that followed it. Summing up this study, the writing process is characterised by three basic activities: planning, revising and producing reader-based texts. The time that is spent on the actual planning of a text depends on the individual writer. The planning phase includes thinking about the purpose of the text and the adaption of the style, i.e. formal or informal register. Even though the planning phase stands at the beginning of each writing process it can be reviewed "at any point in time as the writer critically evaluates the emerging text and thinks of new ideas and new ways to organize and express them" (Hedge 2000: 306). When a writer re-reads the produced text he or she can reflect on the sentences and

may revise, rephrase, add or delete sentences or paragraphs in order to improve the text (see ibid.). Therefore a strong interaction between the skill of writing and the skill of reading can be detected.

3.3. Speaking

When learners are given the opportunity to practice speaking in the classroom it is vital that they "do not only practice speaking in a controlled way in order to produce features of pronunciation, vocabulary, and structure accurately, but also practise using these features more freely […]" (Hedge 2000: 261). This means that the training of the oral abilities should foster spontaneous communication.

When communicating a speaker makes use of one or more communication strategies. On the one hand there is the so-called avoidance behaviour where the non-native speaker tries "to eliminate a problem by changing the topic or not participating in a conversation" and on the other hand there is the achievement behaviour, "in which ways are found to cope with the problem […]" (Hedge 2000: 265-266). These include word coinage, speaking in the first language, gesture or restructuring of a sentence. In order to prevent communicative act to collapse it is central to enable a learner to negotiate, for example with the help of achievement behaviour. Learners should be able to paraphrase their sentences or find other ways of expressing themselves and they have to ensure that they are understood, for example by repeating or clarifying their statements. A great variety of methods has been developed in order to practice these strategies in speaking tasks. Amongst other methods these can take the form of a free discussion, role-plays or gap activities.

3.4. Reading

According to Robert Lado reading „is to grasp language patterns from their written representation. In a second language, reading is usually taught to students who are already literate in the source language" (Lado 1964: 132). It may appear that reading is one of the skills that is tested most commonly because it is "the easiest of the four skills to test" (Kitao et al. 1999: 42) but the following description of the skill of reading wants to show that it certainly is a very complex skill. "A reading skill can

be described roughly as a cognitive ability which a person is able to use when interacting with written texts" (Urquhart and Weir 1998: 88). Accordingly skills in reading can be seen as being a part of the whole process of reading and can therefore "be defined as the receptive handling of a language in written form […]" (Urquhart & Weir 1998: 93).

3.4.1. Various Kinds of Reading

It is undeniable that a student has to develop different strategies in order to fulfil various tasks. The focus of attention varies from text to text and therefore a student has to know what he or she is supposed to do with a certain text in order to know what reading strategy is most helpful to complete a test item. "Learners must be able to switch between the two basic modes of reading, one that emphasizes 'the big picture' or 'a specific surface detail' and another one that requires slower and closer reading" (Gassner et al. 2007: 11).

The following definitions illustrate the various reading strategies defined by a number of authors like Urquhart and Weir (1998), Munby (1987) and Gassner et al. (2007). In general we distinguish between careful reading and expeditious reading. Each of those two types has several kinds of reading that characterise them.

3.4.1.1. Careful Reading

A description of the defining features of careful reading comes from Urquhart and Weir (1998: 103). They characterise it by three important facts. Firstly they describe a a non-selective process, where the reader tries to deal with the majority of textual information. Secondly Urquhart and Weir describe the reader as being submissive regarding the acceptance of the writer's organisation of a text. The third feature of careful reading refers to the attempt by the reader to construct a microstructure based on the first process of grasping the majority of information. This definition includes both abilities that are subsumed under the term careful reading, namely understanding the text as well as understanding the lexis.

Understanding the text

Gassner et al. (2007) summarise five different strategies under the heading of understanding a text. These are "Understanding explicitly stated main idea(s)", "Locating, identifying, understanding and comparing facts, opinions, definitions", "Understanding logical organisation of the text", "Understanding cohesive relationships" and "Inferencing" (11-13). Understanding explicitly stated main idea(s) is a strategy that is commonly used with longer texts where the focus of reading is put on the main idea(s) of a text "as opposed to supporting detail, and not on gist" (Gassner et al. 2007: 11-12). The second strategy mentioned is a quite complex strategy because it includes several different tasks for a reader. There is a strong need for search reading strategies. The third and fourth strategies are closely related as they both refer to the organisation of a text. They include operations like the understanding of relationships among ideas and different cohesive devices, i.e. reference, ellipsis, substitution or conjunction. It is essential that readers learn to deal with cohesive texts as well as with un-cohesive texts and identify the characterising features in order to be able to write logical, cohesive texts themselves. The fifth strategy describes the making of propositional inferences, which means that readers need to be able to deduce information that is not stated explicitly part of the information given in a text.

Understanding the lexis

When Gassner et al. (2007) speak of understanding the lexis they refer to the ability of a student of "[p]redicting the meaning of (unknown) words from context" (ibid. 13). This is a very authentic reading strategy because students may very well be confronted with unknown words in real life situations. Then they will have to figure out the meaning of the word by themselves by using the given context of the word. Even if Gassner et al. (2007) recognise the importance of this reading strategy they still criticise the practical implementation in the language classroom.

> Too often do learners first identify all unknown vocabulary and then get filled in by the teacher before work on a text even starts. It is important to take all opportunities or even create new ones to practise this strategy regularly so that pupils learn to deal with unknown words as a natural part of a more complex text and to pick up content clues that help to make intelligent guesses (ibid. 13).

3.4.1.2. Expeditious Reading

Expeditious reading refers to the kind of reading where a person extracts important information as well as the intended purpose (see Urquhart and Weir 1998: 101). The four kinds of expeditious reading are called scanning, skimming, browsing and search reading and they all "distinguish from careful reading by the factor of selectivity" (Urquhart and Weir 1998: 104).

Scanning

Urquhart and Weir (1998) define scanning as "[r]eading selectively, to achieve very specific reading goals, e.g. finding the number in a dictionary, finding the capital of Bavaria" (103). The selectiveness in this kind of expeditious reading means that anything that does not help to accomplish the reading goal, i.e. does not contain the most important keywords like names or dates, is ignored. Therefore another characteristic of scanning is the speed of reading, even if it is rather difficult or even impossible to control the reading speed of a student. Because of the rather fast way of reading the students often do not grasp the whole topic of the text but they are able to spot very specific information that serves the task. Gassner et al. (2007) claim, that "[…] it is important for readers to be able to handle reading tasks focusing on certain elements or passages of text without fully understanding the whole text in all its details" (10).

Skimming

Skimming is also called reading for gist because the details of a text are dismissed in order to grasp the overall topic of a text. According to Urquhart and Weir (1998)

> [t]he defining characteristics are (a) the reading is selective, with sections of the text either omitted or given very little attention; (b) an attempt is made to build up a macrostructure (the gist) on the basis of as few details from the text as possible (103).

Furthermore skimming refers to the forming of an impression of a text by neglecting the details. Which parts of a text are considered as being important for the set task is determined by the task itself. In their published LTC Technical Report 2 on testing reading, Gassner et al. mention three ways of testing if a reader

is able to skim. The three strategies they describe are referred to as "Recognising text type", "Identifying text topic" and "identifying text purpose" (Gassner et al. 2007: 9-10). Strategy number one is the identification of the text type or genre. "As learners need to be able to distinguish between a variety of text types, prominent features of text types need to be focused on in teaching and recognising them must be practised" (Gassner et al. 2007: 9). The first strategy of recognising a text type is characterised exposure to a variety of text types in order to make the students familiar with specific details of each type of text or genre. Identifying the topic of a text forms the second strategy defined by Gassner et al. and refers mostly to multiple choice questions that need to be assigned the correct title. This can only be done successfully if the students are able to identify the main topic of a text or a passage of a text. The third and last strategy deals with the purpose of a text and according to Gassner et al. it is a strategy that is not practised frequently in classrooms. They suppose that the various purposes of a text are taught at the beginning of the teaching of a foreign language.

Browsing

Browsing describes "the sort of reading where goals are not well defined, parts of a text may be skipped fairly randomly, and there is little attempt to integrate the information into a macrostructure […]" (Urquhart and Weir 1998: 103-104).

Search Reading

As the name indicates, in search reading it is all about finding specific information by searching through a text. It is not about constructing a macrostructure for a text like it is with skimming. "The reader wants information to answer set questions or to provide data, for example, in completing assignments" (Urquhart and Weir 1998: 103).

This discussion of the different kinds of reading serves as the basis for the chapter that follows because in objective testing formats a test constructor can decide on what kind of reading strategy he or she wants to test. Concluding this section it is essential to say, that a reader can change his or her reading behaviour or strategy at

any time throughout the reading of a text, regardless of the length of the text. "[…] [T]hey may switch from careful reading to skimming to search reading to scanning and back to careful reading over a small number of pages" (Urquhart and Weir 1998: 104).

3.5.   Skills versus Strategies

When researching a skill the starting point is an investigation of a taxonomy of skills that can be checked empirically in terms of validity (see Urquhart & Weir 1998: 94). Opposed to skill research stands strategy research where "the researcher begins by having subjects […] read a text and, either retrospectively at the end of reading or at points during reading, report on what they are doing" (ibid).

A definition of strategies comes from Pritchard (1990) who claims that a strategy is "a deliberate action that readers take voluntarily to develop understanding of what they read" (275). Two major studies on strategies were published by Olshavsky (1977) and Sarig (1987). They identified various strategies that were in use by the subjects who read a text. The strategies they found include word, clause and story related strategies as well as skimming, scanning and skipping. These terms describe various moves applied on a text while reading it. On the word level this means for example to use the context to define a word and on the clause level the strategies applied include re-reading or personal identification (see Urquhart & Weir 1998: 94). Needless to say, these two studies may have presented very important findings in the area of strategy research but they are not representative. The categorisation as well as the identification of strategies remains a subjective task for researchers.

The main difference between skills and strategies according to Urquhart & Weir is that strategies are basically reader-oriented and skills are text-oriented. Furthermore, skills are applied unconsciously and rather automatically, whereas strategies always include a conscious act by the reader (see Urquhart & Weir 1998: 96-97). This distinction is supported by Cohen, who believes that "[…] the element of consciousness is what distinguishes strategies from those processes that are not strategic" (Cohen 1998: 5).

3.6. The Construct in the New Matura

The question about the importance of objective testing formats in the language classroom as well as in the New Matura includes the question if and in how far objective testing formats are used in practise.

The reason why selected response tasks and different variations of them are used in the language classroom as well as in the reading section of the New Matura is to determine whether the B2 level in reading has been reached by the test taker or not. The following table taken from the homepage of the university of Innsbruck summarises what purposes the reading test has and what different reading skills are tested in the New Matura.

Criteria	Description	Source
General purpose	To determine if students have reached B2 level in reading.	
Specific purpose	1. Can read with a large degree of independence, adapting style and speed of reading to different texts and purposes, and using appropriate reference sources selectively. Has a broad active reading vocabulary but may experience some difficulty with low frequency idioms (ORC)	CEFR
	2. Can read correspondence relating to his/her field of interest and readily grasp the essential meaning (RC)	CEFR
	3. Can scan quickly through long and complex texts, locating relevant details or a sequence of events. (RO) *(descriptor currently under revision)*	CEFR adapted
	4. Can quickly identify the content and the relevance of news items, articles and reports on a wide range of professional topics. (RO)	CEFR (one part deleted)
	5. Can understand articles and reports concerned with contemporary problems in which the writers adopt particular stances and viewpoints. (RIA)	CEFR
	6. Can understand lengthy, complex instructions in his field, including details on conditions and warnings, provided he/she can reread different sections. (RI)	CEFR
	7. I can understand contemporary literary prose.	CEFR p.27
	8. Can use a variety of strategies to achieve comprehension; checking comprehension by using contextual clues.	CEFR ("identifying clues and inferring", one part deleted)
	General Focus A Gist (G) B Reading for specific information (SI) C Reading for important details (ID) D Search reading (SR) E Reading for main ideas and supporting details (MISD) F Reading to infer (propositional) meaning (IPM) G Reading to deduce the meaning of a word/phrase (DMW)	
Target language situation	⊕ leaving qualification at the end of secondary school ⊕ qualification for future academic studies	
Description of the test taker	Age 17 upwards; majority 18-20; L1 majority German speakers.	
Test level	B2	
Text source	Entire or part of a text from any of the following sources: newspapers, internet [blogs, social	

Figure 1: Reading Test Specifications.
(www.uibk.ac.at/srp/PDFs/Reading_Test_Specifications_B2_v14.pdf, 14. June 2012)

As we can see in this table, the general focus that is tested in the New Matura is in accordance with various kinds of reading that have been defined in section 3.4.1. Because of the new standardised version of the reading section is part of the New Matura these reading tasks are included in a variety of new schoolbook publications. Therefore it is obvious that the role of multiple choice, true/false/not given and matching exercises is growing enormously. This has a great impact on both, the teachers and the students. Teachers will be required to use those items in

the language classroom on a regular basis in order to prepare the students for the written Matura. This also includes the usage of the various objective testing formats in tests and exams, in order to test how the students behave in a testing situation. Students will have to get used to this standardised testing procedure by practising independently as well as in a school context how to approach and solve these testing formats. Due to this growing importance of objective test formats a detailed analysis of fixed response format tasks will be part of this thesis.

3.7. What factors can influence a test taker's performance in a reading test?

There is a great variety of factors that can influence the performance of a test taker on a reading test. These factors can generally be grouped into two main types of variables: text variables and reader variables.

Alderson (2000: 33-60) lists numerous reader variables that affect the reading process. He generally divides the variables that relate to the reader into different categories: the knowledge of the reader, the motivation of the reader to read and the reader's interest in the text as well as other stable characteristics.

Aspects that fall into the category of reader's knowledge are background knowledge of the test taker, knowledge of the language and text type, knowledge of topic as well as cultural knowledge and metalinguistic knowledge such as reading strategies. Background knowledge refers to schemata or scripts that everyone has about specific events that are very common in our daily life. "Schemata are seen as interlocking mental structures presenting the readers' knowledge" (Alderson 2000: 33). These schemata can help a test taker to recall specific information and integrate the new information gained from a text. An aspect that is vital for the understanding of a text is of course the knowledge of language itself. Thus, the ability to read and process a written text depends strongly on the linguistic knowledge of a reader. Similarly, the knowledge about the organisation and structuring of a text depends on the level of proficiency of a reader and can have a strong influence on the performance. Different text types require different reading strategies, as it has been demonstrated in chapter 3.5. Therefore the metalinguistic knowledge of reading strategies is essential for the successful completion of a reading test.

But not only the knowledge that a reader has affects the reading process. Also motivation, interest in the topic and the reader's abilities and skills play an important role in the reading process and therefore form the second category of reader variables. By abilities Spearitt (qtd. in Alderson 2000: 49) refers to information processing abilities that include the recalling of the meaning of words, the drawing of inferences from the content, the recognition of the writer's intention or tone and the ability to follow the structure of a text or text passage. Another variable that can affect test performance is the purpose of reading. "[T]he reason you are reading a text will influence the way you read it, the skills you require or use, and the ultimate understanding and recall you have of that text" (Alderson 2000: 50). The purpose of reading a text can have an effect on the motivation of a reader to read a certain text. It is important to distinguish between extrinsic and intrinsic motivation. Although it is obviously very difficult to induce it, the educationally more desirable motivation is clearly intrinsic motivation because it comes from the reader. Extrinsic motivation affects the reading process by various factors.

> [E]xtrinsically motivated students seem to read at a surface level, paying less attention to facts and details rather than to the main ideas, to what the text is about, to how ideas in the text relate to each other, and to how the ext relates to other texts, or to what the reader knows about the subject or the world (Alderson 2000: 53).

Interrelated to motivation is the variable of interest. Various studies have shown that motivation, interest and other positive factors can have a strong influence on the reading process. Focusing on the aspect of interest, a distinction is presented by Ay and Bartan et al. (2011: 245-246) that differentiates between individual interest and situational interest. While individual interest describes the individual preferences of a reader for a specific topic that exists previous to the reading of a text, situational interest subsumes different feelings that are triggered by a certain situation. Situational interest, as the name implies, can vary according to the situation, text or test.

Furthermore reader variables include the category of "stable characteristics of the reader, like sex, age and personality" as well as "physical characteristics, like eye movements, speed of recognition, […] and such like" (Alderson 2000: 33).

Especially the aspect of sex or gender influencing a test taker's performance was mentioned in several publications (Arabski et al. 2011; Lynn et al. 2009; Doolittle et al. 1989; Norwell 1958; Thorndike 1941) and is therefore considered separately in research question 6.

Over and above those reader variables there are various variables based on the text that can have an effect on the outcomes of a reading comprehension test. (Alderson 2000: 60-79) There is common agreement that the content of a text has an influence on the way a text is processed by a reader. "It is generally assumed (and has long been established) that abstract texts will be harder to understand than texts describing real objects, events or activities" (Alderson 2000: 62). Topics taken from a student's everyday life are more readable and therefore easier to process. Also the difficulty of vocabulary used in the text affects the comprehension. But if a text is completely unfamiliar to a test taker easy vocabulary cannot compensate the unfamiliarity. Other aspects concerning the content of a text that affect the processing are according to Alderson (2000) quantity of information and density of propositions (see 62). In relation to text topic or content of a text stand the style and other differentiating features of a text, i.e. the text type or genre of a text. Basically a distinction between expository and narrative text types is made. The former is harder to understand than the latter based on the greater variety of topics and relationships among units of the text (see Alderson 2000: 64). Given the disparity of expository and narrative texts we also have to distinguish between literary and non-literary text types. One common belief is that "literary texts are somehow harder to process, either because of the multiple layers of meaning they are held to contain, or because of the wider and more complex range of language they exhibit" (Alderson 2000: 65). But the opinions on this issue are rather controversial. On the one hand linguists believe that the language of poems, as a prime example of literary texts, is not very different from non-literary texts and the processing of such texts works in similar ways (Van Peer 1986; Van Dijk 1977). On the other hand critics suggest that a special form of competence is needed in order to process literary texts (Culler 1975).

Besides the topic of a text the way it is organised is a differentiating feature. The organisation of a text includes aspects of the arrangement of the various paragraphs, their relationships to each other as well as to different ideas in the text.

It has been shown that

> texts organised according to the sequence of events could be read faster and were easier to understand than texts whose temporal sequencing was disturbed […]. [A]lso texts with a consistent spatial organisation, e.g. descriptions of objects that followed a clear logical sequence […] were easier to understand and recall (Alderson 2000: 67).

Moreover, typographical features may influence the perception of a text. This accounts for various perceptual features like font types and size as well as amount of information or number of words. Also tables, diagrams or other graphic representations may accompany a text and thereby ease the processing of information. Besides non-verbal information the medium of presentation plays an important role in the understanding of a text. Nowadays there is a great variety of way of how to present a text or test to select from. Next to the traditional print out version the commonly used media include TV-screens, computer based material and overhead slides. When a computer monitor or TV-screen is used it is important to be aware of certain points that have to be considered in such forms of presentation. For example the size and type of font that is used should be adapted to screen size. Other than that there should be a high contrast between the background colour and the colour of the font. If a different medium than paper is used there is "the potential fatigue effect due to screen glare" (Alderson 2000: 78) that should be taken into account. Concluding it can be said that "[r]eading from paper is generally faster, more accurate and less fatiguing" (Alderson 2000: 79).

3.8. Can the format of a reading test affect female and male test takers differently?

As I have mentioned in the discussion of the previous question, one aspect that can have an influence on the performance of a test taker on a reading comprehension test is gender or gender linked to different aspects. "*Gender* is what culture makes out of the 'raw material' of biological sex" (Crawford et al. 1995 qtd. in Halpern 2000: 232).

One aspect gender is commonly linked to is interest.

> Gender is a complex variable because men and women differ not only biologically, but also in their life experiences. More specifically, men and women's lives tend to differ in ways that at least on the face of it, appear to have relevance to the observed gender differences (Leibenluft 1996: 163 qtd. in Halpern 2000: 233).

These gender related interests reflected also in the test scores of various studies. In the PISA survey from 2006 it could be concluded that "gender differences in reading, favouring females, existed in every country" (Brookhart 2009: 120). Similarly, Norvell (1958) and Thorndike (1941) found out that

> males prefer adventure, sports, science and information, while females prefer mystery and romance; males are less likely to read about female protagonists and their activities whereas females are likely to read about male protagonists and their adventures (qtd. in Ay et al. 2011: 246).

Doolittle and Welch (1989) found out that females reached lower scores in science topics and higher scores if text were taken from the field of the humanities. The study conducted by Ay and Bartan (2011) gained similar insights. "It can be observed that fashion, fairy tale and humour topics guarantee the female students significantly higher marks than those received by males overall, regardless of difference in levels" (250). A great number of studies have examined the effect that the assessment format can have on female and male test takers. Diane Halpern (2000) cites a study by Allport, Vernon and Lindzey from 1970 and concludes that "[t]here are strong and consistent differences in the interests, values, and attitudes of females and males in contemporary Western society" (243). Furthermore Susan Brookhart (2009: 130-133) summarises some of the most important findings. She mentioned Willingham and Cole (1997) who considered studies that tested effects of multiple choice and constructed response formats. They found out that evidence from the studies under review did not support their theory that multiple choice format in testing reading is a source of gender differences in the results of a test.

Similar were the outcome of a study conducted by DeMars (1998) who only found very small differences between the two sexes. However, in 2005 Woodfield, Earl-Novell and Solomon were interested in a comparison between course work and

examination performance of British students. "Results indicated that female students did outperform male students, but by about the same amount on both course-work assignments and examination" (qtd. in Brookhart 2009: 131). Another area that can affect test performance is anxiety. In a study conducted by Hembree (1988) findings showed "that test anxiety and performance were significantly related […] [Furthermore] girls exhibited higher mean test anxiety than boys, but their higher test anxiety did not appear to translate into a difference in performance" (ibid).

Given these findings test developers and teachers have to be aware of the fact that the area a text is form can influence the performance on a reading test. "Items that are offensive or that conceptually seem to favour one group over another" (Brookhart 2009: 128) should be revised or eliminated.

4. Objective test formats

A very general distinction concerning the scoring of a test was introduced by Robert Lado (1972), who distinguishes between subjective and objective tests. Subjective tests, as the name suggests, "require an opinion, a judgement on the part of the examiner" (28) that, to a certain degree, is always subjective and therefore there can be a variation in the scoring done by different examiners. Objective tests on the contrary are scored without giving the examiner the possibility to influence the test score. Nonetheless Lado stresses, that "tests are not either subjective or objective but range over a scale with completely objective scoring on one end and completely subjective scoring at the other" (1972: 29). The focus of this chapter will be put on testing formats that are commonly regarded as being rather at the objective end of the scale Lado was talking about.

Objective tests are those that include test items that require the learners to either supply an answer or solution to a given problem or to select the correct answer from a given number of choices. Norman E. Gronlund (1965) provides a classification of these objective test types and divides them into two main categories. The first category is referred to as supply types and includes short answer test items as well as completion tasks. The second category is called selection types and includes multiple choice tasks, matching exercises and true/false/not given items. These types of tests or test items will be explained in detail in the sections 4.3, 4.4 and 4.5.

In addition to those two main categories Alderson et al. (1995) mention a number of other test items that could be considered objective test items. These include information transfer, ordering tasks, editing and gap-filling exercises. Heaton (1983) also lists these types but provides different terms for each item type. Information transfer is called "transformation", ordering tasks are called "rearrangement" and gap-filling exercises are referred to as "completion" (13).

Information transfer items typically ask the test takers to transfer information from a text on to some sort of table, chart or map. These items can be classified as objective item types only if the answers include names or numbers. If the answer takes a more complex form they cannot be marked objectively.

Ordering tasks, as the name indicated, consist of word groups, paragraphs or phrases that need to be put in order. Even if these items provide a good method for the testing of simple as well as complex grammar issues it is difficult to find components of an item that can be ordered in one way only. Furthermore the marking of ordering items can be rather tricky. The test constructor has to decide whether to mark individual ordering errors separately or if the complete item is to be scored incorrect if one or two sentences have been ordered incorrectly (see Alderson et al. 1995: 53).

Editing tasks generally "consist of sentences or passages in which errors have been introduced which the candidate has to identify" (Alderson et al. 1995: 54). Depending on whether the item is posed as a multiple choice question or a more open form the scoring of such items can range from easy to more time consuming. Test takers may be asked to only mark the errors or to write a correction of each incorrect sentence. Either way the test takers should always be informed about the number of errors they have to find.

In gap-filling activities test takers are asked to fill in missing words in a sentence or statements where some words have been deleted. It is essential for the scoring process that there is only one correct solution that can fill the blank. Because this is not always easy to ensure there is the possibility of designing banked gap-filling tasks. "Each of the missing words or phrases is included in a list […]" (ibid) and the candidate has to select the correct word from that list. That way an objective scoring can be guaranteed because the test constructor does not have to decide which words can be accepted and which words are counted as being incorrect.

4.1. Item Characteristics

In general an item can be defined as "[t]he part of a test from which scores are derived" (Davies et al. 1999: 92). A more elaborate definition is provided by Brown et al. (2002) who quote Osterlind (1989). "A unit of measurement with a prompt and a prescriptive form of responding, which is intended to yield a response from an examinee from which performance in some language construct may be inferred in order to make some decision" (57). The prompt is typically called stem, which form the basis for the alternatives in an item that is part of an objective test format. It may occur that the stem takes the forms of an incomplete statement, a complete

statement or a question. The choices a test taker has are divided into the key, a term that refers to the one correct answer, and several alternative answers that are known as distractors (Davies et al. 1999; Brown et al. 2002), distracters (Gronlund 1965) or foils (Oosterhof 2009). An item may be presented either in an isolated form or an additional input may be provided to contextualise the item. In reading tests the latter form of presentation is normally used because the isolated form of presentation of an item is not applicable in reading tests. Test items "require a specified form of response from the test taker. On the basis of this response or set of responses a decision is made about the test taker's knowledge or ability" (Davies et al. 1999: 201).

Every objective test format that will be discussed in the following section makes use of items that can take a variety of forms. According to Robert Lado (1972) "the effectiveness of a test is determined by its items and their arrangement" (342). Therefore it is important to define what an item is and how it can be characterised. Basically Lado distinguishes between three types of items: language items, recognition items and production items. Language items are examples of the performance of the learner on a specific language problem whereas recognition items present a stimulus that includes a problem and make the learner respond in a way that shows if the learner was able to solve the problem (ibid). The main characteristic of a production item is that it "provides a stimulus that forces the student to express the problem utterance or parts of it" (ibid).

There are several factors that should be avoided in a good item that is intended to be part of an objective test format. A very important aspect concerning an item is the balance between the correct answer and the alternatives that are presented. The correct answer should never be a direct repetition of the item stem. It is better to re-phrase the correct option in order to ensure that the test takers have understood the meaning of the stem. Furthermore the incorrect options have to be clearly incorrect, but they should still be logical in relation to the item stem.

Another vital consideration concerning an item is its difficulty, which is "determined as the proportion of correct responses" (Henning 1987: 49). Accordingly, the lower the proportion of correct answers to an item the more difficult it is. Factors that contribute most to the difficulty of an item include the length of the stem, the length of the correct and incorrect options as well as stem

44

unique content words, which are "those nouns, verbs and modifiers that appear in the question but not the passage" (Alderson 2000: 88). Tests that use a variety of multiple choice items often arrange the items in order of the difficulty, meaning that a test starts with the easiest item and the more complicated items come at the end of a test. This has two important reasons. First of all

> [w]hen items are arranged according to difficulty the student can work rapidly through the simple ones until he finds them increasingly difficult and then devotes the proper time to solve them […] thus the arrangement helps the student to show his real level of achievement (Lado 1961: 356).

If the items are arranged randomly it might occur that test takers would have been able to solve items that are positioned at the end of a test but they did not have the time because of the more difficult items at the beginning of the test. Secondly, Lado (1961) mentions mental sets, referring to the effort that a student has to invest in the solving of one item. If a difficult item has been mastered and a fairly easy item follows it will be attacked "with more power than is necessary and will waste time and effort" and the same effect can occur if an easy item has been solved and a difficult item follows "the student will tend to attack it with less than his full power" (357). But the difficulty of an item is not the only information upon which an item can be classified as appropriate. It is also important to test if an item can effectively discriminate between strong and weak test takers. This is known as item discrimination. "[I]tem discrimination is the ability of an item to distinguish between good and poor students; an item with good discrimination will be responded to correctly by many more good students than poor students, most of whom will get the item wrong" (Alderson 2000: 87).

4.2.   Distractor Characteristics

Distractors receive their name from their function, that is to distract those test takers who "are in doubt about the correct answer" (Gronlund 1965: 140). The problem with distractors is that they should be appealing to the test takers but not too appealing and confusing to the test takers. It should always be avoided to provide hints that indicate the correct answer. These hints regard the length of the alternatives, the position of the correct solution or the use of opposites of the

correct answer (see Kitao et al. 1999: 38-39). Henning (1987: 44-48) as well as Heaton (1983: 16-19) identified several problems with distractors that should be avoided in language testing.

The length of the different alternative answers should be approximately equal, otherwise test takers might always pick the option that is the longest and most elaborate one. Another clue that can imply the correct answer to an item is if two of the options are convergent, i.e. are very similar to each other whereas the other two options differ extremely. Inconsistent distractors represent an extreme form of response clues. The term refers to items where the correct option "differs from the other options both in lexical root and in part of speech" (Henning 1987: 45). Furthermore the number of options is an important factor in a successful item. If there are too few alternative answers, as in true or false items, a very large number of items is needed in order to make the test reliable. If there are too many options in one item the test taker may be confused because more than one option seems to be a suitable answer. Henning (1987) and Heaton (1983) also mention nonsense distractors by which they refer to grammatically incorrect alternatives. They should be avoided because they can have a negative backwash effect on the learners because they may adopt those incorrect forms. Another factor that should be considered in the development of an item is the avoidance of redundant information as in the following example from Heaton (1983: 17).

Version 1:

The word "astronauts" is used in the passage to refer to
A. travellers in an ocean liner
B. travellers in a space ship
C. travellers in a submarine
D. travellers in a balloon

Version 2:

The word "astronauts" is used in the passage to refer to travellers in
A. an ocean liner
B. a space ship
C. a submarine
D. a balloon

In the first version of this item it repeats the same sentence fragment in every option, which makes it inefficient because the test taker has to read the whole sentence from the beginning every time without gaining new information. Whereas

46

the second version asks the same question but this time most of the information is part of the item stem, which makes it more efficient and readable.

4.3. Classical Multiple Choice Test

4.3.1. Definition

Multiple choice test items are regarded as being "the most widely applicable and useful type of objective test item" (Gronlund 1965: 140). They can be defined by several factors. Firstly multiple choice tests characteristically fall into the category of selected response. This means that there is no need that the test taker produces language because he or she only has to select the best solution. A multiple choice test offers the possibility to score a test objectively and ensure fair scoring and grading. Furthermore it "can measure at both the knowledge and understanding levels" (ibid), which assures that a test constructor is flexible in what to measure including various complex learning outcomes as well as simple learning outcomes.

4.3.2. Composition

A classical multiple choice item usually consists of two main parts: the prompt or stimulus and three or more possible options, choices or answers. Another differentiation between multiple choice items is suggested by Gronlund (1965: 141). There are items where the learner has to choose only one correct answer from the alternatives. Correct answer types can appear in the form of a direct question as well as in the incomplete-statement form. But not all items can test knowledge in such precise way so that there is only one correct answer. This type is called best-answer type of multiple choice item. These items tend to give a number of plausible reasons and the task of the test taker is to identify the best possible answer. Typically best-answer items include questions of the why and how variety because they offer the possibility to present a more complex problem to the test takers.

Furthermore Gronlund (1965:149-157) suggests a number of guidelines that help to optimise multiple choice items, that are listed in order to sum up the most important factors to consider in a well-written multiple choice item. The item stem should be meaningful in itself without the alternative answers. Therefore, statements where the test taker has to fill in a blank space is, according to

Gronlund, not a proper multiple choice item. Concerning the item stem, it should present a concrete problem and most of the textual information should be included in the stem itself and not in the alternative answers. The formulation of an item should not make use of negative forms or double negative forms in order to prevent the test takers from overlooking words that are typical for negative statements like 'no' or 'not' or get confused by the formulation itself. As far as the answers to an item are concerned, every answer choice that is part of an item should be plausible as well as grammatically correct. A distractor has the purpose of appearing attractive to the test taker and distract from the correct option. "If a distracter is not selected by anyone, it makes not contribution to the functioning of the item and should be eliminated or revised" (Gronlund 1965: 154). Moreover, every item should present a definite problem therefore there can always be only one correct answer to every item. Additionally, several factors can provide irrelevant clues. these comprise verbal associations between the prompt and the correct answer or the length of the correct answer and the alternative options. Those clues should be avoided because they make it possible for test takers to solve an item without understanding it. Over and above that the position of the correct answer should vary from item to item, i.e. "the correct answer should appear in each of the alternative positions approximately an equal number of times, in random order" (Gronlund 1965: 156). Besides answer choices that include real information that has to be reviewed in terms of correctness, there is the option of creating alternative answers like 'none of the above' or 'all of the above'. Gronlund (1965: 157) however suggests that these special alternatives should only be used in appropriate situations.

4.3.3. Types of Questions

There are basically three types of questions that were identified by Johnson (1987). Firstly, textually explicit questions "are those where both the question information and the correct answer are found in the same sentence" (Alderson 2000: 87). Secondly, textually implicit questions on the contrary call for a connection of information from different sentences. Thirdly script-based questions are questions where the test taker's knowledge about a topic is necessary to solve the item because the information needed is not included in the stem. Those three types of

question differ in terms of difficulty, whereby textually explicit questions are regarded the easiest type (ibid). Gronlund (1965) suggests two different types of questions. On the one hand a test constructor has the possibility to pose a direct question, which is regarded as being the easier item to write. On the other hand an item can take the form of an incomplete statement, which has the advantage that it can present a more complex, well-defined problem to the test taker (see 141).

4.3.4. Uses of Multiple Choice Items

Because of their flexible nature multiple choice items are considered to be the most widely applicable and versatile type of test items and are therefore mostly used in standardised tests. Consequently multiple choice questions can measure a variety of learning outcomes, which will be examined in the following section. As it was mentioned above, multiple choice items can measure knowledge outcomes as well as outcomes at the level of understanding. "[T]ypical uses of the multiple-choice form in measuring knowledge outcomes" (Gronlund 1965: 142) include knowledge of terminology, knowledge of specific facts, knowledge of principle, knowledge of methods and procedures and more. When testing the knowledge of terminology a test taker is asked to prove that he or she understands "a particular term by selecting a word, which has the same meaning as the given term or by selecting a definition of the term" (Gronlund 1965: 142f.) An item that aims to test this ability could look like the example taken from Gronlund (1965: 143). The correct answer in all of the following examples will be indicated by this symbol *.

Which one of the following words has the same meaning as the word egress?
A Depress
B Enter
C Exit*
D Regress

In terms of testing the knowledge of specific facts test constructors often use questions of the why, what, when and where variety. As with the previous example there is only one correct answer the test taker has to choose. An example of an item that is able to test the knowledge of specific facts comes again from Gronlund (ibid).

Who was the first United States astronaut to orbit the earth in space?

A   Scott Carpenter

B   John Glenn*

C   Virgil Grissom

D   Alan Shephard

Testing the knowledge of principle is a more complex task than the two uses that have already been mentioned and it is therefore more difficult to construct such items. An item that is testing the knowledge of principles could be formulated in the following way, as Gronlund (1965: 144) suggests.

According to the principle of capillary action, fluids

A   enter solutions of lower concentrations.

B   escape through small openings.

C   pass trough semipermeable membranes.

D   rise in fine tubes.*

The measurement of knowledge of methods and procedures involves a diverse area of testing subjects that include knowledge of common social practices, knowledge used in problem solving or knowledge of governmental procedures (ibid).

To make treaties, the President of the United States must have the consent of the

A   Cabinet

B   House of Representatives

C   Senate*

D   Supreme Court

These examples show what is possible with multiple choice items in terms of measuring knowledge outcomes in learning. But there are far more areas of knowledge that can be tested with multiple choice tests and accordingly, many more forms such items can take.

The other area of measuring learning outcomes involves the level of understanding where more complex problems can be presented to the test takers. One crucial element in these kinds of items is the applications and interpretations have to be new to the test takers. So in the following examples taken from Gronlund (1965: 145-147) this aspect of novelty is taken for granted.

The ability to apply facts and principles is tested by asking the test takers "to identify its correct application in a situation which is new to the pupil" (Gronlund 1965: 145).

Directions: In each of the following sentences circle the word that makes the sentence correct.
1. This is the boy   which asked the question.
                       who*
                       whom

2. This is the dog   which* he asked about.
                       who
                       whom

Furthermore it is possible to test the ability of the learners to interpret cause-effect relationships by "asking pupils to interpret various relationships between facts" that are presented in the item and "to identify the reason which best accounts for it" (Gronlund 1965: 146). These items can ask questions like:

Bread will not become moldy as rapidly if placed in a refrigerator because

A   cooling retards the growth of fungi.*

B   darkness retards the growth of mold.

C   cooling prevents the bread from drying out so rapidly.

D   mold requires both heat and light for best growth.

Another important learning outcome that can be measured by the format of multiple choice testing is the ability of the learner to justify methods and procedures. This can be measured by presenting the learner "with several possible explanations of a method or procedure and asking him [or her] to select the best one" (Gronlund 1965: 147).

Why is adequate lighting necessary in a balanced aquarium?

A   Fish need light to see their food.

B   Fish take in oxygen in the dark.

C   Plants expel carbon dioxide in the dark.*

D   Plant grow too rapidly in the dark.

4.3.5. Advantages & Disadvantages

As it has already been mentioned a main advantage of multiple choice items is the great variety of knowledge that can be tested with them, thus gives a test constructor with a very high flexibility concerning the subject areas that can be tested. Furthermore there is no need for special expertise to score them because there is only one correct answer. In addition the evaluation process is much less time consuming than other test formats.

However there has been a lot of critical voices against multiple choice formats. Popham (2002) criticises that in multiple choice items test takers are only asked to select one correct answer from a set of alternatives and are therefore not expected to generate correct answers themselves (see 134). On the other hand Van Bierkom (2009: 91) names the recognition format of multiple choice items as an advantage because it is more sensitive to partial knowledge because the correct answer has to be one of the alternatives. Another aspect that has often been criticised is the guessing factor that comes with multiple choice testing formats. That is probably the reason why Oosterhof (2009) mentions the nickname "multiple-guess" (93). But the guessing factor can be reduced by two easy methods. Firstly, if the number of possible answers is increased to four or five alternatives the guessing factor will be reduced. Secondly, if a test includes a large number of items the total number of items will also have an effect on the guessing (see Popham 2002: 130). Another disadvantage of multiple choice items is "the difficulty of locating a sufficient number of incorrect but plausible distracters" (Gronlund 1965: 149; Van Bierkom 2009: 91).

4.4. True/False/Not Given

4.4.1. Definition

One of the most common and popular option in testing reading is the true/false/not given format, alternate-choice format (Oosterhof 2009: 112), alternative-response item (Gronlund 1965: 127ff) or binary choice format (Brown et al. 2002: 65). It can be considered as a form of multiple choice testing with the restriction of having only two distractors (see Hughes 2003: 144). Therefore Alderson et al. (1995) refer to items of that kind as "dichotomous items" (51) and include true or false items into their definition as well as Yes or No questions.

4.4.2. Composition

A typical true or false item consists of a statement and presents the test takers with the task to mark the statement as true or false. A variation of a true or false item is to employ a further alternative such as 'not given', which can reduce the guessing factor. But I will come back to this criticism later in this section. The most frequent use of the true or false format is to measure the learner's "ability to identify the correctness of statements of fact, definitions of terms, statements of principles, and the like" (Gronlund 1965: 127) A simple true or false item can look like the following example taken from Gronlund (ibid)

> Directions: Read each of the following statements. If the statement is true, circle the "T". If the statement is false, circle the "F".
>
> *T   F   1. The green coloring material in a plant is called chlorophyll.
>
> T   *F   2. The corolla of a flower includes petals and sepals.
>
> *T   F   3. Photosynthesis is the process by which leaves make the food for a plant.

In order to reduce the guessing factor in this item, there is the possibility of including another option like 'not given'.

Another important use of the alternative-response format is to find out if a learner is able to differentiate between a fact and an opinion, as in the following example (ibid).

> Directions: Read each of the following statements. If the statement is true, circle the "T". If the statement is false, circle the "F". If the statement is an opinion, circle the "O".
>
> *T   F   O   1. The earth is a planet.
>
> T   *F   O   2. The earth revolves around the moon.
>
> T   F   *O   3. There are no plants or animals on Mars.

Furthermore true or false items enable a test constructor to measure a learner's ability to understand relationships of cause and effect as well as basic logical aspects.

4.4.3. Advantages & Disadvantages

One advantage of true/false/not/given items is that they do not require the test takers to produce language themselves and therefore a large number of items can be answered in a relatively short amount of time. Another advantage listed by

Brown et al (2002: 66) is that true or false items test if the test taker has understood a specific point or not, even if the learning outcomes that can be tested are relatively limited in the knowledge area (see Gronlund 1965: 130). Furthermore these items allow the test constructor to guide the test takers by asking questions about specific points of a text or passage. "[T]rue-false and true/false/not given tasks can be used as a valuable teaching device with which the students' attention is directed to the salient points in the [text] by means of the true/false items (Heaton 1988: 114).

Although true/false/not given items seem to be rather easy to write there are several problems that can arise during their construction. A clear disadvantage of true or false items that is mentioned by many experts is that the test takers have to choose between two options, which makes it relatively easy to guess the correct answer without knowing it. Arthur Hughes (1998) classifies true or false questions as "a variety of multiple choice, with only one distractor and a 50 per cent probability of choosing the correct response by chance!" (121) Norman Gronlund (1965) agrees and states that "[w]ith only two alternatives a pupil has a 50-50 opportunity of selecting the correct answer on the basis of chance alone" (130) Therefore Alderson et al. (1995) suggest that "[i]n order to learn anything about a student's ability, it is necessary to have a large number of such items in order to discount the effects of chance" (51) Another solution to this problem that was previously mentioned is to provide a third option that is referred to as "insufficient information" (Kitao et al. 1999: 39) or "not given". This addition can reduce the guessing factor and adds difficulty to the item. Further critics of true or false items claim that the test taker "may be able to recognize a false statement as incorrect but sill not know what *is* correct" (Gronlund 1965: 129). This problem is often eliminated by changing the incorrect parts of a statement or by rewriting the whole statement. This however, adds difficulty to the evaluation and scoring process (ibid). Another factor that item constructors should be aware of is that some items might "require information outside of passage" (Kitao et al. 1999: 40) in order to solve the problem. Some test takers might not be able to answer the item not because they do not understand the passage but because they do not have the necessary background knowledge. As with the conventional multiple choice items the phrasing of the alternative answers requires special attention. Regarding the formulation of true or false items the statements should not be written too broad or

too trivial. Also long and complex sentences should be avoided as well as negative or double negative formulations. As with conventional multiple choice items the correct as well as the incorrect statements should be of equal length (see Gronlund 1965: 131-133). Additionally the incorrect answers "should be based on some possible misunderstanding of the passage" (Kitao et al. 1999: 40). It is important to bear in mind that not every subject matter field is predestined to be tested by true or false items. True or false items are most suitable for "situations where there are only two possible alternatives (e.g. right, left; more, less; 'who', 'whom'; and so on)" (Gronlund 1965: 131).

4.5. Matching

4.5.1. Definition

Matching exercises are also known as 'multiple matching' (Alderson 2000: 215) and represent another objective testing method that uses a fixed response format. Because there is usually a set of choices of options in each matching exercise, Alderson (2000: 218) and Oosterhof (2001: 132) consider them as special form of multiple choice items. Matching exercises are "limited to the measurement of factual information based on simple associations" but they provide "a compact and efficient method of measuring such simple knowledge outcomes" (Gronlund 1965: 134).

4.5.2. Composition

The conventionally known form of a matching item consists of two parallel columns. These two columns include words, phrases, numbers or symbols that need to be matched against each other. Part one can also be referred to as column A and it contains items that are called premises. The items in column B are called responses. The learners that are confronted with matching exercises have to "identify the pairs of items" (Gronlund 1965: 134). A special form of matching exercises are called 'imperfect match'. This means that there are more words or phrases provided to select from in column B than in column A. The reason for this imbalance is to prevent the test takers from identifying the last pair by elimination only. The method of imperfect match is very commonly used in order to make the matching exercises more difficult for the test takers. A very widespread use of

multiple matching exercises can be found in tasks where the test taker has to read one long text that has been separated into several parts. The corresponding questions contain information that can be found in the text. The test takers' task then is to identify which of the parts of the text contain the information in question. Furthermore matching items may require the test taker to match headlines that are provided against passages from a text. In order to avoid guessing or identifying the last headline by elimination only there should be more headlines to choose from than there are paragraphs.

A typical example for a matching item that was constructed with regard to the characteristics mentioned above could look like the following example taken from Gronlund (1965: 136).

Directions: On the line to the left of each achievement listed in Column A, write a letter of the man's name in Column B who is noted for that achievement. Each name in Column B may be used once, more than once, or not at all.

Column A	Column B
(A) 1. Invented the telephone.	A Alexander Graham Bell
(B) 2. Discovered America.	B Christopher Columbus
(C) 3. First United States astronaut to orbit the earth.	C John Glenn
(F) 4. First President of the United States.	D Abraham Lincoln
	E Ferdinand Magellan
	F George Washington
	G Eli Whitney

Another way of presenting a matching exercise is not to ask a question that has to be answered by selecting the best option from a set of alternatives. In that case, parts of a reading paragraph are left out and the test takers are required to choose which sentence can best fill the blank space. Thus paragraphs and not statements are matched against each other. These examples can look like the following matching exercise taken from "Make Your Way Ahead 7" (p. 23). This book was chosen because I have already used this textbook and was already familiar with the form of presentation. There is a great variety of schoolbooks available in Austria, which may or may not have the same structure and range of examples or tasks as "Make Your Way Ahead 7", but it would go beyond the scope of this thesis to analyse all of them. Therefore the examples shown only represent a small part of the available material.

The wife's tail

They used to say that behind every successful man was a devoted wife. Inviting the boss home for dinner was not merely a chance to meet in a relaxed situation, but it was a test of the wife's skills and suitability as a partner to a rising management star.

| 1 | |

These days, however, top executives meet in hotels and restaurants. And I, of course, am glad.

| 2 | |

For if we were back in the dark old days, I would be panicking in the kitchen in case my soufflé wouldn't rise while our guests were sipping Martinis in the lounge. It would be me who handed out after-dinner mints and prayed I'd not done or said the wrong thing.

| 3 | |

Consequently I am invited occasionally to accompany my wife on a team-building weekend away. The explanatory letter begins promisingly enough, "Dear Partner …"

| 4 | |

For example, one weekend we were called to a castle in Ireland. When the time came, I tried to find a medium level of involvement, yet not make myself feel ridiculous. Nevertheless, I was glad I had hired a car and so was able to travel independently, stopping off at points of interest along the way instead of heading for the shops.

| 5 | |

My wife has changed jobs recently so there will be even more entertainment in the future. On the whole I am looking forward to it. After all, every successful woman deserves a supportive man.

A First I am asked to tick the options I would prefer; beauty treatments, perhaps, or aromatherapy? Second, would I prefer a visit to the hair salon or ladies' golf?

B For I am the modern equivalent in trousers and stubble – the Corporate Husband.

C Because, although this change reflects the fact that women are too busy living their own lives to play second fiddle, this development helps me personally and to a large degree.

D Although, superficially, times have changed some of the old terror remains for the poor partner of the ambitious business executive.

E Sometimes though, the frustration of not being recognized for my role in the partnership can cause me sleepless nights and I plot my own assault on the corporate world.

F I joined the other partners for lunch and admired their purchases. Then, as they headed back for their facials and foot-massage, I went for a manly walk along the coast.

Make Your Way Ahead 7 | 23

Figure 2: Exercise "Make Your Way Ahead 7", Unit 4, p. 23.

This top part presents a reading passage where the test taker has to choose which of the extracts that follow the text fit best into the numbered gaps. The second part presents the extracts that need to be filled into the numbered gaps in the text. This matching exercise has been constructed as an imperfect match. One of the extracts does not fit in any of the gaps and is therefore used as a distractor.

4.5.3. Advantages & Disadvantages

Like the two previously described test formats matching exercises have advantages as well as limitations. One advantage "is its compact form, which makes it possible to measure a larger amount of related factual material in a relatively short time" (Gronlund 1965: 135). Popham (2002) agrees and stresses, that "their compact form takes little space on a printed page, thus making it easy to tap a good deal of information efficiently" (140). Furthermore the guessing factor in matching exercises is relatively low, compared to multiple choice items (see Brown et al. 2002: 65). On the one hand matching items are basically easy to construct but on the other hand it is not the easiest task to find plausible responses for each of the premises. Another limitation of matching exercises is the restriction of the kind of information that can be measured with them, which is mainly factual information. Moreover the possible distraction of test takers by options they would not have thought of if they would not have been included in the item which might cause that they cannot solve the problem (see Alderson 2000: 219).

4.6. Comparative Summary

The three selected-response formats that have been described in great detail will now be summarised and compared with regard to their advantages, disadvantages and limitations. Opposed to other test types like constructed-response or personal response, selected-response test types are easier in terms of administering and scoring. Their construction may be more complicated than other test formats but the scoring itself is not only less time-consuming but also absolutely objective. Multiple choice items require the test taker to study some kind of language material and afterwards select the correct answer from three or more options. They can effectively test many language aspects including reading and listening skills, grammar knowledge and phonemic discrimination (see Brown et al. 2002: 68-69). Given that some guidelines are incorporated into the design of multiple choice items they can ensure the objective testing of the language abilities mentioned above. A summary of those guidelines include the grammatical consistency of the alternatives and the avoidance of unintentional clues. Binary-choice items, i.e. true or false items, put a focus on the assessment of the abilities of a test taker to differentiate between two alternatives. Therefore it provides "simple and direct indices of whether a particular point has been comprehended or not" (Brown et al.

2002: 66). Concerning the formulation, long and complex sentences or questions should be avoided. The aspect that is most often discussed in connection to true or false items is the problem of the high guessing factor because test takers have a 50 per cent chance of solving an item without understanding it. Variations of multiple choice items can reduce the guessing factor drastically. By having to choose from more than two options only "[m]ultiple choice items reduce the 50% guessing factor found in binary-choice tests to 33%, 25%, or 20% depending on whether three, four, or five options are used" (Brown et al. 2002: 68). The guessing factor is also an issue related to matching items. In order to prevent the test taker from guessing the correct order or arrangement there should be more options presented in column B than there are prompts in column A. Matching exercises are fairly limited because they can only measure the test takers' ability "to associate one set of facts with another, or one set of definitions with the words they define" (Brown et al. 2002: 67). Summing up, criterion-referenced test items, no matter what form they have, should be designed in such a way that they reduce ambiguity, i.e. they should include concrete, comprehensible instructions and make clear what the test taker is supposed to do.

For the analysis that will follow in the discussion of the research questions I want to analyse the items based on a selection of the most important characteristics that have been collected from several different publications (Kitao et al. 1999; Henning 1987; Gronlund 1965).

- Are the distractors plausible?

- Are there convergent options?

- Is most of the information part of the item stem or is it part of the alternative answers?

- Are the words used in the options taken from the stem or have they been rephrased?

- Are there negative formulations used in the item?

- Is the length of the alternative options approximately the same? Is the correct answer positioned randomly in each of the positions approximately an equal number of times?

- Is the question based on a possible misunderstanding of the text?

- Does the test taker require information outside of the reading passage in order to solve it?

- Is the item constructed as an imperfect match?

These factors will be part of the analysis in the discussion of research question 2. Obviously not all of these characteristics can be applied for every type of multiple choice format because they all have very specific and individual features.

5. Research Questions

5.1. What different ways of preparing learners for objective test formats are there?

Broad guidelines that should help students to prepare for the standardised or New Matura can be found on the homepage of the university of Innsbruck (http://www.uibk.ac.at/srp/aufgaben_englisch.html, 10 June 2012). These study tips include rather universal strategies that learners can follow in order to be well prepared. Careful and exact reading of the instructions that come with every test item is the basis for understanding the task. It is recommended that the learner tries to locate every text passage that fits the given question and answer the question afterwards. Advice that concerns multiple choice questions specifically suggests that the learner should not read the questions concerning a text before having read the whole text. The test taker is advised to formulate his or her own answer before identifying the correct answer among the alternative choices. Preparation includes the use of such reading passages and the items that fit the content of this passage in the classroom setting.

The matching exercises included in the newly developed material represent two different versions. In some matching exercises test takers are required to match headlines to short passages of a text in others the task is to answer short questions concerning a text by matching sentence fragments. Matching exercises test the ability of the test taker to understand the opinion or the detail expressed in the question. The text has to be read several times in order to match the questions and the answer options. One way of preparing for multiple matching tasks is to read a text and highlight opinions, facts and ideas separately.

Guidelines to prepare for true or false items include careful reading because in these items there are mostly only small details that are not conform to the text. It is explained that if a statement is true the whole statement has to be completely true. The same applies for false statements, which have to be totally incorrect. Not given statement are those that present completely new ideas that are not part of the text. It is suggested to read quickly through the whole text for a basic understanding. Afterwards the statements should be read carefully. In order to decide whether a statement is true, false or not given the text has to be re-read with the aim to find the relevant passage.

It is obvious that all of these guidelines can be transferred to the work that is done in the language classroom. Therefore the next important question that has to be answered is concerned with the test taker preparation via textbooks.

5.2. Does the New Matura use examples that are comparable to the textbook examples?

The following discussion of research question 2 presents the focus of my investigation. Generally speaking the examples in the New Matura include the following objective test formats: multiple choice questions, matching, short answer questions, true/false/not given, editing and word formation. The New Matura uses multiple choice items that only have one correct solution. The questions that are asked are always related to a text that is presented previously. The matching exercises used in the New Matura are designed in a way that there are more words or phrases in Column B that can be matched to Column A. There is the possibility that more than one answer is correct. If that should be the case it is clearly indicated. The examples that can be found on the homepage of the university of Innsbruck and www.bifie.at provide a good illustration of the presentation of the items in the New Matura. They will be analysed in terms of several characteristic that have been established in section 4.6. Also textbook examples will be included in the analysis in order to find out if the textbook examples are comparable to those used in the actual Matura exam. The following grid shows an overview of the parts that are to be analysed for each item in each task. There may be slight adjustments made for the individual testing formats.

Characteristics	Item 1	Item 2	Item 3
a. Are the distractors plausible?			
b. Are there convergent options?			
c. Is most of the information part of the stem?			
d. Does the item use the same wording as the text?			

e. Are there negative or double negative formulations?			
f. Is the length of the alternatives approximately equal?			
g. Is the correct answer positioned randomly?			

With the help of these characteristics a comparative analysis of objective testing tasks from textbooks and the New Matura will be established. For my analysis I am going to use examples from "Make Your Way Ahead 8" as well as samples taken from the testing booklet provided by the university of Innsbruck and bifie. I am going to analyse representative examples for each of testing formats from the New Matura individually before comparing them to examples from a recently revised textbook called "Make Your Way Ahead 8" in order to answer the question if the New Matura use examples that are comparable to the text book examples in terms of formulation, difficulty and clues for finding the correct answer. The examples have been chosen randomly, without any preference concerning the topic. But the examples that will be analysed all have approximately the same length. It is important to stress the fact, that the scope of this thesis does not allow for a general conclusion concerning objective testing formats. The analysis will present a limited comparison that holds true for the examples included in this thesis. Using the above grid, only the most interesting points will be discussed. The abbreviations Y and N that will be used in the analysis stand for 'Yes' and 'No' as answers to the questions posed in the left column.

5.2.1. Multiple Choice Tasks

In the latest version of "Make your Way 8" there are several multiple choice tasks that should prepare the learners for the New Matura. These will be used for the following analysis as well as examples from the test booklets provided by bifie and the university of Innsbruck. The instructions for the multiple choice items

is formulated the same way for every task from the textbook as well as from the New Matura, and looks like the following:

> **Read the extract from a novel, then choose the correct answer (A, B, C or D) for questions 1-6 and put a X in the correct box.**

All tasks presented in the textbook consist of a stem that introduces a specific problem in the form of a direct or indirect question. There are no incomplete or complete statements included.

In order to distinguish the various tasks and items from each other they will be referred to as the following table shows. The various tasks will be listed with the according headline. Tasks 1 – 4 have been taken from the "Make Your Way Ahead 8" and tasks 5 and 6 have been taken from the test booklets provided by the bifie.

Task 1 "Keeping it up"	**Task 2** "The Warning"	**Task 3** "Reality TV"	**Task 4** "My flight from fantasy"	**Task 5** "A short history of tractors in Ukrainian".	**Task 6** "The price of the Olympics"
includes Item 5A-5H	Includes Item 6A-6F	includes Item 1A-1F	includes Item 2A-2F	includes Item 3A-3F	includes Item 4A-4F

5.2.2. Multiple Choice Tasks – New Matura

The next section will analyse examples from the New Matura that are already available online. Furthermore tasks will be taken from a testing booklet that has been provided by the bifie and the university of Innsbruck. This booklet represents a selection of various multiple choice tasks that could be part of the standardised Matura.

5.2.2.1. Multiple Choice Task 1

The first multiple choice task that is analysed is entitled "Keeping it up" and it includes eight items based on the reading passage, each with four answer choices. Table 1 provides an overview of multiple choice task 1.

Characteristics	Item 1A	Item 1B	Item 1C	Item 1D	Item 1E	Item 1F	Item 1G	Item 1H
a. Are the distractors plausible?	Y	Y	Y	Y	Y	Y	Y	Y
b. Are there convergent options?	N	N	N	N	N	N	N	N
c. Is most of the information part of the stem?	Y	Y	Y	Y	Y	Y	Y	Y
d. Does the item use the same wording as the text?	N	Y	Y	N	Y	N	Y	Y
e. Are there negative or double negative formulations?	N	N	N	N	N	N	N	N
f. Is the length of the alternatives approximately equal?	Y	Y	Y	Y	N	N	N	Y
g. Is the correct answer positioned randomly	Y							

Table 1: Overview MC Task 1.

If we have a closer look at characteristics a, b and c it is noticeable that all items belonging to MC task 1 were constructed very carefully. All of the distractors are plausible because they spin around the same topic and are in some way part of the text. There are no convergent options included in the answer alternatives and the items were designed in the best possible way so there is no need for rephrasing them. Most of the information is part of the stem and there is no repetition in the alternative options.

Concerning the use of words that come directly from the text it is evident that items 1A, 1D, and 1F do not include words that relate directly to the reading passage. On the contrary items 1B, 1C, 1E, 1G and 1H use words that have been

adopted from the text. Items 1B and 1C only use the words from the passage in some distractors whereas items 1E, 1G and 1H use the words from the text in the distractors as well as in the stem. Interestingly, most of the words are used in the distractors but not in the key in order to make the distractors more attractive to the test takers.

The test items of task 1 do not include any negative or double negative formulations. Moreover, length clues are not part of these items. There are some items where the alternatives are not of equal length, but the longest and most complex answer is never the key to one of the questions. Accordingly it is assumed that the most elaborate answer in items 1E, 1F and 1G are used to distract the test takers.

Rory headed the <mark>little boys' ball</mark> in the <mark>air</mark>

A and proved how skilful he was at football

B although he was annoyed at first when the <mark>ball</mark> hit his <mark>table</mark>.

C because the <mark>boys</mark> wanted him to.

D although he was feeling so hot and <mark>angry</mark>.

These longer answers also include the previously mentioned words that have been taken over from the reading passage. The words that derive from the text are marked in yellow and can be found in the stem as well as in all three distractors. Only option A, which is the key to item 1G, does not contain any of the content words. Over and above that task 1 serves as an ideal example of key distribution. The correct answer is placed randomly in each of the possible positions an equal number of times.

Item 1A	D	Item 1E	D
Item 1B	C	Item 1F	C
Item 1C	B	Item 1G	A
Item 1D	A	Item 1H	B

The second task from the New Matura booklet is called "The Warning" and it includes six items referring to the text. Table 2 provides an overview of MC task 2.

Characteristics	Item 2A	Item 2B	Item 2C	Item 2D	Item 2E	Item 2F
a. Are the distractors plausible?	Y	Y	Y	Y	Y	Y
b. Are there convergent options?	N	Y	N	N	N	N
c. Is most of the information part of the stem?	Y	Y	Y	Y	Y	Y
d. Does the item use the same wording as the text?	Y	N	Y	N	Y	Y
e. Are there negative or double negative formulations?	N	N	N	N	N	N
f. Is the length of the alternatives approximately equal?	Y	N	Y	N	Y	N
g. Is the correct answer positioned randomly?	N					

Table 2: Overview MC Task 2.

All the distractors constructed for MC task 2 are plausible in connection to the text. They include information that is part of the reading passage, thus they fulfil their task of distracting the test takers. Concerning characteristic b it is noticeable that item 2B includes two convergent distractors, where one of them is the key to the question.

Tom's opinion of the new cleaner was that

A his conversation was rather boring.

B he was the best cleaner they ever had.

C he was good but he got in the way.

D his work was first class. *

Distractor B could also be correct, because neither the formulation "first class" nor "best cleaner they ever had" is part of the text. The test taker has to decide, which answer is the best choice according to the text.

Regarding the formulation of the items it is noticeable that they have been designed carefully. None of the items calls for a revision because most of the information is part of the stem.

In most of the items the task developers used the same wording as in the reading passage, except for items 2B and 2D. In item 2B only the stem includes the word "cleaner" that comes directly from the text. And in item 2D only the word "building" is taken from the passage. Therefore these items can be considered as being more difficult than the others because the test taker has to re-read the corresponding sections in order to find the correct solution.

As I have mentioned before, MC task 2 has been constructed carefully so it is not surprising that there are no negative or double negative formulations included. But concerning the length of the distractors it is obvious that items 2B, 2D and 2F include length clues. In the case of item 2F, the longest and most elaborate answer is also the key to the question. In items 2B and 2D the reverse is the case, the shortest answer options form the key to the questions. Thus these items could be revised because most of the criticism concerns items 2B and 2D.

With regard to the position of he key it is clear that not all of the options have been used equally. Options A, B and C only have been used once whereas option D was the key three times. Therefore it would be advisable to rearrange the options in order to reach a random order of answer keys.

Item 2A	A	Item 2D	C
Item 2B	D	Item 2E	D
Item 2C	B	Item 2F	D

5.2.3. Multiple Choice Tasks – Textbook Examples

5.2.3.1. Multiple Choice Task 3

In order to compare these examples with the previously analysed examples from the New Matura testing booklets, the same grid will be used.

Multiple choice task 3 is entitled "Reality TV: a dearth of talent and the death of morality" and contains six items on the reading passage with four answer choices. Table 3 presents an overview of the characteristics of MC item 3.

	Characteristics	Item 3A	Item 3B	Item 3C	Item 3D	Item 3E	Item 3F
a.	Are the distractors plausible?	Y	Y	Y	Y	Y	Y
b.	Are there convergent options?	N	N	N	N	N	N
c.	Is most of the information part of the stem?	Y	N	N	Y	Y	Y
d.	Does the item use the same wording as the text?	N	N	N	N	N	N
e.	Are there negative or double negative formulations?	N	N	N	Y	N	N
f.	Is the length of the alternatives approximately equal?	Y	N	Y	Y	N	Y
g.	Is the correct answer positioned randomly?	N					

Table 3: Overview MC Task 3.

Concerning characteristic c it is noticeable that only items 3B and 3C could be formulated in a more efficient way by changing the direct question to an incomplete statement. The difference becomes obvious if the original item 3B is compared to an alternative formulation. The original version is presented in the textbook as follows:

Why does he believe reality TV shows are worth studying?

A everyone is talking about them.

B because their popularity offers insight into the human race.

C because they are the future of TV.

D because they are everywhere.

Following the characteristics of a well designed item that have been established in section 4.3. an alternative version could look like this:

The author believes that reality TV shows are worth studying because

A everyone is talking about them.

B their popularity offers insight into the human race.

C they are the future of TV.

D they are everywhere.

Interestingly, none of the items uses words that come directly from the reading passage, which makes it more difficult for the test takers to identify the correct option. One feature that is part of item 3D is that is makes use of a negative formulation.

Which of the following activities does the author <u>not</u> refer to in his list of typical reality show behaviour?

A preparing meals.

B arguing with other contestants.

C taking clothes off.

D talking to the camera.

The learner is required to identify the only option that is not included in the text. The negative formulation is emphasised by underlining the word "not" in order to ensure that the test takers do not fall in danger of over reading it.

In items 3B and 3E it is obvious that one option is longer than the other three alternatives, what might provide a clue for the test taker. In the case of those two questions the longest option is the correct answer.

Concerning the spreading of the correct answers it is clear that MC task 3 does not make use of every position. The keys are distributed as follows. Obviously, position D is never used for the correct option, whereas position C is used three times and position A is used twice.

Item 3A	C	Item 3D	A
Item 3B	B	Item 3E	A
Item 3C	C	Item 3F	C

5.2.3.2. Multiple Choice Task 4

The second task that is analysed is entitled "My flight from fantasy" and contains six items on the reading passage with four answer choices. Table 4 presents an overview of the characteristics of MC task 4.

	Characteristics	Item 4A	Item 4B	Item 4C	Item 4D	Item 4E	Item 4F
a.	Are the distractors plausible?	Y	Y	Y	Y	Y	Y
b.	Are there convergent options?	Y	N	N	N	N	N
c.	Is most of the information part of the stem?	Y	Y	N	Y	Y	N
d.	Does the item use the same wording as the text?	Y	Y	N	N	N	Y
e.	Are there negative or double negative formulations?	N	N	N	N	N	N
f.	Is the length of the alternatives approximately equal?	Y	N	N	Y	N	N
g.	Is the correct answer positioned randomly?	Y					

Table 4: Overview MC Task 4.

As we can see in table 4, every possible answer is plausible for the reading passage. Clearly they have been carefully designed in order to fulfil the function of a good distractor. As far as Item 4A is concerned we can see that it includes convergent answers. Therefore a closer look on that question is necessary.

From what she tells us in the first paragraph, which of these might the author be happy doing?

A Eating a steak.

B Reading a map.

C Eating Marmite in Italy.

D Talking to someone she doesn't know.

Noticeably answer options A and C are convergent because they concern the same topic. It is obvious for a learner that one of those options has to be the correct solution. As a possible solution to this problem the alternative answers should either be completely different to each other or all of the distractors should be concerned with the activity of eating.

As far as the characteristic c is concerned it is notable that not all of the items have been formulated ideally. Items 4A, 4B, 4D and 4E are formulated in such a way that most of the information is part of the item stem. But items 4C and 4F could have been written in a more efficient way. Both items repeat a sentence fragment in every answer option, which makes the question more time consuming to read. The question could be changed similarly to MC task 3, by changing the direct question to an incomplete statement in order to avoid repetition.

In items 4A, 4B and 4F the test taker will be able to identify words that come directly from the text. This repetition of the words from the reading passage might represent a response clue for some test takers to choose the option that uses the same wording. Item 4F serves as an example.

Why does the author believe she will not be taking any more long flights?

A Because she is concerned about the damage flying does to the environment.

B Because she is afraid that other journeys won't be as good.

C Because she is afraid of bird flu.

D Because she's still really afraid of flying.

The words "damage" and "environment" in option A have been taken directly from the passage in the text. Also alternative C uses the word "bird flu" which is also part of the text but very easy to eliminate as correct option.

In terms of formulation none of the items in MC task 4 use negative or double negative formulations. With regard to characteristic f it is obvious that only two items have been designed carefully without implying any answer clues. But items

4C, 4E and 4F make use of response clues in the form of length clues that might indicate the correct answer. As an example of length clues in task 4, a difference in answer length is presented in item 4C:

> Why did she find her US experience exciting?
>
> **A** Because she got to discuss US politics with real Americans.
>
> **B** Because things were so cheap there.
>
> **C** Because she liked being in an environment that she didn't understand 100%.
>
> **D** Because the people were genuinely friendly and ready to help.

Answer option C, which is the correct solution, is noticeably longer compared to the other alternatives and thereby offers a hint to the test takers. Besides, also item 4B includes a distractor that is fairly long compared to the other alternatives, but the longest alternative is not the correct option in item 4C. Therefore it is assumed that distractor D is supposed to irritate test takers by its length.

As far as the position of the correct answer is concerned, in task 4 the key is presented in every answer option. The correct answers are distributed as follows. As we can see options A and B are only used once as the position of the key whereas options C and D are used as the correct answer twice. Therefore it can be said that the key has been positioned randomly and makes use of all four positions.

Item 4A	C	Item 4D	D
Item 4B	B	Item 4E	D
Item 4C	C	Item 4F	A

5.2.3.3. Multiple Choice Task 5

The third task displays a more difficult task because a conversation between father and daughter is part of the reading passage. It is entitled "A short history of tractors in Ukrainian". The reason for the higher level of difficulty is that the protagonists of the story express their feelings, emotions and opinions, which can be more difficult to identify and interpret than factual information. An overview is provided by table 5.

Characteristics	Item 5A	Item 5B	Item 5C	Item 5D	Item 5E	Item 5F
a. Are the distractors plausible?	Y	Y	Y	Y	Y	Y
b. Are there convergent options?	N	N	N	N	N	N
c. Is most of the information part of the stem?	Y	Y	Y	Y	Y	Y
d. Does the item use the same wording as the text?	N	Y	N	N	N	Y
e. Are there negative or double negative formulations?	Y	N	N	N	N	N
f. Is the length of the alternatives approximately equal?	Y	Y	Y	Y	Y	Y
g. Is the correct answer positioned randomly?	N					

Table 5: Overview MC Task 5.

In the case of multiple choice task 5 the most interesting part is described by characteristic d. Items 5B and 5F use the same words that also occur in the passage itself. But other than usually the words that are repeated in the item only appear in the incorrect options. Thus these distractors clearly want to make the test takers to read carefully in order to find the correct answer.

Which of these things do we learn about Valentina in the last paragraph?

A She has a relative in the UK.

B She drives a Lada.

C she speaks English really well.

D She wants to work as a cleaner.

The corresponding passage from the text is quoted in order to understand the following discussion:

> She has an ==uncle in Selby==, and has come to visit him on a tourist visa. She wants to make a new life for herself and her son in the West, a good life, with good job, good money, nice car – absolutely no ==Lada== no Skoda – good education for son – must be Oxford Cambridge, nothing less. She is an educated woman, by the way. Has a diploma in pharmacy. She will easily find well-paid work here, once she learns ==English==. In the meantime, he is helping her with her ==English==, and she is ==cleaning== the house and looking after him. (excerpt Task 5)

In item 5F the words "Lada", "English" and "cleaner" (marked yellow) come directly from the last paragraph of the text, but they do not imply the correct answer. In order to be able to solve the item the test taker has to understand that "Selby" (marked green) is a town in the UK. But also without the background knowledge about "Selby" being a town in the UK, option A is still an attractive option because the other options can be eliminated by reading the text carefully.

Item 5A includes a negative formulation that is, like it was the case in MC task 3, emphasised by underlining the negative word "not".

> Which of the following is <u>not</u> true about Nadezhda's father's new girlfriend?
> **A** She is almost fifty years younger than him.
> **B** She has been married before.
> **C** She has a child.
> **D** She is not very attractive.

Again the key to the questions does not occur in every position. The answers appear in the following order. Options A, B and D are used twice but the correct answer never occurs in position C.

Item 5A	D	Item 5D	A
Item 5B	B	Item 5E	B
Item 5C	D	Item 5F	A

5.2.3.4. Multiple Choice Task 6

The sixth multiple choice task that is analysed is entitled "The price of the Olympics" and contains again six items based on the reading passage with four answer choices. Table 6 presents an overview of the characteristics of MC task 6.

Characteristics	Item 6A	Item 6B	Item 6C	Item 6D	Item 6E	Item 6F
a. Are the distractors plausible?	Y	Y	Y	Y	Y	Y
b. Are there convergent options?	N	N	N	N	N	N
c. Is most of the information part of the stem?	Y	Y	Y	Y	Y	Y
d. Does the item use the same wording as the text?	Y	Y	N	Y	N	N
e. Are there negative or double negative formulations?	N	N	N	N	N	N
f. Is the length of the alternatives approximately equal?	Y	Y	N	Y	Y	Y
g. Is the correct answer positioned randomly?	N					

Table 6: Overview MC Task 6.

Looking at characteristic d it is clear that items 6B and 6D use language from the text. An interesting aspect in item 6D is that only the distractors use the word "views" (marked yellow) that occurs in the reading passage several times. Merely the key does not include language from the text.

How are the developers able to keep offering apartments with great views?
A They have bought up all the waterfront where the views are the best.
B They keep the buildings low so everyone can enjoy the views.
C They build taller and taller buildings.
D They buy up the houses with the good views and redevelop them.

These distractors A, B and D have the potential of irritating the test takers because they could either assume that all distractors are concerned with the view and therefore option C has to be correct, or they could assume that one of the answer that include the word "views" has to be the key.

Another aspect that is worth noticing is that the positioning of the key is not balanced. Like in tasks 3 and 5, there is no balance in the distribution of the key

because option A is never the position of the correct answer whereas option C is used as the key position three times.

Item 6A	D	Item 6D	C
Item 6B	C	Item 6E	C
Item 6C	D	Item 6F	B

5.2.4. Evaluation of the Multiple Choice Tasks

Overall, the multiple choice tasks could be improved in various areas. As far as the examples from the New Matura are concerned, only a few improvements should be considered. Three items of task 1 include length clues that should be corrected in order to avoid guessing. Furthermore in task 2 the distribution of the correct answer should be revised in order to achieve a balance in the arrangement of the answer key. Moreover the convergent options in item 2B should be rewritten in order to avoid the elimination of the remaining two options.

Concerning the textbook examples, task 3 lacks efficiency in the formulation of items 3B and 3C. They could be formulated as an incomplete statement rather than a direct question. By reformulating these items repetition can be avoided and the test taker would need less time to read all alternatives. The same holds true for multiple choice task 4, where items 4C and 4F should be rephrased in order to include most of the information in the item stem and not in the alternatives.

Concerning the length of clues that are given in some items, it would be advisable to adapt those answer alternatives. This is the case in task 3 where items 3B and 3E indicate the correct answer by sentence length and also task 4 where items 4C, 4E and 4F include length hints that might encourage test takers to choose the most elaborate answer.

As far as convergent answers are concerned, only item 3A of task 3 could be modified by making all of the distractors treat the topic of eating or by changing option A to a different subject area in order to avoid convergent options.

Regarding the positioning of the key only task 4 presents the correct solution every position. But answer D is used twice in questions that directly follow each other,

which could be changed in the task in order to achieve a balanced distribution of correct answers. Tasks 3, 5 and 6 do not balance their positioning of the correct solution. The key is positioned randomly but does not appear in every possible position.

5.2.5. True/False/Not Given Tasks

As a special type of multiple choice question, true/false/not given tasks are used to identify the correctness of statements of fact or definitions. Therefore test constructors have to pay attention to various factors of those alternative-response format tasks.

The instructions for the true/false/not given items is formulated the same way for every item from the textbook as well as from the New Matura, and it appears before the actual task starts. Here is an illustration of the instructions provided.

> **Read the text below, then decide whether the statements (24-30) are true (T), false (F) or not given (NG) in the text. Put a T in the correct box.**

In the case of true/false/not given exercises the grid used for the analysis of the multiple choice items has to be adjusted in order to be able to capture those special features of true/false/not given items. The grid that follows includes an overview of the criteria that will analysed in the true/false/not given items.

Characteristics	Item 7A	Item 7B	Item 7C	Item 7D	Item 7E	Item 7F
a. Is the question based on a misunderstanding?						
b. Does the item use the same wording as the text?						
c. Are there negative or double negative formulations?						
d. Is outside information required?						

e. Is the key positioned randomly?						

The characteristics that have been added to the grid investigate if the question asked is based on a possible misunderstanding in the text or if the test takers require information that is not included in the text. Furthermore characteristics that have already been introduced will be analysed such as wording, negative formulation or positioning of the key.

The following table should serve as a guideline in order to have a better overview of the true/false/not given tasks that will be analysed. Task 7 has been taken from the bifie test booklet whereas tasks 8 and 9 come from "Make Your Way Ahead 8".

Task 7	Task 8	Task 9
"I Skied Down Mt. Everest"	"Golfer in gambling debt shock"	"Touch the top of the world".
includes	includes	includes
Item 7A-7J	Item 8A-8J	Item 9A-9J

5.2.6. True/False/Not Given Tasks – New Matura

5.2.6.1. True/False/Not Given Task 7

As far as the examples from the New Matura are concerned, they are very similar to those taken from the textbook. Therefore one example should serve as a representation for this category. The item is entitled "I Skied Down Mt. Everest" and it includes seven items that have to be answered with the help of the reading passage. Table 7 provides an overview of T/F/NG task 7.

Characteristics	7A	7B	7C	7D	7E	7F	7G	7H	7I	7J
a. Is the question based on a misunderstanding?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
b. Does the item use the same wording as the text?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
c. Are there negative or double negative formulations?	N	N	N	N	N	N	N	N	N	N

d. Is outside information required?	N	N	N	N	N	N	N	N	N	N
e. Is the key positioned randomly?	Y									

Table 7: Overview T/F/NG Task 7.

Every item in T/F/NG task 7 uses the opportunity to challenge the test takers by using statements that refer to the text. The reading passage needs to be read carefully in order to answer correctly.

Regarding the formulation of the items in T/F/NG task 7 it is noticeable that every item includes words that come directly from the text. These items provide hints that can help the test taker to locate the information that is needed to solve the task. But similarly to the examples taken from the textbook it is necessary to examine the reading passage carefully

The weight of the ==oxygen packs== slowed the climbers down.	T	F	**NG**

If the test taker re-reads the relevant paragraph in the text the word "oxygen" and "oxygen pack" will be found but the answer to the question is still 'not given'.

[…] lack of ==oxygen== sucks the warmth right out of you. Carrying ==oxygen packs== to stay alive was new for me, but 37 days after arriving at Mt. Everest Base Camp, we reached the summit.

Furthermore, neither negative formulations are included in the items nor do the test takers need information that goes beyond the text to answer the questions.

The key is positioned randomly and appears in every available position. Similarly to the true/false/not given tasks from the textbook there is an equal number of true and false keys and only 2 items where the correct answer is 'not given'.

5.2.7. True/False/Not Given Tasks– Textbook Examples

5.2.7.1. True/False/Not Given Task 8

The second true/false/not given task from the textbook in use to be analysed is entitled "Golfer in gambling debt shock" and it includes 10 items concerning the reading passage. Table 8 shows an overview of T/F/NG task 8.

Characteristics	8A	8B	8C	8D	8E	8F	8G	8H	8I	8J
a. Is the question based on a misunderstanding?	Y	Y	N	Y	Y	Y	Y	Y	Y	Y
b. Does the item use the same wording as the text?	Y	Y	N	Y	Y	Y	N	N	Y	N
c. Are there negative or double negative formulations?	N	N	N	N	N	N	N	N	N	N
d. Is outside information required?	N	N	N	N	N	N	N	N	N	N
e. Is the key positioned randomly?	Y									

Table 8: Overview T/F/NG Task 8.

As far as characteristic a is concerned, it is visible that only item 8C does not refer to a problem or statement presented in the reading passage. Because the correct answer to item 8C is "not given" it is very easy for the test taker to answer this question because the content of the item is not mentioned in the text. This does not hold true for all items where the correct answer is "not given", because very often the information needed to answer the question is formulated in a different way in order to encourage careful reading. Except for item 8C, every other item uses the opportunity to challenge the test taker by referring to events, comments or facts from the text that need to be read carefully in order to answer correctly.

Regarding the formulation of the questions it is noticeable that six out of ten items use words that come directly from the text. These items provide hints that can help the test taker to locate the required information in the reading passage. But even if these clues are included in the question itself the test taker still needs to read carefully in order to answer the question correctly.

As an illustration item 8B is compared to the passage where the corresponding information can be found.

| His addiction is the first thing he discusses in his new book. | T | **F** | NG |

Daly discussed his addiction to gambling in the final chapter of his autobiography, "John Daly: My Life In and Out of the Rough," to be released next Monday.

None of the items in this T/F/NG task makes use of negative or double negative formulations. Moreover, the test takers do not need any information outside of the reading passage to answer the questions. Also the distribution of the key was done in an optimal way. There is an equal number of true and false keys and only 2 items where the correct answer is 'not given'. The key is positioned randomly.

5.2.7.2. True/False/Not Given Task 9

The third true/false/not given task from "Make Your Way Ahead 8" is called "Touch the top of the world". Equal to the previously analysed true/false/not given task it includes as reading passage as well as ten corresponding items. Table 9 gives a brief overview of T/F/NG task 9.

Characteristics	9A	9B	9C	9D	9E	9F	9G	9H	9I	9J
a. Is the question based on a misunderstanding?	Y	Y	Y	Y	Y	N	Y	Y	Y	Y
b. Does the item use the same wording as the text?	Y	Y	Y	Y	Y	N	Y	Y	Y	N
c. Are there negative or double negative formulations?	N	N	N	N	N	N	N	N	N	N
d. Is outside information required?	N	N	N	N	N	N	N	N	N	N
e. Is the key positioned randomly?	Y									

Table 9: Overview T/F/NG Task 9.

Regarding characteristic a, only item 9F differs from the other items.

He used hot water to bring the feeling back to his hands.	T	**F**	NG

It presents a statement that could be considered as being related to the text but it is not based on a possible misunderstanding because the author of the text does not explain in detail what happens to the hand after putting it back in the glove.

I stuffed my lifeless hand back inside the glove and beat it against my knee.

He could either be using cold or hot water but the information is not given in the text. Deducing from the text passage I would rather disagree with the authors of this item because the correct answer should be 'not given' instead of 'false'.

Concerning characteristic b, it is obvious that most of the items include words that are taken over from the reading passage. Only items 9F and 9J do not use words directly related to the text. There are no negative or double negative formulations and there is no need for any information outside of the text itself.

The positioning of the key is similar to the second true/false/not given item that has been analysed. The key was positioned randomly in each option an equal number of times.

5.2.8. Matching Tasks

Concerning the matching exercises in the textbook and the New Matura, they cannot be analysed in the same way that the multiple choice and true/false/not given items have been described. They are not comparable to each other in a similar way because there are several types of matching exercises that are included in both publications. Therefore the variations of the matching exercises will briefly be described in order to see what different tasks can be part of the learning and testing of multiple matching exercises.

5.2.9. Matching Tasks– Textbook Examples

5.2.9.1. Matching Task 10

In the case of the textbook under review there are basically three types of matching exercises included in the book. Firstly, a common type of activity is included where

the students have to match two columns against each other after having been introduced to a certain topic. Matching task 10 demonstrates how these exercises may be constructed.

These days there are many different types of family. Here are some. **Match these with the definitions.**			
1	nuclear family	**a**	Used to talk about all the relatives in a family, including grand- parents, aunts, uncles etc.
2	extended family	**b**	A family with only a mother or a father caring for the children.
3	single parent family	**c**	The family created when two divorced people remarry and bring their children with them.
4	blended family	**d**	A family which offers temporary care for parentless children, until a permanent family can be found for them.
5	adoptive family	**e**	The 'traditional' family consisting of the mother, father and children.
6	foster family	**f**	A family in which the children are not biologically related to their parents.

Figure 3: "Make Your Way Ahead 8", Matching Exercise, p. 90.

Noticeably, on the left side there is column A, which presents a term that has to be matched against the corresponding definition in column B. Against what has been described in section 4.5. there is an equal amount of definitions and terms.

5.2.9.2. Matching Task 11

Secondly, a large amount of matching exercises in the textbook focus on the matching of extracts from the text to the equivalent passages where they have been taken from. Matching task 11 shows an excerpt of such a matching task.

An invisible army of workers strikes for the right to be seen   **US wants it both ways – closed borders and cheap labour**
The best place to witness America's immigration dilemmas is not the US-Mexican border. Go instead to any suburban school and see the Latino nanny, an illegal immigrant, picking up the children while their middle-class parents are still at work. Or stand outside any office block in the evening and watch the Mexican janitors go in to scrub out the toilets as the suited workers go home.

1 _____

America's reliance on illeagal immigrant labour – to keep the economy buzzing, prices low and the standard of living higher than anywhere else in the wolrd – is the worst.kept secret in the US economy, and everyone knows it.

2 _____

Publicly, most people deplore it. Privately, everyone does it – often while simultaneously demanding tighter border controls and greater restrictions on immigration. At best, the general American attitude towards employing illegal immigrants involves a strange form of double-think. At worst, it is hypocritical.

3 _____

Figure 4: "Make Your Way Ahead 8", "An invisible army of workers strikes for the right to be seen", p. 144.

The learners are required to choose which of the extracts A-F fit into the numbered gaps in the article. There is always one extra paragraph that should distract the students because it does not fit in any of the gaps.

A	American conservatives now want the millions of illegal immigrants in the US classified as felons and a fence built quickly along the Mexican border.
B	Since 1986, it has been illegal to employ workers without the necessary work permit and residency papers, yet the employment of illegal labour is endemic. Everyone knows about it.
C	On the other hand, no media attention is ever directed at the plights of the illegal immigrants themselves; the long hours of work, the appealing conditions many face and the constant fear they live in that at any moment they may be found out and deported.

Figure 5: "Make Your Way Ahead 8", "An invisible army of workers strikes for the right to be seen", p. 145.

5.2.9.3. Matching Task 12

The third type of matching exercise included in the textbook requires the students to identify which of the sections included information that is listed before or after the text. Matching task 12 is called "UFO study finds no sign of aliens" and demonstrates how theses exercises have been constructed.

Which section mentions the following?	
the length of time it took to produce the report.	1.....
why some people will not believe the findings of the report.	2.....
the atmospheric conditions that might look like UFOs.	3.....
the fact that there are people who are convinced they've had alien contact.	4.....
the lack of proof of the existence of objects that could crash and damage Earth.	5.....
the reasons why the report has been made public.	6.....
the anonymity of the report's author.	7.....
the medical reasons why some people believe to have been abducted.	8.....
some of the airborne object that might get confused for UFOs.	9.....
the reasons why the government does not plan any further investigations.	10....

A

A confidential Ministry of Defence report on Unidentified Flying Objects has concluded that there is no proof of alien life forms. In spite of the secrecy surrounding the UFO study, it seems citizens of planet Earth have little to worry about.
The report, which was completed in 2000 and stamped "Secret: UK Eyes Only", has been made public for the first time. Only a small number of copies were produced and the identity of the man who wrote it has been protected. His findings were only made public thanks to the Freedom of Information Act, after a request by Sheffield Hallam University academic Dr David Clarke.

B

The four-year study – entitled "Unidentified Aerial Phenomena in the UK" – tackles the long-running question of UFO-spotters: "Is anyone out there?" The answer, it seems, is "no". The 400-page report puts it like this: "No evidence exists to suggest that the phenomena seen are hostile or under any type of control, other than that of natural physical forces." It adds: "There is no evidence that 'solid' objects exist which could cause a collision hazard." So if there are no such things as little green men in spaceships or flying saucers, why have so many people reported seeing them? Well, here is the science bit.

Figure 6: "Make Your Way Ahead 8", "UFO study finds no sign of aliens", p. 22-23.

The students have to answer ten questions by choosing from sections A-D. Following the instructions, some choices can be used more than once. The questions asked require the students to read the passages very carefully in order to locate the matching information.

5.2.10. Matching Tasks – New Matura

In the New Matura only one type of matching exercises is included, which is very different from the matching tasks in the textbook that has been analysed. The type of matching used in the New Matura can however be compared to the matching

exercises from the book where extracts from the text need to be matched against the equivalent passages.

The matching tasks in the New Matura require the test takers to read a text passage where parts have been removed. Afterwards students have to choose the correct part for each gap. In every matching exercise appearing in the New Matura there are more parts that could be matched than there are gaps to be filled. Concerning the number of gaps to be filled, there is a variation depending on the text passage. The number of questions ranges from six to ten questions or gaps to be filled. Similar to the true/false/not given and multiple choice tasks, there is always one example provided that has already been solved.

5.2.10.1. Matching Task 13

Matching task 13 shows the design and construction of these tasks with an example taken from the bifie test booklet (p. 4-5).

All the Wolrd's a Stage

When he saw Jamie's name on the display, Sid pressed the answer button.
"Have I got news for you!" Jamie sounded excited.
"What's up?"
"You know how Mr Crane said)0) _____ for our next production? Well, I know
     who's getting the lead part."

Sid's heart stood still. Mr Crane was the drama teacher at Darton High and director
     of the Drama Club of which (Q7) _____. Both were intending to apply for
     Drama College when they left school and a main part in King Lear or
     Hamlet would be a dream come true.

"How about starting at the beginning and telling me the whole story." Sid said
     slowly. Jamie could get rather carried away at times.
"Right! Well, (Q8) _____ after the rehearsal – this was after you'd left – when I
     heard him on the phone…"
"Who?" Sid interrupted.
"Mr Crane, of course. He was waiting for me to finish so he could lock up."

A	he so much wanted it for himself
B	he was sure to be accepted
C	we were going to do Shakespeare
D	Mr Crane wanted a word with Sid and Jamie
E	Jamie and Sid were the two most enthusiastic members
F	Shakespeare wrote all the best plays.
G	Things were not turning out
H	Macbeth was the best name for him
I	he can definitely handle Macbeth
J	I was putting the stuff away

Figure 7: Schriftliche Reifeprüfung aus Englisch (7 Mai 2012.), "All the Wolrd's a Stage", p. 4-5.

5.2.10.2. Matching Task 14

The second type of matching exercise that occurs in the New Matura is a variation of the previously described task. The test takers are confronted with an interview where the eliminated parts are the questions the interviewer asks. Matching task 14 illustrates the design of such items by showing an extract of the reading passage and the corresponding questions.

Interview with Bode Miller

0 _____

The team cut my funding, [so] I was going to be paying for
myself anyway. There were other issues. [For example,] I
thought it would be productive to have a bus with a gym in
it and our own chefs so that we can control our food. Not that it's ever been an issue, [but]
somebody could put a little of any banned substance in our mass pastas in our hotel, and
we'd all be disqualified, no questions asked.

Q15 _____

I do. The sport is really individual when you're competing, but outside of the competition you
definitely rely on your team-mates for support. But it's not gone. We're all still friends. I see
those guys all the time.

Q16 _____

Yeah. It perpetuates a positive energy. I think guys generally race at their highest level when
they're feeling positive about themselves rather than trying to bring other guys around them
down to a lower level to beat them.

A	What are your ultimate goals for your career and your private life?
B	Do you ever think about giving up skiing?
C	Do you miss the camaraderie?
D	Have you always wanted to become a skier?
E	Why did you separate from the U.S. ski team?
F	Are you still the crazy guy you used to be?
G	How would you describe your racing spirit?
H	You had some disappointing races in the past few weeks. Will your luck change?
I	What ist he most difficult choice you have ever had to make?
J	Are you on good terms with the other racers?
K	What will you do when you are no longer competing?
L	How do you deal with accidents?

Figure 8: Schriftliche Reifeprüfung aus Englisch (3. Mai 2010.) "Interview with
Bode Miller", p. 6-7.

5.2.11. Comparison

The multiple choice tasks used in the New Matura include more items and therefore more questions on one reading passage than the tasks taken from the schoolbook. Regarding the number of items, the tasks from the textbook always use six items that have to be answered per task, whereas the examples from the New Matura vary between six and eight items per reading task. Concerning the length of the individual reading passages there is no noticeable difference. The texts from the schoolbook range between 528 - 645 words in each text. Similarly the items taken from the New Matura include texts that have between 600 - 623 words. As it was mentioned above, the instructions take the same form in each of the tasks, with the only difference that the New Matura includes one extra item that has already been answered in order to demonstrate how the item should be solved.

All the distractors included in the items that have been analysed are plausible. They manage to distract the test takers by using information from the text so that the corresponding passages have to be read carefully.

Concerning the formulation of the distractors and the keys it is noticeable that none of the examples taken from the New Matura call for reformulation. They have been constructed in an efficient way, as most of the information is part of the stem and not of the alternatives. However, some of the questions in connection with texts from the schoolbook need to be revised in order to make the items more efficient. This also applies for negative formulations, which only occur in items from the textbook. Furthermore the textbook examples include more length clues than the samples from the New Matura.

In connection with the distribution of the correct answer both sources of items need to be careful when it comes to arrangement of the answers. Some items use all the positions for the key to a question, whereas others neglect some positions completely. In order to construct a good item it is necessary to consider all positions for the key in a random order.

Regarding true/false/not given items the most obvious dissimilarity between the items used in the New Matura and the textbook is that in the Matura examples the first question has already been answered. This is done in order to show the test takers the way they should solve the other questions. This is not done in any of the textbook examples that have been analysed. Another divergence can be found in

the number of items that is used in each task. In the textbook examples students have to answer 10 items on a reading passage, whereas the Matura tasks only present 8 items of which only 7 have to be answered by the test takers themselves.

Besides these differences the true/false/not given tasks presented in both cases are very similar to each other. Both sources include well-constructed items that use texts of approximately the same length. A majority of the items include words or phrases that have been taken directly from the text. Most of the items under consideration are based on a misunderstanding of the text and they do not ask questions that would need information that goes beyond the reading passage. Concerning the positioning of the key it is evident that all true/false/not given tasks try to position the correct answer in a random order.

As a concluding remark on true/false/not given items it can be said that the items under consideration can be compared to each other in terms of the features that have been analysed. They have the same way of construction and present the questions in a similar manner.

Finally, the matching items shall be compared to each other. It became obvious that in this case the biggest contrast could be identified. The textbook provides a great number of practising materials that include a variation of three different types of activities. Unfortunately those variations of matching exercises are in a large part not included in the New Matura. Only one type of matching exercise from the book can be compared to the examples used in the New Matura. In the Matura examples that have been analysed the test takers had to choose the correct part, question or headline to fill a gap. Similarly, in some textbook examples the students were asked to fill gaps with extract eliminated from the text passage.

Overall, the book offers a good preparation for most of the exercises that have to be solved in the New Matura. Especially the multiple choice tasks can be trained by the use of a textbook like "Make Your Way Ahead" or other comparable books on the market. The tasks are presented in a similar way, which makes it easy for test takers to get used to the visual presentation and the formulation of such tasks. Furthermore, frequent exposure to true/false/not given items as they appear in the textbook can help to prepare test takers for the New Matura, as the exercises have been constructed very similar. Unfortunately, the matching tasks that were part of the textbook cannot be seen as an adequate preparation for the New Matura. There

is a wide range of multiple matching exercises and therefore the schoolbooks should be revised in order to provide a wide range of matching exercises. By that the test takers get used to this testing format and may be able to deal with new formats in an effective way.

5.3.   What could be improved in objective testing formats?

As with every method of testing it should be ensured that objective testing items are only used if they are the best option for assessing a given objective. Clearly, only well-written selected-response items should be used therefore it is necessary to control, pilot and improve given items. Each item should be concerned with one specific problem hat is stated clearly in the stem. In the items that have been analysed, there was a great number that included various response clues. It should be considered to revise those items in order to guarantee that the test taker is not influenced by the formulation in any way. This includes convergent answers as well as length clues.

Furthermore a great number of tasks did not present the answer or key to a question in each of the alternative positions randomly. Sometimes positions were not even used once, which should be avoided.

Concerning the topic of the reading passage there is a great potential of improving the multiple choice items. The topic areas of reading tests that use multiple choice formats should be chosen carefully because, as it has been shown in section 3.8., interests in specific areas can affect male and female test takers differently. Discrimination should be avoided to ensure fair test conditions for both sexes. This criticism does not apply for the category of text type or genre because a broad variety has been used in the tasks included in the analysis. "Good tests of reading and good assessment procedures in general will ensure that readers have been assessed for their ability to understand texts in a range of different topics" (Alderson 2000: 63). Furthermore Alderson states that texts from popular fiction would be more suitable for reading tests than texts from the field of non-fiction because the former genre is expected to be more balanced in terms of difficulty (ibid). In the tasks included in the textbook it is noticeable that the authors tried to provide a variety of topics in order to attract both female and male readers.

As usual, every unit is devoted to a different topic such as "A star is born", "Lifestyles" and "The Emerald Isle". Within the units it is ensured that, for example, not only female stars are discussed to address both sexes equally. Regarding the examples in the New Matura booklets that I used for my analysis, they consisted of texts from different fields including sports, fashion, environment, education as well as relationships.

Overall there were five texts in all of the booklets, a minimum of one text in each booklet, that were dealing with sportive topics, which may be rather uninteresting to some female test takers. In order to achieve a balance in terms of content of the text it would be advisable to use more neutral topics that would interest both, female and male test takers equally or at least try to balance female and male topics in order to avoid discrimination.

6. Conclusion

The significance of selected-response formats in testing is evident. Thus, with the introduction of the New Matura in Austria the importance is even rising. It is essential for students in the Austrian school system to be familiar with these forms of testing. Therefore this thesis should serve as a guideline that demonstrates effective and successful ways of using those testing formats. It is necessary for test developers, teachers, students and everyone else who is involved in the testing process to be aware of certain factors that evolve around language testing. Clearly there has been a lot of criticism evolving around the topic of multiple choice testing but it is inevitable to become acquainted with the advantages this testing format holds for a test taker as well as for a test constructor. An obvious advantage is the versatility of objective testing formats. They are suitable in a great number of subject areas and can be used to measure a variety of educational objectives. Besides they are adjustable to various learning outcomes from simple to more complex levels such as the student's ability to apply principles to new situations, comprehend concepts and interpret cause-effect relationships. Moreover they are very quickly to answer compared to a task presented in the form of an essay. Therefore a test taker can answer a large number of questions in a relatively short time. Thus, validity can be granted because the outcomes of the test will be representative of the student's achievement. The time issue also comes into play when it comes to the scoring of the test. Because there is only one correct answer the correction of the test will not take as much time as if the test taker would have to produce language instead of selecting it. Based on this objective way of scoring multiple choice tasks are considered as being very reliable and can ensure fair grading.

Undoubtedly, there are also limitations to the usage of objective testing tasks. Versatility cannot only be seen positively. Because a test taker has to select his or her answer from a given set of alternatives, multiple choice tasks for example cannot measure various learner abilities such as the organisation of personal thoughts or the production of original ideas. Therefore it is important to distinguish between objectives, that can be appropriately assessed by using multiple choice items and objectives, which would be better measured by some other test format. Furthermore it is worth mentioning that, despite the common belief, multiple

choice items are not easy to construct. Good items are generally very time-consuming to construct because there are so many aspects to consider. These aspects include the invention of plausible distractors, the formulation of the item itself or the avoidance of unintentional clues.

When it comes to true/false/not given tasks as a variation of multiple choice tasks, it is difficult to eliminate the guessing factor, which can reduce the reliability of a reading test. As a solution it has been suggested to present a large number of items in order to reduce the guessing factor.

It has been shown that there are numerous ways of preparing the students for the New Matura, which should be easy to integrate into the language classroom, due to the fact that there are various publications that have been adapted for the purpose of the standardised Matura. Even if the practising material that is available in the schoolbook is not always comparable to the samples that appear in the written Matura these items serve as a good basis for familiarising the learners with the new testing formats.

Over and above that this thesis touched upon the question what different factors can affect the performance of the test takers in a reading test that includes different varieties of multiple choice items. It has been demonstrated, that not only text variables like topic, content, vocabulary and genre can influence the performance on a test but also various reader variables affect the score of a test. Therefore, awareness of metalinguistic knowledge like reading strategies but also stable characteristics like sex, age, personality and interest in the topic should be considered in test development.

If test developers evaluate the existing items critically and improve them according to the various principles that characterise a well-written task, future generations will be versed in the handling of selected-response items. Multiple choice tasks and the variations of them should be appreciated for being an effective method to test a great variety of abilities in a short amount of time.

7. References

Alderson, J. Charles. 2000. Assessing reading. Cambridge: Cambridge University Press.

Alderson, J. Charles. 2006. Diagnosing foreign language proficiency: the interface between learning and assessment. London: Continuum.

Alderson, J. Charles; Beretta, A. (eds.). 1992. Evaluating second language education. Cambridge: Cambridge Univ. Press.

Alderson, J. Charles; Clapham, Caroline; Wall, Dianne. 1995. Language Test Construction and Evaluation. Cambridge: Cambridge University Press.

Alderson, J. Charles; Urquhart, Alexander H. 1984. Reading in a foreign language. London: Longman.

Alatis, James. 1989. Language teaching, testing and technology. Lessons from the Past with a View Toward the Future. Washington D.C.: Georgetown University Press.

ALTE. 1998. Multilingual Glossary of Language Testing Terms. Cambridge: Cambridge Univ. Press.

Bachman, Lyle F. 2004. Statistical Analyses for Language Assessment. Cambridge: Cambridge University Press.

Bachman, Lyle F.; Palmer, Adrian S. 1996. Language testing in practice. Oxford: Oxford University Press.

Bachman, Lyle. F. 1990. Fundamental considerations in language testing. Oxford: Oxford Univ. Press.

Baker, David. 1989. Language Testing: A Critical Survey and a Practical Guide. London: Edward Arnold.

Bright, J.A., and McGregor, G.P. 1970. Teaching English as a second language. London: Longman.

Brown, James Dean. 1988. Understanding research in second language learning: a teacher's guide to statistics and research design. Cambridge: Cambridge University Press.

Brown, James Dean; Hudson, Thom. 2002. Criterion-referenced language testing. Cambridge: Cambridge University Press.

Bygate, M. (ed.). 2001. Researching pedagogic tasks: second language learning, teaching and testing. Harlow: Longman.

Cohen, Andrew, D. 1998. Strategies in Learning and Using a Second Language. London: Longman.

Cohen, Andrew, D. 1994. Assessing Language Ability in the Classroom. (2nd edition). New York: Heinle and Heinle.

Cohen, Andrew, D. 1980. Testing Language Ability in the Classroom. Massachusetts: Newbury House Publishers.

Cohen, Jacob. 1988. Statistical Power Analysis for the Behavioral Sciences. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Cumming, Alister; Berwick, Richard. (eds.). 1996. Validation in language testing. Clevedon: Multilingual Matters.

Davies, Alan. 1990. Principles of language testing. Oxford: Blackwell.

Davies, Alan; et al. 1999. Dictionary of language testing. Cambridge: Cambridge University Press.

Gassner, Otmar; Mewald, Claudia; Sigott, Guenther. 2007. Testing Reading. Specifications for the E8-Standard Reading Tests. LTC Technical Report 2. Klagenfurt: Language Testing Centre.

Gronlund, Norman, E. 1965. Measurement and Evaluation in Teaching. New York: The Macmillan Company.

Gronlund, Norman, E. 2003. Assessment and Student Achievement. (7th edition). Boston: AB Longman.

Halper, Diane. 2000. Sex Differences in Cognitive Abilities. (3rd edition). Mahwah: Erlbaum.

Harris, M.; McCann, P. 1994. Assessment. Oxford: Heinemann.

Heaton, John, B. 1988. Writing English Language Tests. New York: Longman.

Hedge, Tricia. 2000. Teaching and learning in the language classroom. Oxford: Oxford University Press.

Henning, Grant. 1987. A Guide to Language Testing. Cambridge, Massachusetts: Newbury House Publishers.

Horn, Manfred; Vollandt, Rüdiger. 1995. Multiple Tests und Auswahlverfahren. Stuttgart: Gustav Fischer Verlag.

Hughes, Arthur. 2003. Testing for Language Teachers. Cambridge: Cambridge University Press.

Kitao, Kathleen S.; Kitao, Kenji. 1999. Essentials of English Language Testing. Tokyo: Eichosha Co Ltd.

Lado, Robert. 1961. Language Testing. London: Longman.

Lado, Robert. 1964. Language Teaching. A Scientific Approach. New York: McGraw-Hill Inc.

Lynch, Brian, K. 2003. Language assessment and programme evaluation. Edinburgh: Edinburgh Univ. Press.

McNamara, Tim. 2000. Language Testing. Oxford: Oxford University Press.

McNamara, Tim; Roever, Carsten. 2006. Language testing: the social dimension. Malden, Mass: Blackwell.

Oller, John W.; Perkins, Kyle. 1980. Research in Language Testing. Massachusetts: Newbury House Publishers.

Oosterhof, Albert. 2009. Developing and Using Classroom Assessment. Upper Saddle River, New Jersey: Merrill.

Popham, W. James. 2002. Classroom Assessment: What teachers need to know. (3rd edition). Boston: Allyn & Bacon.

Urquhart, Alexander H.; Weir, Cyril J. 1998. Reading in a second language: process, product and practice. Essex: Pearson Education Limited

Van Blerkom, Malcolm L. 2009. Measurement and Statistics for Teachers. New York and London: Routhledge.

INTERNET:

"SRP - Standardisierte Reifeprüfung"
http://www.uibk.ac.at/srp/aufgaben_englisch.html (10. June 2012).

"TOEFL testing"
 http://www.testpreppractice.net/TOEFL/toefl-testing-1.aspx.   (09.   February 2010).

8. Appendix

Schriftliche Reiferpüfung aus Englisch. Reading Test: Haupttermin 03. Mai 2010.

LESEN E HT1 2009/10

Name	Klasse
Punkte	von **30**

Schriftliche Reifeprüfung aus Englisch

1. Haupttermin
03. Mai 2010

Reading Test

Instructions

1. This test contains 4 tasks and 30 questions.

2. Write all your answers in this test booklet.

universität
innsbruck

© Bundesinstitut für Bildungsforschung, Innovation &
Entwicklung des österreichischen Schulwesens
1010 Wien, Schreyvogelgasse 2 / 5. Stock

Bundesinstitut
bifie

Bildungsforschung, Innovation & Entwicklung
des österreichischen Schulwesens

"Keeping It Up" p. 2-3.

Task 1 ___ / 8 P.

Read the text below, then choose the correct answer (A, B, C or D) for
questions 1-8. Put a ☒ in the correct box. The first one (0) has been done for
you.

Keeping It Up

Rory pushed his chair back into the shade and raised his glass to his lips. He watched Pam
crossing the square to the little shopping centre. She had seen a belt; she just had to have it,
and taking a handful of notes out of her purse had left him here at the café, with all her
shopping trophies and her bag to look after.

Tomorrow they were returning to Dublin, and for Rory it wasn't a day too soon. With his fair
colouring he was easy prey for the Mediterranean sun and he wondered, not for the first
time, why he had ever agreed to this holiday. Pam was now scanning a shop window,
elegantly poised on her high heels. Long days of careful tanning had turned her into a
bronzed goddess in her white mini-dress. A couple of local youths eyed her up
appreciatively. Rory chuckled: at one time he would have been consumed with jealousy; all
he felt now was amusement. If only they knew what putting up with Pam on a day-to-day
basis was like!

Pam had disappeared. Rory sat sipping his beer. Baking in the afternoon sun, the square
was almost deserted. There was only a handful of little boys, kicking a ball around. Rory
observed them affectionately. With a sigh he wondered what the guys would be doing at that
moment, probably enjoying a beer in some cosy pub, discussing the team's chances before
moving on to Ibrox Stadium. The trip to Glasgow for the Cup Final had been planned for
ages when Pam suddenly came up with the Mediterranean holiday idea. If he didn't pull out
of the Glasgow trip, Pam had warned, it would mean that football meant more to him than
she did. And apparently there was somebody called Jon, who would be delighted to
accompany her if Rory refused. And so here he was, burnt bright pink and miserable. Having
followed Pam around from the beach to the shops and back again all week, he couldn't say
right now whether football or Pam was higher on his list of favourite things.

Rory was startled by the little boys' ball bouncing off the café table, narrowly missing his
empty glass, and landing in one of Pam's carrier bags. The boys froze, perhaps expecting an
angry response. Rory got up, retrieved the ball and on an impulse tossed it in the air and
kept it up with his head for a length of time that even he found impressive. When the ball
eventually bounced to the ground, the little boys cheered and applauded wildly, chattering
excitedly and whistling for more.

A few repeat performances later, sticky with sweat, Rory laughingly called a halt. He was hot
and out of breath but he hadn't had this much fun all week. He was beginning to feel his old
self again. "Ice cream?" he asked the boys, pointing to the sign outside the café, and again a
cheer went up. But when Rory opened his wallet to pay, he discovered he had no change.
He glanced at Pam's bag. "You owe me it," he thought to himself, rummaging around for her
purse. What fell out of Pam's purse when he opened it was a surprise, but not really an
unpleasant one. It was a photograph of a guy, quite a decent sort, by the look of him. Rory
turned it over. On the back it said, "All my love, Jon." Grinning, Rory carefully replaced the
photo. "Good luck to you, Jon," he murmured.

Having paid for the ice cream, Rory gathered up Pam's belongings and headed back for the
hotel. He had a case to pack and he didn't want to miss the coverage of the match on TV. He
had kept up the pretence long enough.

0   Rory observed Pam

    A   from a park bench. ☐
    B   buying a new belt. ☐
    C   going into the shopping centre. ☐
    D   walking over to the shops. ☒

Q1   Pam

    A had told Rory to wait at the cafe. ☐
    B had hardly any money left. ☐
    C had not bought anything yet that day. ☐
    D knew what she was going to buy. ☐

Q2   Rory

    A wished they were going home a day earlier. ☐
    B wanted to spend his next holiday in Ireland. ☐
    C had not really enjoyed this holiday. ☐
    D had been been here on holiday with Pam before. ☐

Q3   Pam was

    A pleased that she was being watched. ☐
    B attracting attention. ☐
    C flirting with two young men. ☐
    D aware that she was being watched. ☐

Q4   Rory found it funny that

    A the young men had no idea what a difficult person Pam was. ☐
    B he knew Pam would want him to feel jealous and he didn't. ☐
    C people did not realize that Pam went shopping every day. ☐
    D the young men had no idea how wonderful it was living with Pam. ☐

Q5   Watching the little boys playing football made Rory

    A think of his team winning at Ibrox Stadium. ☐
    B feel like joining in. ☐
    C wish he was back home in Dublin. ☐
    D think of his friends. ☐

Q6   Rory had cancelled his trip to the Cup Final because

    A he preferred to go on holiday with Pam. ☐
    B Pam meant more to him than football did. ☐
    C he had to do so in order to please Pam. ☐
    D Jon had told him he wanted to go on holiday with Pam. ☐

Q7   Rory headed the little boys' ball in the air

    A and proved how skilful he was at football. ☐
    B although he was annoyed at first when the ball hit his table. ☐
    C because the boys wanted him to. ☐
    D although he was feeling so hot and angry. ☐

Q8   Rory returned to the hotel without Pam, because he

    A had planned to watch the match before she got back. ☐
    B had decided his relationship with her was over. ☐
    C had said he would get the luggage ready. ☐
    D was shocked at finding Jon's photo in her purse. ☐

3

102

"Interview with Bode Miller" p. 6-7.

Task 3   ___/ 9 P.

Read the text below, then choose the correct heading (A-L) for each paragraph (15-23). There are two extra headings that you should not use. Write your answers in the boxes provided at the end of the task. The first one (0) has been done for you.

Interview with Bode Miller

0 _____

The team cut my funding, [so] I was going to be paying for myself anyway. There were other issues. [For example,] I thought it would be productive to have a bus with a gym in it and our own chefs so that we can control our food. Not that it's ever been an issue, [but] somebody could put a little of any banned substance in our mass pastas in our hotel, and we'd all be disqualified, no questions asked.

Q15 _____

I do. The sport is really individual when you're competing, but outside of the competition you definitely rely on your team-mates for support. But it's not gone. We're all still friends. I see those guys all the time.

Q16 _____

Yeah. It perpetuates a positive energy. I think guys generally race at their highest level when they're feeling positive about themselves rather than trying to bring other guys around them down to a lower level to beat them.

Q17 _____

The results have been pretty poor this season. But when I'm at speed and not making mistakes, I'm much, much faster than the rest of the world right now.

Q18 _____

I'm always out to ski hard. If I get good results, that's ideal. But I feel I've been true to myself my entire career with my effort. My intensity is really second to none on the World Cup. The effort and intensity are the only things I can control. If other guys ski better, you don't get the results.

Q19 _____

It's a matter of perception. I can make everyone think that I'm not partying, or I could easily make people think the other side. In the past, it's been a matter of where the media have put the focus. This year with my team separating and other things, there's a lot of other stuff to focus on.

Q20 _____

Athletically, it's to not be hurt. I enjoy being outside, and eventually I'd like to have a family. I'd like to not be limping around when I'm 50 years old.

6

103

Q21 _____

I've been exploring different options for when I'm done skiing. I have the Turtle Ridge Foundation, which is helping a bunch of worthy causes around the Northeast. I've also started SkiSpace, which is an online social network that basically deals with all things based around any snow sport.

Q22 _____

I grew up snowboarding for a while. I went through a period at boarding school when my coaches wanted me to switch to snowboarding because they thought I was no good at skiing. I was too skinny. I had terrible technique. They were saying I should be a snowboarder, and luckily, I resisted.

Q23 _____

I've had a lot of them. I had crashes when I was small and Gumby-like that would have killed me now. I would just fly off jumps and go 40 or 50 meters when I was 6 years old—break skis, smash my goggles and get a bloody nose and go crawl inside for a little while and then come back out and ski more in the afternoon.

A	What are your ultimate goals for your career and your private life?
B	Do you ever think about giving up skiing?
C	Do you miss the camaraderie?
D	Have you always wanted to become a skier?
E	Why did you separate from the U.S. ski team?
F	Are you still the crazy guy you used to be?
G	How would you describe your racing spirit?
H	You had some disappointing races in the past few weeks. Will your luck change?
I	What is the most difficult choice you have ever had to make?
J	Are you on good terms with the other racers?
K	What will you do when you are no longer competing?
L	How do you deal with accidents?

0	Q15	Q16	Q17	Q18	Q19	Q20	Q21	Q22	Q23
E									

7

104

"I Skied Down Mt. Everest" p. 8-9.

Task 4   __ / 7 P.

Read the text below, then decide whether the statements (24-30) are true (T), false (F) or not given (NG) in the text. Put a ☒ in the correct box. The first one (0) has been done for you.

"I Skied Down Mt. Everest"

Kit Deslauriers, 37, as told to Emily Brower Auchard

I tried alpine skiing for the first time when I was 14 years old - I was immediately hooked. So the day I graduated from college, I moved to Telluride in Colorado to build a life centered on skiing. I completed my EMT training, worked on ski patrol, and was introduced to the thrill and beauty of backcountry skiing, where there are no chairlifts or lodges - just you and the mountain. Nine years ago, after I climbed a mountain in India and found myself wishing I had my skis with me, I decided to perfect my skills so that I could ski anything, anywhere. I met my husband, Rob, on a mountain in Siberia a year later, and together we've climbed and skied peaks all over the world.

Two years ago, we set a goal to ski the seven highest summits on the seven continents. Mt. Everest was the final and most challenging of them all. It took several days of traveling to Nepal and an 11-day trek just to reach base camp. We still had to climb the mountain itself, which has several different, complicated terrains. For example, the Khumbu Ice Fall is a maze of ice and deep crevasses that are crossed by ladder, so every step counts. More so than on other mountains I've climbed, on Everest we also had to contend with the effects of extreme altitude and cold temperatures; at high altitudes, lack of oxygen sucks the warmth right out of you. Carrying oxygen packs to stay alive was new for me, but 37 days after arriving at Mt. Everest Base Camp, we reached the summit.

Standing on the Southeast Ridge of Mt. Everest, looking over an 8,000-foot drop to Nepal on my right and a 9,000-foot drop to China on my left, I felt like I was on top of the world. Below us, the snow was hard-packed and windblown with firm ripples.

We planned to ski down 6,000 vertical feet, but our oxygen started running low, so I took off my skis to get us through a dangerous area more quickly. When the weather took a turn for the worse, we had to stop and camp for the night. We'd been out for 15 hours - to continue would have been suicide.

After a cold night with very little food and water, we faced the most extreme skiing of the trip. The Lhotse Face, a 45-degree slope, is much steeper than any commercial ski run. And that day, it was also a sheet of ice.

If one of us were to fall on this kind of terrain, we'd surely die. I was scared, but I couldn't let the fear take control. If I did, my balance would have been thrown off. To focus my attention, I developed a mantra on the spot that I repeated for each turn: "Like your life depends on it. Turn." Eight hours later, we were down and greeted by our support team with cheers. Everyone was rooting for us - and we did it.

a

Looking up at the mountain, I was in awe at what I had just accomplished. I was the first woman to descend Mt. Everest on skis - and I had completed my last of the seven summits!

Today, Rob and I are still finding mountains to climb up and ski down, and our life is rich with the experience. They're not all as challenging as Mt. Everest, but each is thrilling, and every trip makes my heart sing.

	Statements	T	F	NG
0	Kit became fascinated with alpine skiing the first time she tried it.	X		
Q24	Kit left college in order to be able to focus only on skiing.			
Q25	Kit decided to become a better skier when a mountain in India proved too difficult for her to ski down.			
Q26	Rob was the one who encouraged Kit to ski on challenging mountains.			
Q27	Before Mt. Everest, Kit had never had to deal with such difficult conditions.			
Q28	The weight of the oxygen packs slowed the climbers down.			
Q29	When they reached the Lhotse Face, conditions were no worse than usual.			
Q30	Kit was overwhelmed by what she had achieved when she looked back.			

LESEN E HT 2011/12

Name	Klasse

Schriftliche Reifeprüfung aus Englisch

Haupttermin
07. Mai 2012

Reading Test

Instructions

1. This test contains 4 tasks and 27 questions.

2. Write all your answers on the answer sheet.

19

Universität
Innsbruck

Bundesinstitut
bifie

"All the World's a Stage" p. 4-5.

Task 2   7 P.

Read the text about two young actors. Parts of the text have been removed.
Choose the correct part (A-J) for each gap (7-13). There are two extra parts that
you should not use. Write your answers in the spaces provided [on the answer
sheet]. The first one (0) has been done for you.

All the World's a Stage

When he saw Jamie's name on the display, Sid pressed the answer button.
"Have I got news for you!" Jamie sounded excited.
"What's up?"
"You know how Mr Crane said (0) ___ for our next production? Well, I know who's getting the
lead part."

Sid's heart stood still. Mr Crane was the drama teacher at Darton High and director of the
Drama Club of which (Q7) ___. Both were intending to apply for Drama College when they
left school and a main part in King Lear or Hamlet would be a dream come true.

"How about starting at the beginning and telling me the whole story," Sid said slowly. Jamie
could get rather carried away at times.
"Right! Well, (Q8) ___ after the rehearsal – this was after you'd left – when I heard him on the
phone..."
"Who?" Sid interrupted.
"Mr Crane, of course. He was waiting for me to finish so he could lock up."
"And..." Sid prompted.
"And I heard him say 'Macbeth'."
"You're kidding!" Sid breathed.
"No, honestly. I heard him say it with my own ears."
"Who was he speaking to on the phone?"
"Mr Chalmers, probably. He's helping to stage the next performance, isn't he?"
"And what about the lead part?" Sid said hesitantly; it had to be either him or Jamie, and
(Q9) ___.
"Well, after I heard 'Macbeth', I started listening in earnest, obviously. There was a pause
and then he said, 'I'm sure Sid could do it' and then 'We'll give him the lead. He's very
capable; (Q10) ___. I'll speak to him about it tomorrow'. So, I guess congratulations are in
order. Lucky thing! I'm hoping for Banquo myself. I could do a lot with that ghost scene."
Sid was overcome by a mixture of emotions. With Macbeth on his college application,
(Q11) ___. At the same time, he was touched by Jamie's reaction: it would have meant so
much to him, too, to get the lead in Macbeth. A murmured "Thanks, mate," was all Sid could
manage before hanging up.

4

108

Next morning Sid and Jamie could think of nothing but Macbeth. Eventually, during French, a junior came in to say that (Q12) ___ and could they go down to his room at break. Sid and Jamie exchanged meaningful looks. Neither heard a word of what Mlle Coindet said for the rest of the period.

When they entered Mr Crane's office, he was on the phone again.
"Yep, he's here now. I'll ask him right away."
Mr Crane took the receiver from his ear and laid it against his shoulder.
"Sid, got something big lined up for you. I feel you are the very man for the job." He grinned broadly. "We need somebody to walk our German Shepherd weekdays, and since you live nearby and have a dog of your own, we thought you might be interested in making some extra pocket money. How about it? I've got my wife on the phone right now - can you give us a quick answer?"
Sid shrugged and nodded.
"Good man. He's a bit of a handful, Macbeth, but we'll give you the lead and as long as he's on it, he's quite well behaved."

(Q13) ___ quite as Sid had expected. He glanced at Jamie, whose mouth had dropped open. Mr Crane said goodbye to his wife, hung up and turned to face the boys, smiling.
"Well, back to class, Sid," he said. "I'll give you my address later. Jamie, have you got a moment? I'd like to discuss the new production with you. Do you remember there was some talk of us having a go at Shakespeare?"

A	he so much wanted it for himself
B	he was sure to be accepted
C	we were going to do Shakespeare
D	Mr Crane wanted a word with Sid and Jamie
E	Jamie and Sid were the two most enthusiastic members
F	Shakespeare wrote all the best plays
G	Things were not turning out
H	Macbeth was the best name for him
I	he can definitely handle Macbeth
J	I was putting the stuff away

LESEN E HT2 2009/10

Name	Klasse
Punkte	
von **30**	

Schriftliche Reifeprüfung aus Englisch

2. Haupttermin
10. Mai 2010

Reading Test

Instructions

1. This test contains 4 tasks and 30 questions.

2. Write all your answers in this test booklet.

innsbruck

"The Warning" p. 2-3.

Task 1 ___ / 6 P.

Read the text below, then choose the correct answer (A, B, C or D) for questions 1-6.
Put a ☒ in the correct box. The first one (0) has been done for you.

The Warning

Tom shrugged off his white lab coat, hung it up on its peg and went over to the sink to wash his hands. It had been another long day. A glance at the clock told him it wasn't far short of nine. He ripped a paper towel out of the dispenser and dried his hands, studying his face in the mirror as he did so. Red-rimmed eyes with dark circles under them, and he wasn't even thirty yet. High-profile work took its toll. The money was a joke, the hours were criminal, but Tom knew that it was unusual for somebody his age to be trusted with a project of this magnitude. He had spent the day putting the final touches to the report that the local authorities were waiting for and he knew it was going to be a bombshell. TYP Cosmetics would not be pouring their toxic waste into the Darrie for much longer when people found out exactly what it was doing to the wildlife. He and his team were going to be the heroes of the day when the media got hold of it and promotion - and money – would surely follow.

Tom yawned and stretched and looked round his lab one last time to check that all was in order: apparatus switched off, no bottles uncorked, the test tubes standing in neat rows. The work surfaces were gleaming, polished to perfection by the cleaner, who had mopped and wiped around Tom while he was still busy a few hours ago. Tom liked the new cleaner. He wasn't exactly chatty but he did a grand job; he was much more conscientious than the last one. Even the floor was spotless. Tom's eyes came to rest on a scrap of paper lying just inside the door. "Oh, well," he thought aloud. "Nobody's perfect." He bent to pick up the piece of paper and was on the point of binning it, when he noticed there was writing on the other side: DONT DO IT OR ELSE. Frowning, Tom turned the paper this way and that, then jerking the door open, he looked up and down the dark corridor. It was deserted; there was no sign of life.

Tom sat down again at his desk and examined the writing. Seconds later he was smiling wryly and shaking his head. This had to be Jake, the department's self-appointed practical joker. Tom reflected that what Jake chiefly needed was an elementary grammar lesson, a better sense of humour, and possibly a little more work. He had poked his head in at four thirty this afternoon to say good-bye. On the other hand, Tom wondered, why hadn't he seen the scrap of paper before if it had been lying there all that time? Why hadn't the cleaner discovered it?

Tom studied the words and licked lips. The building was cloaked in a silence broken only by the humming of the strip lighting. He fingered the paper. His mouth felt dry and his knee had started to shake uncontrollably under the table. Those words OR ELSE, those black capital letters... Unwillingly, desperately trying to evade the thought, he remembered how much would be at stake for TYP Cosmetics when the results of his study were released. They would be facing financial ruin, or takeover at the very least. A noise outside the door, ever so slight, made Tom look up in panic, his heart in his throat. Wide-eyed, he watched as a large sheet of paper slid silently into view. On it was written, plain to see: YOU HAVE BEEN WARNED.

2

111

0   Tom stopped work

   A   because he was exhausted.   ☐
   B   for a few minutes to wash his hands.   ☐
   C   because it was time for him to stop.   ☐
   D   just before nine o'clock.   ☒

Q1   Tom was certain that

   A   his report was going to attract a lot of attention.   ☐
   B   TYP Cosmetics would be pleased with his report.   ☐
   C   his work was important but only for very few people.   ☐
   D   he would be rewarded more than his team.   ☐

Q2   Tom's opinion of the new cleaner was that

   A   his conversation was rather boring.   ☐
   B   he was the best cleaner they had ever had.   ☐
   C   he was good but he got in the way.   ☐
   D   his work was first class.   ☐

Q3   When Tom found the piece of paper, it occurred to him that

   A   his colleagues did not all work as hard as him.   ☐
   B   one of his colleagues was playing a joke on him.   ☐
   C   he had had enough of Jake's jokes.   ☐
   D   Jake was not good enough to be on his team.   ☐

Q4   The more Tom thought about the message, the more

   A   aware he became of noises in the building.   ☐
   B   he wished he had gone home earlier.   ☐
   C   anxious he became.   ☐
   D   worried he was about Jake.   ☐

Q5   Tom gradually realized there might be a connection between

   A   the message and one of his colleagues.   ☐
   B   Jake and TYP Cosmetics.   ☐
   C   the cleaner and TYP Cosmetics.   ☐
   D   the message and his report.   ☐

Q6   It was quite clear to Tom that the second message was

   A   only frightening because he was alone .   ☐
   B   threatening him personally.   ☐
   C   a joke in very bad taste.   ☐
   D   a warning from somebody trying to help him.   ☐

3

112

"Make Your Way Ahead 8"

http://www.oebv.at/sixcms/list.php?page=titelfamilie&titelfamilie=Make%20Your%20Way&modul=aktuelles&id=322598&aktuelles=Unterwegs+zur+neuen+Matura

Make Your Way to the Written Matura

Read the newspaper article, then choose the correct answer (A, B, C or D) for questions 1–6 and put a ⊠ in the correct box.

Reality TV: a dearth of talent and the death of morality

I've managed to miss out on reality TV until now. In spite of all the talk in Britain about nasty Nick and flighty Mel, and in America about the fat, naked bastard Richard manipulating his way to desert-island victory, I have somehow preserved my purity. I wouldn't recognise Nick or Mel if I passed them in the street, or Richard if he was standing in front of me unclothed.

Ask me where the Big Brother house is, or how to reach Temptation Island, and I have no answer. I do remember the American Survivor contestant who managed to fry his own hand but that's because he got onto the main evening news. Otherwise, search me. Who won? Who lost? Who cares?

The subject of reality TV shows, however, has been impossible to avoid. Their success is the media story of the (new) century. Success on this scale insists on being examined, because it tells us things about ourselves; or ought to.

And what cheap narcissism is here revealed! The television set, once so idealistically thought of as our window on the world, has become a dime-store mirror instead. Who needs images of the world's rich otherness, when you can watch these half-familiar avatars of yourself – these half-attractive half-persons – enacting ordinary life under weird conditions?

Who needs talent, when the unashamed self-display of the talentless is constantly on offer?

"Famous" and "rich" are now the two most important concepts in western society, and ethical questions are simply obliterated by the potency of their appeal. In order to be famous and rich, it's OK – it's actually "good" – to be devious. It's "good" to be exhibitionistic. It's "good" to be bad. And what dulls the moral edge is boredom. It's impossible to maintain a sense of outrage about people being so trivially self-serving for so long.

Oh, the dullness! Here are people becoming famous for being asleep, for keeping a fire alight, for letting a fire go out, for videotaping their cliched thoughts, for flashing their breasts, for lounging around, for quarrelling, for being unpopular, and (this is too interesting to happen often) for kissing! Here, in short, are people becoming famous for doing nothing much at all, but doing it where everyone can see them.

The problem with this kind of engineered realism is that, like all fads, it's likely to have a short shelf-life, unless it finds ways of renewing itself. The probability is that our voyeurism will become more demanding. It won't be enough to watch somebody being catty, or weeping when evicted from the house of hell, or "revealing everything" on subsequent talk shows.

What is gradually being reinvented is the gladiatorial combat. The TV set is the Colosseum and the contestants are both gladiators and lions; their job is to eat one another until only one remains alive. But how long, in our jaded culture, before "real" lions, actual dangers, are introduced to these various forms of fantasy island, to feed our hunger for more action, more pain, more vicarious thrills?

In the world outside TV, our numbed senses already require increasing doses of titillation. One murder is barely enough; only the mass murderers make the front pages. You have to blow up a building full of people or machine-gun a whole royal family to get our attention. Soon, perhaps, you'll have to kill off a whole species of wildlife or unleash a virus that wipes out people by the thousand.

And as in reality, so on "reality TV". How long until the first TV death? How long until the second? By the end of Orwell's great novel 1984, Winston Smith has been brainwashed. "He loved Big Brother." As, now, do we. We are the Winstonians now.

(Salman Rushdie on the perils of voyeurism in The Guardian Saturday June 9, 2001)

1 What opinion of TV reality shows does the author express in the first paragraph?
  A He only takes an interest when they make news. ☐
  B He can't understand why they are so popular. ☐
  C He has absolutely no interest in watching them at all. ☐
  D He finds them interesting from an academic point of view. ☐

2 Why does he believe reality TV shows are worth studying?
  A Because everyone is talking about them. ☐
  B Because their popularity offers insight into the human race. ☐
  C Because they are the future of TV. ☐
  D Because they are everywhere. ☐

3 Why is the author so worried about fame?
  A Because fame has become devalued. ☐
  B Because fame has become over-important in modern society. ☐
  C Because our fascination with fame has distorted our moral values. ☐
  D Because famous people are too concerned with showing off their money. ☐

4 Which of the following activities does the author not refer to in his list of typical reality show behaviour?
  A preparing meals. ☐
  B arguing with other contestants. ☐
  C taking clothes off. ☐
  D talking to the camera. ☐

5 What is the danger for TV reality shows?
  A People will get bored of them unless they find more exciting formats. ☐
  B Governments are under pressure to ban them from our screens. ☐
  C They could lead to murder. ☐
  D They are becoming very expensive. ☐

6 What point does the author make about modern life in the real world in the penultimate paragraph?
  A We have become obsessed by violence. ☐
  B Many people live in fear of a chemical attack. ☐
  C We have become insensitive to violence. ☐
  D People are becoming more and more violent. ☐

Read the newspaper article, then choose the correct answer (A, B, C or D) for questions 1–6 and put a ⊠ in the correct box.

My flight from fantasy

If you had asked me a year or so ago, I would have told you with smug certainty that travel doesn't broaden the mind, it merely raises the blood pressure. I'm not one of nature's travellers. I hate tunnels, fear air crashes, loathe coaches and get sea-sick. I'm also vegetarian, so 98% of the menu in any restaurant is a no-go area. I'm a lousy photographer, have a poor sense of direction, and never engage strangers in conversation because ... they are strangers. The one time I attempted to drive on the right, I displaced the wing mirrors of a row of cars. In short, I always felt I was better off at home, or in Tuscany, which is nicer than home but still sells Marmite.

But then, on a relatively stress-free three-hour flight, I started to think that it wouldn't be so bad to stay in my seat a few hours longer and find a different continent. I began to view travel not as a problem but a possibility. And one night, as my husband eyed a cheap-flight website and asked me for the thousandth time if I wouldn't fancy, say, San Francisco for a change, I said yes. So last month, I found myself hurtling down hills in a cable car and eating fresh raspberries in the middle of winter. But it wasn't the Golden Gate Bridge that broadened my mind; it was simply being elsewhere, encountering difference.

In your own country, it's easy to make assumptions about people based on accents, about areas based on architecture. But even in the US, ostensibly very like Britain, I found myself unqualified to make those judgments, and it was rather thrilling seeing things without the usual filter of prejudice or certainty.

I heard a black bus passenger's take on the Dick Cheney shooting: "I would have got seven years in jail for that, accident or no accident!" A homeless man of no more than 20, with few teeth and track marks on his neck, was so pleased with the pitiful couple of dollars we gave him, he told us all the best vantage points for photographing his city.

Perhaps most remarkable was a conversation overheard between a young woman with Down's syndrome, a young man with speech, hearing and mobility difficulties and their carer, who was taking them downtown on a cable car. Using an ingenious combination of signing, touching, pointing and shouting, they held a conversation that began with "Do you like butter?" and ended with "I love you". It was a complete relationship in microcosm, taking no more than the 10-minute journey, and with each breakthrough in communication greeted with howls of infectious laughter.

Travel, I have belatedly realised, is not just about sightseeing. Away from home, work, the school run, our routine, we are bombarded with new information, suddenly alert and alive. You've probably known that for years, but for me it's a revelation.

So the irony is not lost on me that just as I have overcome my fears about long-haul flights, the outlook for air travel has become rather bleak. If the aftermath of the Buncefield fire doesn't ground us, then it's quite possible that plans to restrict the spread of bird flu will. And even without that, it's hard to see how we can go on flying willy-nilly about the globe with oil supplies politically vulnerable and dwindling, and each long-haul round trip damaging the environment more than driving a car for a year.

Travel may well broaden the mind, but it also knackers the planet, and the more your mind is broadened, the more unacceptable that trade-off seems. On my journey through life I have always been accompanied by fear and guilt. Just as the former has loosened its grip, it seems as though the latter is going to keep me grounded.

(Rebecca Front in The Guardian March 8, 2006)

1 From what she tells us in the first paragraph, which of these might the author be happy doing?
A Eating a steak. ☐
B Reading a map. ☐
C Eating Marmite in Italy. ☐
D Talking to someone she doesn't know. ☐

2 Why did the author decide to go to San Francisco?
A She'd always wanted to see the Golden Gate bridge. ☐
B Her husband kept asking her to go with him. ☐
C They got some really cheap tickets. ☐
D She thought the extra few hours in an airplane might be worth the new experience. ☐

3 Why did she find her US experience exciting?
A Because she got to discuss US politics with real Americans. ☐
B Because things were so cheap there. ☐
C Because she liked being in an environment that she didn't understand 100%. ☐
D Because the people were genuinely friendly and ready to help. ☐

4 How did the three people on the cable car communicate?
A By getting the carer to translate. ☐
B By telling a lot of jokes. ☐
C By speaking quietly. ☐
D By mainly using their hands. ☐

5 What revelation about travel has the author had?
A That it makes our normal lives seem quite attractive. ☐
B It's not as dangerous as she thought. ☐
C That it doesn't need a lot of planning to be exciting. ☐
D That visiting new countries offers us a totally new experience. ☐

6 Why does the author believe she will not be taking any more long flights?
A Because she is concerned about the damage flying does to the environment. ☐
B Because she is afraid that other journeys won't be as good. ☐
C Because she is afraid of bird flu. ☐
D Because she's still really afraid of flying. ☐

116

Read the extract from a novel, then choose the correct answer (A, B, C or D) for questions 1–6 and put a ☒ in the correct box.

A short history of tractors in Ukrainian

Two years after my mother died, my father fell in love with a glamorous blonde Ukrainian divorcee. He was eighty-four and she was thirty-six. She exploded into our lives like a fluffy pink grenade, churning up the murky water, bringing to the surface a sludge of sloughed-off memories, giving the family ghosts a kick up the backside.

It all started with a phone call.

My father's voice, quavery with excitement, crackles down the line. 'Good news, Nadezhda. I'm getting married!'

I remember the rush of blood to my head. Please let it be a joke! Oh, he's gone bonkers! Oh, you foolish old man! But I don't say any of those things.

'Oh, that's nice, Pappa,' I say.

'Yes, yes. She is coming with her son from Ukraina. Ternopil in Ukraina.'

Ukraina: he sighs, breathing in the remembered scent of mown hay and cherry blossom. But I catch the distinct synthetic whiff of New Russia.

Her name is Valentina, he tells me. But she is more like Venus. 'Botticelli's Venus rising from waves. Golden hair. Charming eyes. Superior breasts. When you see her you will understand.'

The grown-up me is indulgent. How sweet – this last late flowering of love. The daughter me is outraged. The traitor! The randy old beast! And our mother barely two years dead. I am angry and curious. I can't wait to see her – this woman who is usurping my mother.

'She sounds gorgeous. When can I meet her?' 'After marriage you can meet.' 'I think it might be better if we could meet her first, don't you?'

'Why you want to meet? You not marrying her.' (He knows something's not quite right, but he thinks he can get away with it.)

'But Pappa, have you really thought this through? It seems very sudden. I mean, she must be a lot younger than you.'

I modulate my voice carefully, to conceal any signs of disapproval, like a worldly-wise adult dealing with a love-struck adolescent.

'Thirty-six. She's thirty-six and I'm eighty-four. So what?'

(He pronounces it 'vat'.)

There is a snap in his voice. He has anticipated this question.

'Well, it's quite an age difference …'

'Nadezhda, I never thought you would be so bourgeois.' (He puts the emphasis on the last syllable – wah!)

'No, no.' He has me on the defensive. 'It's just that … there could be problems.'

There will be no problems, says Pappa. He has anticipated all problems. He has known her for three months. She has an uncle in Selby, and has come to visit him on a tourist visa. She wants to make a new life for herself and her son in the West, a good life, with good job, good money, nice car – absolutely no Lada no Skoda – good education for son – must be Oxford Cambridge, nothing less. She is an educated woman, by the way. Has a diploma in pharmacy. She will easily find well-paid work here, once she learns English. In the meantime, he is helping her with her English, and she is cleaning the house and looking after him. She sits on his lap and allows him to fondle her breasts. They are happy together.

(From: Marina Lewycka, A Short History of Tractors in Ukrainian.- Penguin Books 2006)

1 Which of the following is <u>not</u> true about Nadezhda's father's new girlfriend?
 A She is almost fifty years younger than him. □
 B She has been married before. □
 C She has a child. □
 D She is not very attractive. □

2 What was Nadezhda's reaction when her father first told her the news?
 A She was overjoyed for him. □
 B She secretly disapproved. □
 C She thought he was joking. □
 D She told him he was mad. □

3 How does Nadezhda feel when her father tells her more about Valentina?
 A She wants to meet her new mother. □
 B She starts to feel jealous. □
 C She starts to feel happy for her father. □
 D She has mixed feelings. □

4 How does Nadezhda's father react when she says she wants to meet Valentina?
 A He is not so keen for them to meet immediately. □
 B He says he's too busy to meet her now. □
 C He gets angry with her. □
 D He invites her to the wedding. □

5 How does the Nadezhda's father respond to her question about the age difference?
 A He is surprised by it. □
 B He accuses her of being too conservative. □
 C He doesn't see any problem. □
 D He gets very angry with her. □

6 Which of these things do we learn about Valentina in the last paragraph?
 A She has a relative in the UK. □
 B She drives a Lada. □
 C She speaks English really well. □
 D She wants to work as a cleaner. □

Read the web page, then choose the correct answer (A, B, C or D) for questions 1–6 and put a ⊠ in the correct box.

The price of the Olympics

I first came to Vancouver when it was one of the prettiest cities in the world. Tucked into the foothills of the Rockies, built on a natural harbour, Vancouver seemed to have been built with an eye on its natural surroundings. It was typically Canadian, I thought. Proud, but in an understated way. Modern, but respectful. But oh dear – the Winter Olympics are coming to town and Vancouver's gone and messed itself up big time.

No doubt I'm going to be accused of generally having a bit of a 'down on progress', a one-sided view of 'development' … well, like I could care less…because what's happening to Vancouver's beautiful waterfront is absolutely typical of the impact greed has on a city's environment. Where there was once a relatively relaxed city centre all that can be heard now is the constant hammering of construction. Where there was once wide open vistas to the mountains, the trees, and the water, there is now a forest of glass and steel and a major fight on to grab the best views for those who can afford them at the expense of everyone else. And, according to the hotel staff, it's all because Vancouver is set to host the Winter Olympics in 2010.

The story goes that property here is 'relatively' cheap. There is going to be a huge influx of visitors to Vancouver for the Olympics (which puts its tent up for a couple of weeks and then moves on of course), and the City hopes that the attendant publicity will put Vancouver "on the map". Speculators and developers know that means that property prices will soar. They can then sell to the new incomers for a vast profit, take their money and go and smash up a low-cost, "undeveloped" area somewhere else.

The people who move in, of course, want views of the mountains, the trees, and the water – as did the people who already live here of course. There's really only one way that the developers can keep offering those spectacular views – and that's by putting up ever-higher towers or by putting towers in front of the towers that are already there, which then block off the views of the people who moved in a year or so ago. No, the developers aren't stupid – they know that the apartments that will sell for the highest prices are those with the best views. The solution? Put up higher towers to give the next lot of residents the views they want, which will inevitably block the views of those people who have only just moved in themselves. And so it goes on and on...

And buildings don't just exist as structures – they're built to be full of people. All of those new people will inevitably bring hundreds of extra vehicles, a demand for new roads and for more recreation facilities etc etc – and within a very short time a city that felt small and low-impact will have turned itself into a grid-locked metropolis that makes no effort at all to fit into the environment anymore, and all those people who moved here because of the quality of life start to wonder why they bothered...

You know, just for once wouldn't it have been wonderful if the Bald Eagles that hunt over Stanley Park, or the thousands of waterfowl that roost in English Bay, had had a say in all this development madness. "No you can't build here, I need this space." "No, come on, we fly hundreds of miles to find safety here and you want to put up huge glass buildings in our way – that's just not right...."

Of course, they're just ducks and eagles – what right have they got to say that a few rich developers can't get richer and a few more miles of beautiful waterfront shouldn't be turned into a playground for the human race...?

(From: http://www.charliesbirdblog.com/)

1 What was the writer's impression of Vancouver the first time he visited the city?
   A It was behind the times.   ☐
   B It was unlike any other city in Canada.   ☐
   C It was very futuristic.   ☐
   D It appreciated it's natural setting.   ☐

2 What is his impression on his second visit?
   A He is happy to see how it has become more modern.   ☐
   B It's more relaxed that it was before.   ☐
   C It's noisy and most of the views of the mountains have gone.   ☐
   D The rich people have bought all the best houses for themselves.   ☐

3 Why is Vancouver an ideal city for property developers?
   A Because they need new building for the 2010 winter Olympics.   ☐
   B Because Vancouver is fast becoming a very fashionable place to live.   ☐
   C Because it's very easy to get planning permission there.   ☐
   D A combination of low house prices and the hosting of a major
international event.   ☐

4 How are the developers able to keep offering apartments with great views?
   A They have bought up all the waterfront where the views are the best.   ☐
   B They keep the buildings low so everyone can enjoy the views.   ☐
   C They build taller and taller buildings.   ☐
   D They buy up the houses with the good views and redevelop them.   ☐

5 What is the danger of all this development for the city?
   A It will lose a lot of its wildlife.   ☐
   B People will move away because the city will lose its charm.   ☐
   C It will lose the thing that originally attracted people to live there.   ☐
   D It will become overcrowded.   ☐

6 Which of these descriptions do you think best describes the writer?
   A anti-Olympics   ☐
   B pro-conservation   ☐
   C anti-progress   ☐
   D pro-Canadian   ☐

Read the text below, then decide whether the statements (1-10) are true (T),
false (F) or not given (NG) in the text. Put a ☒ in the correct box.

		T	F	NG
1	John Daly has lost an average of around $5 million a year over the last 12 years.			
2	His addiction is the first thing he discusses in his new book.			
3	His father also had problems with gambling.			
4	He once celebrated beating Tiger Woods by gambling away his prize money and more.			
5	He has broken PGA rules with his gambling habit.			
6	He has neglected his family in order to pay off his gambling debts.			
7	Daly has also had problems with alcoholism.			
8	Daly is in danger of losing sponsorship because of his gambling.			
9	Daly uses sponsorship money to help clear his debts.			
10	He finally managed to beat his gambling addiction last year.			

Golfer in gambling debt shock

John Daly says he has lost between $50 million and $60 million during 12 years of heavy gambling, and that it has become a problem that could "flat-out ruin me" if he doesn't bring it under control.

Daly discussed his addiction to gambling in the final chapter of his autobiography, "John Daly: My Life In and Out of the Rough," to be released next Monday.

He told one story of earning $750,000 when he lost in a playoff to Tiger Woods last fall in San Francisco at a World Golf Championship. Instead of going home, he drove to Las Vegas and says he lost $1.65 million in five hours playing mostly $5,000 slot machines.

"If I don't get control of my gambling, it's going to flat-out ruin me," he says in the book, co-written with Glen Waggoner and published by HarperCollins.

The book got the attention of PGA Tour headquarters, and commissioner Tim Finchem met with Daly to discuss the problem.

Finchem said the book does not violate PGA Tour regulations, although "it is clear that he continues to be concerned about and grapple with significant personal challenges."

"I have expressed to John the tour's concern for his well-being, as well as his ongoing need to uphold the image and standards of the PGA Tour," Finchem said. "While we will continue to enforce the regulations and policies of the PGA Tour, I have advised John of the tour's willingness to support him in his efforts to deal with his personal issues."

The two-time major champion wrote that he has spent the last 10 years paying off gambling debts with his sponsorship income, hustling appearance money and "running myself ragged doing corporate outings instead of spending time with my family and working on my game."

He recalled former Dallas Cowboys linebacker Thomas "Hollywood" Henderson telling him at a Tucson, Ariz., rehab center in 1993 that Daly would find something he loves as much as drinking, and that he would have to be careful.

"The people around me ... were hoping, of course, that the 'something' would be practising golf. No such luck," Daly wrote. "What I found was gambling."

He said he owed $4 million to casinos in two years of gambling until he won the 1995 British Open at St. Andrews, his second major. That victory and the ability to get handsome appearance fees, enabled him to pay off the debt.

But the gambling continued.

Daly three-putted from 15 feet on the second playoff hole against Woods at Harding Park. He headed to Las Vegas and lost $600,000 within 30 minutes. He said he took out another $600,000 line of credit and lost that in two hours.

"And here's how my sick mind analyzed the situation," Daly wrote. "My sponsorship payments would be coming through in January, so I'd be able to pay everything off and get back to even by the beginning of the new year. Everything's fine. Everything's OK. No problem. Hell, yes, there's a problem."

Daly says he has taken more control of his life in the last six years.

"I'm off those ... medications. I don't drink JD (Jack Daniels) anymore. I don't beat up on hotel rooms and cars as much. Only gambling remains a problem," he wrote.

He said he plans to start at the $25 slots in the casinos and set a "walkout loss number," which would tell him it's time to leave.

"If I make a little bit, then maybe I move up to the $100 slots or the $500 slots, or maybe I take it to the blackjack table," he wrote. "It's their money. Why not give it a shot, try to double it? And if I make a lot, I can ...

Well, that's my plan."

(source http://sports.yahoo.com/golf/pga/news)

Read the text below, then decide whether the statements (1-10) are true (T),
false (F) or not given (NG) in the text. Put a ⊠ in the correct box.

		T	F	NG
1	The author was surprised by the heat on the mountain.			
2	The men used the ice to protect them from the sun.			
3	It was the first time that the author had put up a tent in such cold conditions.			
4	In training the author found it difficult to use his hands wearing such thick gloves so he took them off.			
5	The moment he put his hand back in the glove he was sick.			
6	He used hot water to bring the feeling back to his hands.			
7	The author felt embarrassed when his teammates started to help him.			
8	The author was a geography teacher in Phoenix.			
9	The author practised hard how to put up a tent because he didn't want to let down the team.			
10	The author was the first blind person to ever try and climb McKinley.			

Touch the top of the world

That afternoon in the blazing heat, we built snow walls around our campsite. I knew about McKinley's legendary cold, but no one had told me about the heat, reflecting like a mirror off the snow and burning my eyeballs through the leather flaps of my glacier glasses. Then the wind, chilled by the glacier, whipped past me, taking much of my body's warmth with it. "Windburn on top of sunburn. Get used to it," Chris laughed, observing Sam and me constructing snow fortresses out of the glacier. Sam cut blocks of blue ice from the floor while I placed them in a rectangle around the site to block the wind. Soon the walls were as tall as me. Then I cut steps into our fort while Sam packed the small gaps in the walls with snow. Finally the site was ready for the tent. I held one side and threw the other into the wind, which caught it and unfurled it. I laid it on the ground and oriented it by feeling the loops and pockets on the corners through my layers of gloves.

Months before on a training climb on Mount Rainier, a teammate assigned me to set up a tent on the Muir snowfield, on which wind and cold seemed to be the only constants. I was beginning to shiver as I knelt in the snow with the tent laid out in front of me. Through my thick gloves I couldn't feel the delicate sleeves of the fabric. I fumbled with it, clumsily trying to jam the pole through. Then I took my glove off so I could actually touch it. My hands were my eyes, but three frustrating layers of material over them made me feel blind. Only for a second, I thought. Just enough time to get the pole started in the sleeve. But sharp splinters of sleet pricked my bare skin and it went instantly numb. I stuffed my lifeless hand back inside the glove and beat it against my knee. When it came back to life, the pain was so intense I almost vomited from nausea. Not wanting to give up, I whipped off the other glove, but this hand too went numb before it even touched the tent fabric. Sam and Jeff approached. They had finished with the other tents and without saying a word started working on mine. The pain in my hands was nothing compared with my frustration and embarrassment, like a balloon expanding in my chest. I knelt in the snow, listening to the tent lifting up under the pressure of the poles, and I made a promise to myself. The things I could not do, I would let go; but the things I could do, I would learn to do well.

Afterwards in Phoenix, when the temperature was hovering above a hundred degrees, I took the tent to a field near the school where I taught and, with my thick gloves on, worked on setting it up and breaking it down and setting it up again. I heard cars slowing down on the nearby road, to gape, I imagined, at the lunatic in the blazing heat, in a tank top and mountaineering gloves, kneeling over a tent. But I refused to be the weak link of the team. I wanted them to put their lives in my hands, as I would put mine in theirs. I would carry my share. I would contribute as any other team member. I would not be carried up the mountain and spiked on top like a football. If I were to reach the summit, I would reach it with dignity.

(From: Erik Weihenmayer, "Touch the Top of the World: A Blind Man's Journey to Climb Farther Than the Eye Can See", Dutton Adult 2001)

For questions 1–10 answer by choosing from sections A–D. Some of the choices may be required more than once.

Which section mentions the following?

the length of time it took to produce the report.	1 ......
why some people will not believe the findings of the report.	2 ......
the atmospheric conditions that might look like UFOs.	3 ......
the fact that there are people who are convinced they've had alien contact.	4 ......
the lack of proof of the existence of objects that could crash and damage Earth.	5 ......
the reasons why the report has been made public.	6 ......
the anonymity of the report's author.	7 ......
the medical reasons why some people believe to have been abducted	8 ......
some of the airborne object that might get confused for UFOs.	9 ......
the reasons why the government does not plan any further investigations.	10 .....

UFO study finds no sign of aliens

A

A confidential Ministry of Defence report on Unidentified Flying Objects has concluded that there is no proof of alien life forms. In spite of the secrecy surrounding the UFO study, it seems citizens of planet Earth have little to worry about.

The report, which was completed in 2000 and stamped "Secret: UK Eyes Only", has been made public for the first time. Only a small number of copies were produced and the identity of the man who wrote it has been protected. His findings were only made public thanks to the Freedom of Information Act, after a request by Sheffield Hallam University academic Dr David Clarke.

B

The four-year study – entitled "Unidentified Aerial Phenomena in the UK" – tackles the long-running question of UFO-spotters: "Is anyone out there?" The answer, it seems, is "no". The 400-page report puts it like this: "No evidence exists to suggest that the phenomena seen are hostile or under any type of control, other than that of natural physical forces." It adds: "There is no evidence that 'solid' objects exist which could cause a collision hazard."

So if there are no such things as little green men in spaceships or flying saucers, why have so many people reported seeing them? Well, here is the science bit.

"Evidence suggests that meteors and their well-known effects and, possibly some other less-known effects are responsible for some unidentified aerial phenomena," concludes the report.

"Considerable evidence exists to support the thesis that the events are almost certainly attributable to physical, electrical and magnetic phenomena in the atmosphere, mesosphere and ionosphere."

"They appear to originate due to more than one set of weather and electrically charged conditions, and are observed so infrequently as to make them unique to the majority of observers."

C

People who claim to have had a "close encounter" are often difficult to persuade that they did not really see what they thought they saw. The report offers a possible medical explanation.

"The close proximity of plasma related fields can adversely affect a vehicle or person," states the report.

"Local fields of this type have been medically proven to cause responses in the temporal lobes of the human brain. These result in the observer sustaining (and later describing and retaining) his or her own vivid, but mainly incorrect, description of what is experienced."

D

There are, of course, other causes of UFOs – aeroplanes with particularly bright lights, stray odd-shaped balloons and strange flocks of birds, to name but a few.

Yet, it will be difficult to convince everyone that there is a rational explanation for all mysterious movements in the sky.

Some UFO-spotters believe governments will always cover up the truth about UFOs, because they are afraid of admitting that there is something beyond their control.

It is not clear how much time and effort the MoD has spent looking at the skies in recent years, but it appears there are no plans for an in-depth UFO report like the one written in 2000.

An MoD spokesperson said: "Both this study and the original 'Flying Saucer Working Party' (already in the public domain in the National Archives) concluded that there is insufficient evidence to indicate the presence of any genuine unidentified aerial phenomena.

"It is unlikely that we would carry out any future studies unless such evidence were to emerge."

(By Mark Simpson BBC News http://news.bbc.co.uk/2/hi/uk_news/4981720.stm)

For questions 1–5 you must choose which of the extracts A–F below fit into the numbered gaps in the article. There is one extra paragraph that does not fit into any of the gaps.

An invisible army of workers strikes for the right to be seen

By Ben Macintyre

US wants it both ways – closed borders and cheap labour

The best place to witness America's immigration dilemma is not the US-Mexican border. Go instead to any suburban school and see the Latino nanny, an illegal immigrant, picking up the children while their middle-class parents are still at work. Or stand outside any office block in the evening, and watch the Mexican janitors go in to scrub out the toilets as the suited workers go home.

1	

America's reliance on illegal immigrant labour — to keep the economy buzzing, prices low and the standard of living higher than anywhere else in the world — is the worst-kept secret in the US economy, and everyone knows it.

2	

Publicly, most people deplore it. Privately, everyone does it — often while simultaneously demanding tighter border controls and greater restrictions on immigration. At best, the general American attitude towards employing illegal immigrants involves a strange form of double-think. At worst, it is hypocritical.

3	

Some have compared the immigration issue to that of Prohibition in the 1920s, when alcohol was illegal in the country. Public figures loudly condemn the employment of illegal workers, just as they once denounced the demon drink, but trying to persuade them to give it up is another matter. The illegal nanny has replaced the illicit mistress as the figure most likely to topple a modern American politician.

4	

"Nannygate" rumbles on today. Several big names have already fallen. Just a few years ago, Bernard Kerik, President Bush's choice for homeland security chief, was dropped when it emerged that he, too, had employed an illegal immigrant as a nanny. In 2000, Linda Chavez, a prominent Latino Republican, wrote a trenchant newspaper column asking: "Wouldn't it be better to change the immigration laws to allow more people to reside here legally than simply to turn a blind eye to those who are violating the laws now on the books?" A year later, Mr Bush nominated Ms Chavez to be Labour Secretary. But then it was revealed that she had given money to an illegal immigrant from Guatemala who had done her household chores. Her nomination was withdrawn.

126

5

In the meantime, many of those same conservatives, like everyone else, will continue to employ those same immigrants to cut their grass, look after their children and scrub their lavatories.

(From: The Times, April 29, 2006)

A   American conservatives now want the millions of illegal immigrants in the US classified as felons and a fence built quickly along the Mexican border.

B   Since 1986, it has been illegal to employ workers without the necessary work permits and residency papers, yet the employment of illegal labour is endemic. Everyone knows about it.

C   On the other hand, no media attention is ever directed at the plights of the illegal immigrants themselves; the long hours of work, the appalling conditions many face and the constant fear they live in that at any moment they may be found out and deported.

D   The US Government finds itself in the peculiar position of spending huge amounts of money to police a vast border, ineffectively, in order to keep out the cheap labour on which the US economy depends.

E   Zoe Baird, Bill Clinton's first choice as Attorney-General, was scuppered by her failure to pay taxes for a domestic worker without papers. Kimba Wood, his second nominee, went the same way when it emerged that her baby-sitter was an illegal immigrant.

F   Or go to any American meat-packing plant, construction site or tomato field and see the hundreds of thousands of illegal immigrants employed, at low wages, to do jobs most native-born Americans do not want.

127

SECTION 2 | The modern family

1 Work in pairs. Choose one of the "family" members below and describe them for your partner to guess.

Example: *"It's your brother or sister's daughter."*
"That would be my niece."

nephew	half-sister	brother-in-law	cousin	stepfather
niece	first cousin once removed	adopted son	great uncle	godmother

2 These days there are many different types of family. Here are some.
Match them with the definitions.

1   nuclear family

a   Used to talk about all the relatives in a family, including grand-parents, aunts, uncles etc.

2   extended family

b   A family with only a mother or a father caring for the children.

3   single parent family

c   The family created when two divorced people remarry and bring their children with them.

4   blended family

d   A family which offers temporary care for parentless children, until a permanent family can be found for them.

5   adoptive family

e   The 'traditional' family consisting of the mother, father and children.

6   foster family

f   A family in which the children are not biologically related to their parents.

Christina Moser

Adresse	Gonzagagasse 2/31
	1010 Wien
Telefon	0699 / 176 172 90
E-Mail	christinamoser_@gmx.at
Geburtstag	01.03.1987
Geburtsort	Grieskirchen
Staatsbürgerschaft	Österreich

Ausbildung

2006 – 2012	Universität Wien, Lehramt Englisch & PP
1997 – 2005	Wirtschaftskundliches Realgymnasium Wels
1993 – 1997	Volksschule Wels/Pernau

Praktische Erfahrung

– Unterrichtspraktikum Englisch, Sacré-Cœur
– Unterrichtspraktikum PP, BG/BRG Kundmanngasse

– In der Zusammenarbeit mit der Institution Buch.Zeit Wels:
Betreuung von Kindern und Jugendlichen im Rahmen diverser
Buch- und Lesefestivals wie Lesetopia Wels.
Aufgabenbereich: Stationenspielbetreuung & Aufsicht

– Privater Nachhilfeunterricht Englisch

Interessen

Reisen
Tauchen
Fotografie

Sprachkenntnisse

Englisch
Französisch