The interaction between affect and meta-cognition in language use: the case of foreign language anxiety

Ph.D. Thesis

Vita Kalnberzina

Supervisor: Dr. Caroline M. Clapham

Lancaster University

September 23, 2001

Abstract

The thesis explores the interaction between affect and meta-cognition as conceptualised by Bachman and Palmer's (1996) model of language use. Anxiety is used to operationalize the concept of affect and develop an empirical model of interaction between meta-cognitive strategies, affect and language performance.

The theoretical part examines the use of concepts of affect and meta-cognition in psychology and linguistics and the existing models of interaction between the two. Cognitive theories of affect and anxiety are reviewed to define the functions of affect in language use and provide the basis for the practical studies.

The practical research is concerned with exploring the causes of foreign language anxiety and detecting the effects of foreign language classroom, test trait and test state anxiety on language performance. Meta-cognitive competence is explored using Purpura's (1999) Meta-cognitive strategy questionnaire. Research methods include an observation, two interviews and two questionnaire studies.

The mathematical modelling method (SEM) is used to develop an empirical model of interaction between meta-cognition, anxiety and explore their impact on language proficiency.

Acknowledgements

This paper was created with the help of many people.

First of all I would like to thank my colleagues who participated in the development of the centralised Year 12 examination in Latvia: Dr. Dianne Wall (Lancaster University) the trainer of the team of examiners (including myself) and the team members without whom there would never have been a Year 12 examination as we know it now: Rita Kursite (developer and chief examiner of the Listening test), Inga Lande (Reading test), Dace Dumpe (Language Use Test), Ilze Straupmane (Speaking Test).

Secondly I would like to thank the staff of the Curriculum and examination centre for graciously allowing me to use their data and the help I received in administration of my questionnaires.

Thirdly I would like to thank approximately ten thousand test-takers for filling in all my questionnaires and answering all my questions.

Fourthly I would like to thank the British Council for sponsoring my PhD studies at Lancaster University.

Fifthly I would like to thank Dr. Caroline Clapham for her time and effort spent on Year 12 examination while training us how to use statistics, but especially for her support in the role of supervisor when I was trying to make sense of the Year 12 examination, the competences it evaluated and the individual variables that affected the measurement. I can easily say that without her this thesis would not have been written.

And finally I would like to thank my husband for taking responsibility for the data processing during the examination development period, as well as for the help and support I received while writing my PhD thesis.

Contents

List of tables

List of figures

List of appendices

Appendix 1 Year 12 Examination materials 1999

Appendix 2 Year 12 Examination materials 2000

Appendix 3 Transcript of interview-1999 (Translation)

Appendix 4 Transcript of interview-2000 (Translation)

Appendix 5 Questionnaire of Study 3 -1999 (Translation and original)

Appendix 6 Questionnaire of Study 4-2000 (Translation and original)

Appendix 7 Histograms for Year 12 Examination results 1999

Appendix 8 Histograms of Year 12 Examination results 2000

Appendix 9 Item-analyses of Year 12 Examination results 1999

Appendix 10 Item-analyses of Year 12 Examination results 2000

Appendix 11 Factor analyses of Study 3-1999

Appendix 12 Factor analyses of Study 4-2000

Appendix 13 Intercorrelations Study 3-1999

Appendix 14 Intercorrelations Study 4-2000

Appendix 15 EQS Foreign language anxiety model-1999

Appendix 16 EQS Foreign language anxiety model -2000

Chapter 1   Introduction

The topic of the thesis 'the interaction between affect and meta-cognition in language use' brings together language ability, test situation and test-taker characteristics (meta-cognition and affect in this study). If we treat test as a product we can separate test, language use and test-taker characteristics and use separate theories to investigate and explain them. If we view a foreign language test as a process we see that it is impossible to perform well without understanding the task needs and selecting the appropriate language elements. The test-taker has to assess the task demands and his or her ability as well as assess his or her performance during the test. Bachman and Palmer (1990) say that assessment strategy is a meta-cognitive strategy. Stevick (1999) considers that providing feedback on our performance is one of the functions of affect. LeDoux (1999) and Damasio (2000) consider that neither cognition nor affect can be understood separately.

A similar observation can be found in applied linguistics: Gardner (1997) considers that language achievement is influenced by language aptitude, attitudes and motivation, anxiety and language learning strategies, but that none of them operates in isolation. He says that some of these have elements in common and it would be beneficial to determine how the various variables relate to one another and how they operate in unison to influence individual differences in second language acquisition. Gardner proposes that *the time is ripe for meta-analyses that focus not only on the relationships of different variables to language achievement but also on the relationships of these variables to each other and to other variables associated with the acquisition of a second language. Such research is necessary to permit the formal modeling that is now due* (Gardner 1997, p. 40).

The interaction between meta-cognition and motivation is well researched and recorded (see Okada et al 1996, Weinert 1987), but the interaction between anxiety and meta-cognition, has not been much researched in spite of the fact that recent findings of research on meta-cognition and anxiety suggest that 'assessment' is the basis of both phenomena:

1. Purpura's (1999) research suggests that *meta-cognition* is a unitary concept consisting of different 'assessment' and evaluation variables

2. Bachman and Palmer (1996) say that *affective schemata provide the basis on which language users assess, consciously or unconsciously, the characteristics of language use task and its setting in terms of past emotional experiences in similar contexts* (Bachman and Palmer 1996, p. 65).

The absence of research on interaction between anxiety and meta-cognition prevents researchers to examine the role of 'affect' [emotion minus motivation and bodily sensation (Cacioppo et al 1999)] in language use in general. For example, Purpura (1999) says that although Bachman and Palmer (1996) *do state that meta-cognitive strategies interact with language knowledge and affect, it remains unclear as to how affect might be operationalized and how it might interact with meta-cognitive strategies* (Purpura 1999, p.21). Therefore although anxiety is focus of this study, it is examined as an affective variable that not only affects language performance, but also simultaneously interacts with meta-cognition.

1.1 Theories on interaction between cognition and affect

My thesis could be roughly divided into two parts: theoretical research and practical research. The theoretical part of the thesis consists of three chapters: the interaction between meta-cognition and affect in psychology (Chapter 2), and in language use

(Chapter 3), and anxiety as a result of interaction between affect and meta-cognition (Chapter 4).

1.1.1 Psychology on interaction between cognition and emotion

Chapter 2 deals with research into the interaction between meta-cognition and affect in psychology. It is divided into three parts: meta-cognition (2.1), affect (2.2) and models depicting interaction between the two (2.3) (I will follow this pattern, meta-cognition, affect and interaction between the two, in most chapters).

The first section on meta-cognition (2.1) is largely based on Brown's (1987) views on meta-cognition and its role in social psychology, the connection between meta-cognition and consciousness on the one hand, and the link between the meta-cognition and cognition on the other hand. Liddell (1949) considers that exploring anxiety helps us to understand intelligence, I could say just the opposite, exploring the theories of meta-cognition helped me to understand the basis of anxiety.

The second section (2.2) discussing affect and emotion is largely inspired by the views of Damasio (2000) and LeDoux (1999). My thesis could not have been written in the time when emotion and cognition were seen as opposites: cognition as the salvation of civilization and emotion symbolizing chaos that had to be kept in control by cognition (see Cohen 1998). Now that the situation in psychology has changed and it is admitted that cognitive appraisal is the basis of emotion (Scherer 2000) and emotion is an integral part of reasoning (Damasio 2000), the interaction between the two aspects of personality has become a popular field of research and the lists of the functions of emotions discussed in this section not only describe the nature of emotions, but also help us to understand how we think and why we think in the way we do.

The third section (2.3) explores the models of interaction between emotion and cognition as well as Scherer's (2000) model of emotion-production, which shows explicitly the interaction between the emotion and cognition in spite of its name. Another reason for placing this theory in this section is the fact that Scherer (2000) uses self-organisation theory as its basis, a theory that to my mind can finally put Humpty Dumpty (emotion and cognition) together again (Lazarus et al 1984).

1.1.2 Meta-cognition and affect in language learning and testing

The third chapter deals with the interaction between meta-cognition and affect in applied linguistics.

The first section of the chapter (3.1) explores the utilization of the concept of meta-cognition in language use models: how it was first incorporated just as supporting strategies (Canale and Swain 1980, Tarone and Yule 1989) and later developed as an integral part of language competence in Bachman and Palmer's (1996) model of language use.

The second section of this chapter (3.2) is largely based on Scovel's (1991) views on effect of affect and on Stevick's (1999) framework of roles of affect in language use. Although Stevick proposes affect as a system and does interpret many of the latest findings in psychology of the role of emotion in the applied linguistics context, he does not incorporate the system of affect into the existing language use models explicitly.

Bachman and Palmer, on the contrary, have provided a framework that incorporates both cognition and affect in the language use model [discussed in the third section of this chapter (3.3)] without naming explicitly the functions of affect

in language use. Therefore my research (both theoretical and practical) utilizes both Bachman's framework and Stevick's list of functions or roles of affect to explore the role of test anxiety in language use.

1.1.3 Anxiety as an interaction between cognition and affect

Chapter 4 starts with a glimpse of the different views of anxiety and its correlates (stress, worry and fear) in psychology. It uses the opposition of healthy and neurotic anxiety (May 1979) to introduce the opposite roles anxiety can play in human development: from motivating one's personal development (Kierkegard 1849 and May 1979) to destroying human consciousness (Beck and Emery 1985).

The test anxiety section (4.2) explores two different theories of test-anxiety: cognitive interference theory (Sarason1978, Wine 1980) and skills deficit theory (Kirkland and Hollandsworth 1980, Paulman and Kennely 1984). These two theories to my mind explore the cause and effect of test anxiety: skills deficit could be considered as a cause of anxiety, but cognitive interference as one of the effects of the negative evaluation of the situation (one of the roles of affect, Stevick 1999).

The last section on anxiety explores the findings of foreign language anxiety research (section 4.3). The traditional view of foreign language anxiety as a social, situational anxiety (Gardner and MacIntyre 1991) is now further developed by Onwuegbuzie et al (2000) as input, processing and output anxieties. A different view of foreign language anxiety is offered by Cheng et al (1999) and Saito et al (1999): they see foreign language anxiety as a complex consisting of reading, listening, writing and speaking anxiety.

The existence of different views on foreign language anxiety defined my research questions: what is foreign language anxiety and what are its causes and effects, what is the interaction between anxiety, meta-cognition and language performance? My main question is whether test anxiety has a cognitive basis (as Eysenck 1992 proposes) and how this affects test-takers' language performance.

1.2   Practical research

The place of research is the Year 12 English language examination in Latvia. This is a proficiency test; the focus of the test is English language performance rather than knowledge assessment. Therefore examination of the results of the Year 12 examination and the data provided by the research instruments (interviews and questionnaires) allowed me to investigate the interaction between language performance and individual characteristics (both cognitive and affective). The data were analyzed using both qualitative (Studies 1 and 2) and quantitative research methods (Study 3).

Different data elicitation methods were used to cross-validate the information obtained: Lickert scale measurement, answers to questionnaires, comments on reasons of anxiety, observation and semi-structured interviews.

The focus of Study 1 (Chapter 8) was to register signs of anxiety in all language skill tests using an observation method using Oxford's (1999) framework of anxiety in language acquisition process. All the signs of foreign language acquisition anxiety could be observed also during a language test.

The aim of Study 2 (Chapter 9) was to examine the causes of test anxiety with the help of interviews. Discussing the causes of anxiety the test-takers made frequent references to their ability and task demands, therefore the transcripts of the interview were analysed using both theories of anxiety and meta-cognition.

Study 3 (Chapter 10) aimed to measure the level of anxiety and its interaction with the use of meta-cognitive strategies with the help of several questionnaires. The test-takers were divided into groups according to their level of proficiency and anxiety. The comparison of the use of meta-cognitive strategies in groups of different anxiety level suggested interconnection between the two concepts. Correlation coefficients and Structural Equation Modeling provided further insights into the nature of the interaction between cognitive and emotional variables during language use.

I will start with the analysis of the existing theories of interaction between cognition and emotion in psychology.

Chapter 2   The interaction between meta-cognition and affect in psychology

Interaction between affect and cognition is researched by psychologists and linguists and referencing across the fields is common. Nevertheless I decided to separate their findings into two separate chapters (Chapters 2 and 3) because the focus of linguistics is different from that of psychology. Psychologists use language competence as a tool to understand how the mind works (see the intelligence tests) while linguists use the theories on interaction between cognition and affect to understand how language is produced and processed.

This chapter will review literature on meta-cognition (section 2.1), affect (section 2.2) and the interaction of the two (section 2.3) as seen in psychology. I will be using both social and experimental psychology texts which might cause some difficulty with the terminology (for example, the difference between affect and emotion, cognition and meta-cognition), but as each field offers a different kind of information on the topic under discussion both need to be addressed and I will do my best to explain the differences in terminology as they appear.

2.1   Meta-cognition

There are different views in social psychology of what the term 'meta-cognition' means: does it mean just knowing, or does it also imply ability; does it cover procedural knowledge, declarative knowledge or both?

Brown (1987) defines meta-cognition as 'understanding': *meta-cognition refers to understanding of knowledge, an understanding that can be reflected in either effective use or overt description of the knowledge in question* (Brown 1987, p.65). She

considers that *what is at issue is the concept of degree of understanding* (ibid.). The same idea was proposed earlier by Piaget (1976) and supported by Gombert (1992).

Brown (1987) considers that: *it is often difficult to distinguish between what is meta and what is cognitive... the confusion that follows the use of a single term for a multifaceted problem is the inevitable outcome of mixing metaphors* (Brown 1987, p. 66). Although most researchers distinguish between cognition and meta-cognition, Brown considers that because the interpretations differ, it is not possible to reach a common understanding.

Brown considers that when we use the term 'meta-cognition', it is necessary to distinguish between

1. knowledge about cognition (stable, statable, often fallible and late developing information) and

2. ability to control the use of knowledge (activities to regulate and oversee learning: predicting outcomes, scheduling strategies, monitoring activities and checking outcomes).

Because of all the various uses and misuses of the term of 'meta-cognition', Brown proposes that the term should be *pensioned-off* or at least severely restricted in its purposes and should refer to knowledge about cognition where it is statable. Brown suggests that *process terms, such as 'planning ahead', 'monitoring', 'resource allocation', 'self questioning' and 'self directing' should be used alone, without the addendum 'meta-cognition'* (Brown 1987, p. 106).

If we take a look at a contemporary definition of cognition we find that cognition is seen not only as knowledge and the ability to control knowledge during an activity, but is also seen as a process of control: according to Sparrow and Davies (2000) *the*

term "cognition" refers to the highest levels of various mental processes such as perception, memory, abstract thinking and reasoning, and problem solving as well as the more integrative and control processes related to executive functions such as planning, choosing strategies and the enactment of these strategies. They propose that although there are different approaches to cognition, *almost all of them are concerned with the existence of multiple component processes* (Sparrow and Davies 2000, p.117). This suggests that in psychology the term 'meta-cognition' seems to have been taken over by 'cognition'.

2.1.1   Levels of access to meta-cognition

One of the reasons why there are various views of what the term 'meta-cognition' stands for, is that it is used to denote both conscious and subconscious processing (Brown 1987). As Purpura (1999) found that the level of processing (conscious or automatic processing) affects language performance I will focus here on the difference between automatic and controlled processing.

2.1.1.1 Conscious access to meta-cognition

Brown (1987) proposes that the difficulty in investigating meta-cognition is caused by the human inability to report strategies, which can be explained by the roots of meta-cognition (its connection with human consciousness). She comes to the conclusion that *conscious access to the routines available to the system is the highest form of mature human intelligence* (Brown 1987, p.71).

Damasio (2000) makes a similar observation on public consciousness. He says that understanding of Western thought probably marched in a reverse order to the complexity of the phenomena connected with consciousness; the more complex notions were discussed first when the deeper, more primitive notions were not even

thought of. For example, the word 'consciousness' was developed on the basis of the more complex concept represented by the word 'conscience' (in Latin 'gathering of knowledge') (Damasio 2000, p.231).

Brown explains this in terms of evolution: *in the course of evolution, cognitive programmes become more accessible to other units of the system and therefore, may be used flexibly in a variety of situations. This flexibility is a hallmark of higher intelligence, reaching its zenith at the level of conscious access and control, which affords wide applicability across a wide range of mental functioning* (Brown 1987, p.71).

Brown differentiates between 'multiple' and 'reflective' access to meta-cognitive knowledge:

1. 'multiple access' refers to the ability to use knowledge flexibly and the ability to vary the use of knowledge systematically to fit a wide range of situations,

2. 'reflective access' refers to the ability to mention, as well as use, the components of the system.

2.1.1.2 Automatic versus controlled processing

One of the reasons why it is difficult to research the use of meta-cognitive strategies is the fact that it is difficult to establish whether a person simply is not aware of his or her strategy use or whether he or she does not use that strategy. This difficulty is the result of the existence of different levels of consciousness of processing.

Brown (1987) distinguishes between 'controlled' and 'automatic' processing. She defines 'controlled processing' as a comparatively slow, serial processing limited by

short-term memory constraints, requiring effort and providing a large degree of control.

She says that 'automatic processing' is a fast parallel process, not limited by short-term memory; it requires little effort and demands little direct control. Brown differentiates between two forms of automatic processing:

1. activities that appear to be common to all age groups and rarely demand intensive strategy effect (for example, some forms of recognition in reading)

2. and activities that were originally effortful but because of extensive training and experience have become automatic.

Brown uses the example of an efficient reader to explain the difference between the two processes: *the skilled reader's top-down and bottom up reading processes are so fluent that he or she can proceed merrily on automatic pilot until alerted to a comprehension failure by some triggering event... In the process of disambiguation and clarification, the individual enters a controlled, deliberately planful, strategic state that is quite distinct from the automatic pilot state* (Brown 1987, p. 80).

Here one could add that the event could be real or imagined. For example, a test-taker, who suddenly remembers a previous failure may stop working on auto-pilot because of a sudden loss of confidence and may start checking and rechecking all the answers to try and consciously monitor the reading process. If we accept this possibility, we have to accept the fact that active consciousness and self-awareness is not always the highest state of intelligence. It is rather the ability to control not only one's knowledge, but also one's emotions to achieve one's aims.

2.1.1.3 'Flow': full control of consciousness

'Flow' is a term that was introduced by Csikszentmihalyi (1998) to describe a 'loss' of self-consciousness that can occur if there is a unity between cognition and emotion: *consciousness is full of experiences and these experiences are in harmony with each other. Contrary to what happens all too often in everyday life, in moments such as these, what we feel, what we wish, and what we think are in harmony* (Csikszentmihalyi 1998, p.28).

He considers that to achieve 'flow' we need to be in a specific situation:

1. flow tends to occur when a person faces a clear set of goals that are mutually compatible (as in games like tennis or chess) when we do not have to question what should be done at each moment

2. activities need to provide immediate feedback to tell us how well we are doing

3. a person's skills need to be fully involved in what he or she is involved in overcoming a challenge that is just about manageable. Optimal experiences usually involve a fine balance between one's ability to act and available opportunities for action.

Csikszentmihalyi investigated the experiences of artists, musicians, athletes and has found that although they all did different things, when experiencing flow, their description of their experiences were very similar: *when goals are clear, feedback relevant and challenges and skills are in balance, attention becomes ordered and fully invested. There is no space in consciousness for distracting thoughts, irrelevant feelings. Self-consciousness disappears, yet one feels stronger than usual* (Csikszentmihalyi 1998, p. 29). This is similar to some of the descriptions of test-takers' experiences during a Speaking test (see section 9.1), when the task is clear, but

is nevertheless felt to be challenging and the interviewer's natural reaction provides an immediate feedback. The test-takers feel fully immersed in conversation and are in fact enjoying their experience.

2.1.2 Role of consciousness

Brown (1987) connects meta-cognition with understanding and consciousness. According to Piaget (1978), the role of consciousness is to liberate the regulatory functions from active testing and instead *elaborate operations on operations. The child thereby becomes capable of varying the factors in his experiments, of envisaging the various models that might explain the phenomenon, and of checking the latter through actual experimentation* (Piaget 1976, p.352).

Damasio (2000) proposes a similar idea: *Creatures with consciousness have some advantage over those that do not have consciousness. They can establish a link between the world of automatic regulation and the world of imagination, the world in which images of different modalities can be combined to produce novel images of situations that have not yet happened, the world of planning, the world of formulation of scenarios and prediction of outcomes. The sense of self links forethought on the one hand to pre-existing automation on the other* (Damasio 2000, p. 304).

Piaget's (1978) and Damasio's (2000) definition of consciousness contains elements of meta-cognition. Damasio considers that the power of consciousness comes from the effective connection it establishes between the biological machinery of the individual life regulation and the biological machinery of thought. That connection is the basis for the creation of an individual concern, which permeates all aspects of thought processing, focuses all problem-solving activities and inspires the ensuing solutions.

2.1.3 Levels of control

The term 'levels of control' was introduced in psychology from models of information processing and could be paraphrased as levels of ability to control one's own actions: from total lack of awareness to full awareness and control of one's self. Brown (1987) considers that apart from the historical connection between meta-cognition and consciousness research, the next most influential influence on our understanding of meta-cognition comes from information processing research. Its impact on our understanding of the functioning of meta-cognition is most important.

Brown draws parallels between our understanding of human consciousness and technological development: as the technological capabilities of problem solving systems grew in the 1970s, so did the systems' self-awareness. The executive system of an information-processing model had to be aware of its own resources, its shortcomings and its abilities. In 1978, Brown, reviewing the executive functions of problem solving systems, wrote: *The basic requirements of an executive system include the following abilities:*

1. *to predict the system's capacity limitations*

2. *to be aware of its inner resources and their appropriate use*

3. *to identify and characterise the problem at hand*

4. *to plan and schedule appropriate problem solving strategies*

5. *to monitor and supervise the effectiveness of the routines*

6. *to dynamically evaluate its operations in the face of success or failure* (Brown 1978 p.152).

If we compare these abilities to the executive functions in the definition of cognition (Sparrow and Davies 2000) we cannot help noticing the similarities. If, on the other hand, we compare the list of abilities of an executive system with Bachman and

Palmer's (1996) model of language use (see section 3.2.3) we notice that two of the meta-cognitive strategies (planning and assessment) can be matched with abilities in the executive system's list, but that the first one (goal-setting) does not have a counterpart. This differentiates the understanding of the concept of meta-cognition in applied linguistics from cognition; in linguistics it means not just controlling cognition, it means managing and using it for one's own goals. This is where Brown's (1987) metaphor no longer applies and the human mind starts differing from information processing models.

To exemplify the impact of the information technologies on our understanding of the human mind, I will just mention one example. Piaget (1976) writing at approximately the same time uses the phrase 'levels of regulation'. He considers that there are three different levels of regulation that are used by learners (autonomous, active and conscious regulation):

1. 'autonomous regulation' is a part of any knowing act: learners continually regulate their performance, fine-tuning and modulating their actions and making unconscious adjustments

2. 'active regulation' is a trial error regulation where the learner constructs theories in action and tests them to develop some unifying principles that can lead to successful problem solving

3. 'conscious regulation' involves the mental formulation of a hypothesis that can be tested via imaginary confirmatory evidence or counter examples. For example, we can consciously invent, test, modify and generalize theories and discuss these operations with others.

As Piaget does not make a connection between cognition and emotion, his model of mind at the highest level can control actions, invent theories and even discuss them with others, but it is not clear what for. What we have here in Piaget's description is a high quality processor that has an inner control over its operations, which can lead to successful problem solutions and even invention of new theories, but it is not clear who is to choose which problems to solve and what theories to invent. If there is no mechanism that chooses one's goals, all the operations mind can accomplish seem machine-like: it needs to be manipulated by somebody else, or something else (for example, emotions). The next section will examine the notion of emotions and their role in decision-making.

2.2   Affect and emotion

This section will review the present concept of 'emotion' in psychology and its views on the role of emotion in human consciousness. It will also explore the interaction between emotion and cognition since most researchers when discussing the role of emotion describe its influence on cognition.

The task is daunting because emotion is examined by several sciences: sociologists argue that emotion should be treated in terms of social functions because the most important characteristic of emotion is its role in communication with others (Oatley and Jenkins 1996), but psychologists Power and Dalgliesh (1998) say that the social role of emotion is only the tip of an iceberg, while the psychological role forms most of the rest of the iceberg.

Even definition of the subject matter is a problem: Caccioppo et al (1999) define emotion as 'affect' plus 'motivation' plus 'bodily sensation'. Damasio (2000), however, considers that 'affect' is often used as a synonym of 'mood' or 'emotion', although it is more general and can mean both mood (states of emotion that are

frequent over long periods of time) and emotion (publicly observable collection of responses) and feeling (private, mental experience of an emotion, *often poised at the very threshold that separates being from knowing*) (Damasio 2000, p. 43).

LeDoux (1999), a physiologist, shows the scope of the views on emotions of the present day sciences. These propose that emotions are:

1. bodily responses that evolved as part of the struggle to survive,

2. mental states that result when bodily responses are sensed by brain,

3. ways of acting or ways of talking,

4. thoughts about situations in which people find themselves,

5. social constructions that happen between rather than within individuals.

As language testing is a social science it might be sensible to restrict this overview to social psychology. However, the feedback from the test-takers suggests that bodily sensations, such as tiredness, cause additional nervousness and play an important role in the test-taking experience. In addition Fransson's research (1984) suggests that goals or types of motivation determine the level of anxiety and its influence on performance during a language test and it is psychology that examines the relationship between goals and emotions. I will, therefore refer to research that explains emotion regardless whether it comes from social or cognitive psychology and will use the studies' ability to shed light on the interaction of affective and cognitive variables as the only criterion for including research into my review.

2.2.1 Emotion in research

I will start this review with the recent change in attitude towards emotion in psychology. Damasio (2000), when discussing the change of the role of emotion in psychology, says:

Throughout most of the 20th century emotion was not trusted in the laboratory. Emotion was too subjective, it was said. Emotion was too elusive and vague. Emotion was at the opposite end from reason, easily the finest human ability, and reason was presumed to be entirely independent from emotion (Damasio 2000, p.40).

One could disagree with this absolute statement and point out that Bartlett (1932) in his schema theory showed how important attitude was in our thinking. Nevertheless, the overall tendency of the research is captured by Damasio with precision and refuted with vigour as he describes the role of the 20th century in the research into emotion:

This was a perverse twist on the Romantic view of humanity. Romantics placed emotion in the body and reason in the brain. 20th century science left out the body, moved emotion back into the brain, but relegated it to the lower neural strata associated with ancestors who no one worshipped. In the end not only was emotion not rational, even studying it was probably not rational (Damasio 2000, p.40).

Damasio is aware of the existence of other viewpoints, but considers them to have been neglected by 20th century science:

The notion of an integrated organism – the idea of an ensemble made up of a body proper and a nervous system was available in the works of thinkers such as L.von Bertalanffy, K.Goldstein and P.Weiss, but had little impact in shaping the standard conceptions of mind and brain (Damasio 2000, p.40).

To exemplify the neglect of these differing viewpoints Damasio quotes a 17th century French thinker, Malebranche, who saw emotions and cognition in the way we see them today:

It is through light and through clear idea that the mind sees the essence of things, numbers and extensions. It is through feeling that the mind judges the existence of creatures and that it knows of its own existence (Damasio 2000, p.313).

Similar ideas have been expressed by other psychologists and sociologists (see Scovel 1991, Oatley and Jenkins1997, LeDoux 1999, Coulter 1986), but Malenbrach, to my mind, brings out the very essence of feelings (or emotions), that of judgement which is impossible without the interaction between emotion and mind. He also says that through feeling we know of our own existence, thus proposing that feeling can be a source of information which is also important in showing the nature of the link between cognition and emotion and which nowadays is proposed as a totally new approach to investigating emotion (see Lewis 2000).

Coulter (1986) laments the neglect of emotion in sociology: *Sociology has had little to say about the nature of 'affective' or emotional conduct, perhaps primaily because, following Max Weber's lead, it has generally been hived off theoretically from the bulk of 'rational' action in human affairs, downgraded to a sort of appendage to social relations and consigned to a permanantly residual status* (Coulter 1986 p.120). To improve the situation he suggests studying emotion on a totally different basis, as a part of rationality.

Damasio also points out the intimate relationship between reasoning and the emotions: *Emotion is integral to the process of reasoning and decision making, for better and for worse. Selective reduction of emotion is at least as prejudical for rationality as excessive emotion* (Damasio 2000, p.42).

LeDoux represents the point of view of experimental psychology: *Emotions evolved not as conscious feelings, linguistically diferentiated or otherwise, but as brain states and bodily responses. The brain states and bodily responses are the fundamental facts of an emotion, and the conscious feelings are the frills that have added icing to the emotional cake* (LeDoux 1999, p.302).

On the contrary, in social psychology 'affect' has traditionally been treated as the *conscious subjective aspect of an emotion considered apart from bodily changes*

(Cacioppo, Wendi and Gardner 1999, p.848). Here we can see the crucial difference between social psychology and experimental psychology that lies in the concentration on the social aspect and the overlooking of the 'bodily changes' as if a human being did not have to come to terms with his or her body before becoming a social being.

Oatley and Jenkins consider that emotion is at the basis of our plans and actions: *The core of emotion is readiness to act and the prompting of plans: an emotion gives priority for one or a few kinds of action to which it gives urgency – so it can interrupt, or compete with alternative mental processes or actions* (Oatley and Jenkins 1976, p.96).

Damasio (2000) sees emotion as the basis of our consciousness: *The fabric of our minds and of our behaviour is woven around continuous cycles of emotions followed by feelings that become known and beget new emotions* (Damasio 2000, p. 43). He considers that consciousness is born when we become aware of a relationship between ourselves and some object. He calls this awareness 'feeling' which is a stepping-stone to 'knowing'. He even suggests that *the mechanisms which permit consciousness may have prevailed because it was useful for the organisms to know of their emotions* Damasio 2000, p. 285).

2.2.2 Production of emotions

In this section I will give a short overview of the most influential theories of the generation of emotions to provide the context and introduce the vocabulary necessary for the discussion of the topic of anxiety (for example appraisal, arousal and unconscious affect).

2.2.2.1 Drive theory

The question that dominated emotion research agenda during the 20th century, 'where do emotions come from?', was raised by James in 1884. LeDoux says that the modern era in emotion research began when James asked whether feelings cause emotional responses or responses cause feelings. James' solution to the stimulus-to-feeling sequence problem was that feedback from responses determines feelings.

This view dominated till the 1920s when Cannon introduced the concept of 'emergency reaction', *a specific physiological response of the body that accompanies any state in which physical energy must be exerted. The flow of blood is redistributed to the body areas that will be active during an emergency situation and will need more energy* (LeDoux 1999 p.45).

In 1929 Cannon proposed that human behaviour was driven by states of imbalance, which occur when basic needs are not satisfied. The brain and the body mechanisms that control respiration, heart rate and body temperature, hunger, thirst and sexual desires, called 'homeostats', place value on stimuli that allow the organism to maintain balance and enhance survival. This theory says that an organism is motivated to return to a state of balance. Later this idea was developed to explain the human need for interaction with others and social relations, the so-called 'sociostats'.

According to Schuman (1999) sociostats are the innate tendencies of the human organism to seek interaction with members of the same species. They are the drives for attachment and social affiliation. In addition to innate homeostatic and sociostatic values organisms develop highly individual value systems that make people like one thing and dislike another. They are based on the past experiences and influence cognition that is devoted to learning and other activities. Because of these

motivational differences the same stimulus can be evaluated differently by different people and the same thing can be liked by one person and hated by another.

2.2.2.2 Arousal theory

Arousal theory became popular at the beginning of the 20th century. The research of arousal started at its primary stages as trying to establish the relationship between the state of the person and his or her performance: the Yerkes-Dodson Law (1908) suggests that when the arousal is too low or too high, performance is depressed. The same was found by Davis and Harvey in 1992 when they researched the influence of arousal on sport: the major league baseball players performed less well in the closing stages if the arousal was too high (Flannagan 1954, p.288).

LeDoux (1998) says that nowadays it is possible to register the level of arousal by putting electrodes around the human scull. Electrodes pick up the electrical activity of cortical cells through the skull triggered off by chemicals released by different systems located in the brain stem. LeDoux (1998) says that the arousal level of the cortex is related to the difference between being awake and alert as opposed to drowsy or asleep. When we are paying attention to something, our cortex is aroused, when we are not focusing on anything our cortex is in an unaroused state. LeDoux considers that arousal is important in all mental processes as it contributes to attention, perception, memory, emotion and problem solving.

2.2.2.3 Appraisal theory

In 1960s Schachter and Singer suggested that bodily arousal or feedback was crucial to emotional reaction: sweaty palms, rapid heart beat, muscle tension inform the brain that a state of heightened arousal exists. However, these responses are similar in many different emotions and do not identify what kind of aroused state we are in. On the

basis of physical and social context as well as knowledge about what kind of emotions occur in these situations we label the aroused state as fear or love or shock.

According to Shachter and Singer (1962) emotional feelings result when we explain emotionally ambiguous bodily states on the basis of cognitive interpretations (called 'attributions'). Thus it is the cognitive representation of the physiological arousal, not the arousal itself, which interacts with thoughts about situation in the generation of feelings.

According to LeDoux (1999) *the real impact of this* (Schachter and Singer's) *work, though, was not so much that it explained where our emotions came from, but instead it revitalised an old notion, one that was implicit in the philosophical writings of Aristotle, Descartes and Spinoza, that emotions might be cognitive interpretations of situations (* LeDoux 1998, p.49).

The same idea was further developed by Arnold (1960). She defines appraisal as the mental assessment of the potential harm or benefit of a situation; emotion is the felt tendency towards anything appraised as good or away from anything appraised as bad. Although the appraisal process occurs unconsciously, its effects are registered in consciousness as emotional feeling.

Power and Dalgliesh (1998) criticise Schachter and Singer for limiting cognitive interpretation to a minor supporting role, that of simply labelling the aroused state and think that as a result they provide *a simplistic and inadequate cognitive basis in comparison to recent appraisal theories* (Power and Dalgliesh 1998, p.81). They consider that Mandler (1984) made the next crucial step connecting bodily arousal, cognitive attributions and ongoing goals or plans. In Mandler's (1984) theory physiological arousal is considered to arise from a perceived discrepancy or from the interruption to an ongoing plan or goal. Arousal provides the intensity of the emotional state and cognition provides its quality (Mandler 1984, p.119).

The same was acknowledged by May in 1981; when discussing the causes of human anxiety he wrote: *The distinctive quality of human anxiety arises from the fact that man is a valuing animal, who interprets his life and world in terms of symbols and meanings. It is the threat to these values, specifically, to some value that the individual holds essential to his existence as a self, that causes anxiety* (May 1981, p.241).

On the basis of appraisal theory several researchers developed their own theories. One of them was Weiner (1986) who proposed the so-called attributional theory in 1985/6 that was based on Schachter and Singer's model and their work on achievement and success or failure in the classroom. He suggested that the undifferentiated physiological arousal be replaced by two different emotional states that have a range of motivational consequences. It is not the initial affective state that provides the input for cognitive processing but the causal explanation, which is determined by whether the person perceives the cause to be

1. external or internal

2. stable or variable over time

3. controllable or uncontrollable

4. global or specific (Power and Dalgliesh 1998, p.84).

Lazarus proposed a cognitive-motivational-relational theory in 1966-91 in which he elaborated on the notion of appraisal. According to his theory there are two stages of appraisal: primary and secondary. During primary appraisal we decide whether the encounter is relevant or irrelevant, positive or stressful; we compare our goals with the situation for the goal relevance, goal congruency (enabling versus blocking goals) and ego involvement.

During secondary appraisal we evaluate our coping resources. Lazarus et al (1984) consider that the coping resources can be emotion or problem focused. If we focus on emotions, then we use defence mechanism, but if we decide to focus on the problem solving processes, then the situation can be appraised as changeable and the individual attempts to alter the situation instead of coping with the stress itself.

If we compare the two appraisal stages to the description of goal-setting and assessment strategy description in Bachman and Palmer's (1996) model of meta-cognitive strategies it is difficult not to notice parallels with the analyses of the goals in the primary appraisal and coping resource evaluation during the secondary appraisal. Evidently this is the area of interaction between cognition and emotion and the researchers of both emotion and cognition use their specific terminology to describe the same phenomena.

2.2.2.4 Unconscious affect

Zajonc (1984) proposed that cognitive appraisal of the stimulus and the bodily reaction to it does not comprise the whole experience of emotion. In his experiments he found that if subjects are exposed to some novel visual patterns and then asked to choose whether they prefer the previously exposed or the new visual patterns, they reliably choose the previously exposed patterns although they had been exposed so briefly to these that the subjects were unable to state whether they had ever seen them. According to Zajonc these results go against the common sense assumption that we must know what something is before we can determine whether we like it or not. This allowed him to conclude that emotional processing can occur in the absence of conscious awareness.

Erdelyi (1992) used a different technique but the finding was the same: subconscious processing can take place outside conscious awareness.

Bargh (1992) carried out an experiment where the subjects were given words on cards and they had to make sentences out of them. For some subjects the sentences were about elderly people, whereas other subjects received sentences on other topics. After completing the task, the subjects left the room. Unbeknowst to them, the amount of time taken to walk down the hall was timed by the experimenters. The subjects that had unscrambled the sentences about elderly people took longer to walk the same distance although there had been no reference to elderly people being slow or weak. This and other similar experiments allowed Bargh to propose that even if the meaning of the stimuli is implicit, subconscious processing influences our actions. In these cases, we do not question their influence, but try to provide some logical explanation instead. *The fact that emotions, attitudes and goals are activated automatically means that their presence in the mind and their influence on thoughts and behaviour are not questioned. They are trusted like we would trust any other kind of perception* (LeDoux 1999, p.63).

According to LeDoux, *The perception in oneself of an attitude (disguised as a fact) about a racial group can seem to be as valid as their colour of skin. When one is aware of biases and posesses values against having these, he or she can exercise control over them* (LeDoux 1999 p.63).

According to Bargh (1992), a goal of social psychology should be to make people aware of these nonintuitive, scientifically discovered unconscious factors that affect thought and behaviour.

Another of LeDoux's conclusions about subconscious processing is that the cause of an emotion can be very different from the reasons we use to explain the emotions to ourselves or others after the fact. This conclusion suggests that self report and introspection are not as reliable as we would like them to be and *the success of cognitive science is due in large measure to its ability to investigate the mind without*

relying exclusively on introspection as there are mental events that we do not have access to (LeDoux 1999, p.66) .

Unconscious affect is a striking example of how our cognition is manipulated by emotion: not only cognition follows the lead of affect, but also provides its own fake reasons for doing it, thus making it impossible for a human being to distinguish between reason and emotion guided actions.

2.2.2.5 Emotion generation via schematic models

Power and Dalgliesh (1998) consider that the production of emotions is based not only on a stimulus and our response to it, but also on our interpretation of the stimulus (which they call event). Interpretation could be seen as a projection into future how this event will affect our future. To develop such a projection, we would have to use all three forms of mental representations: analogue, propositional and schematic. I will now give a brief overview of the two basic forms of mental representation (analogue and propositional) and then turn to schema theories.

Pylyshyn (1984) proposed that experience is stored in the form of images and symbols. If someone mentions, for example, Big Ben, then we can see it in the mind's eye in colour, we can also recall the sound of traffic around it and the sight of the Houses of Parliament and lawns and the river in the background. Power and Dalgliesh consider that analogical mental representations are non-discrete, they represent things implicitely, they have loose rules of combination and they are tied to a particular sense of modality, for example, vision (also olfactory, auditory, gustatory, proprioceptive, and tactile images).

Fodor (1987), however, proposed that human experience is stored in an unspecified form of language of thought which can be articulated if necessary in words and sentences to describe the images. He called this 'propositional representation' form.

According to Power and Dalgliesh, *the propositional representations are considered to be explicit, discrete and abstract. They represent beliefs, ideas, objects, concepts and relations between them* (Power and Dalgliesh 1998, p.163). They are a sort of natural language of thoughts. They represent the ideational content of the mind.

2.2.2.5.1 Bartlett's schema theory

Bartlett (1932) introduced the third, more complex form, that of schema. He developed a series of experiments in which he tried to discover how people remember things. He presented a group of students and colleagues with a story from an unfamiliar culture and asked them to reproduce the story after different intervals of time. Bartlett discovered that when the participants in the experiment reproduced the story after some time they substituted the more unfamiliar facts with phenomena from their local culture. This allowed him to propose that memory was not storing every single fact separately, but was organised in larger units of cognitive representation (schemas). He also found that the regular mistakes were caused by the attitude of the participants to the stories and by images belonging to unfamiliar cultures. Bartlett concluded that *the organisation and activation of knowledge is crucially affected by the interests, attitudes and the previous experiences of the participants of his experiments* (Bartlett 1932 p.206-207).

2.2.2.5.2 Schank and Abelson's theory of scripts

 A variation of schema theory is script theory. Schank and Abelson (1977) proposed that we construct schematic representations of frequently encountered situations, in their example, of going to a restaurant, which they called 'scripts'. Scripts are developed by abstracting large amounts of propositional level information (what actions can be expected in a restaurant: search for a table, sitting down, ordering food); this will also include analogical information (smells, sounds and feelings about going to a restaurant). The restaurant script is activated whenever restaurant-related

information is encountered. The schematic representation is called up from long-term memory and combined with analogue and propositional information to build an online model of the current situation.

2.2.2.5.3 Eckland's interpretation of Piaget's schema theory

Eckland (1981) used Piaget's approach to schema theory to explain motivation. For Piaget, the central feature of behaviour is the intimate interaction between the schemas and the environment in which they function. This is captured in the types of adaptation: in the forms of assimilation (as the individual deals with the environment) or accommodation, the individual's change in response to the environment's needs (Eckland 1981 p.18).

A schema is assumed to contain a number of tests where input or the result of an internal operation is compared with a standard for that condition. If there is a match, operations proceed as normal; the schema is in equilibrium; if the input fails, the schema enters a state of disequilibria. This is a motivational state, when other parts of schema can be activated. If this attempt is also unsuccessful, a permanent change will gradually take place in the form of differentiation. This process is called *accommodation* and is the basis of learning.

Behaviour, thus conceived in terms of functional interaction, presupposes two essential and closely interdependent aspects: affective, because the interaction with the environment involves both structuring and valuation (cognitive and affective aspects). Schemas develop by becoming differentiated, and coordinated into larger structures. If, however, an event has not been assimilated, it will have no meaning.

Eckland considers that in spite of the fact that Piaget has not written on affects and motives, a complete conceptual framework for motivation lies implicit in Piaget's central concepts: Desirability is the indication of a rupture or of an uncompleted

totality to whose formation some element is lacking and which tends toward this element in order to realise its equilibrium.

Eckland considers that Piaget also explains the basis of expertise and familiarity of the situation that is necessary for us to feel in control (for example, during a test situation). Piaget considers that we cannot instantly feel at ease in a new situation, we need time to acquire and to process and systematize the results of experience, so that the situation becomes controllable and we experience satisfaction with the result of our effort: *No amount of combination and recombination of earlier schemas will produce the expert's eye - only slow differentiation of schemes in a field where only more or less random variation was seen before. Through the progressive differentiation, recognitory assimilation becomes possible, namely at the moment when there is a standard and input matches it. Recognition is assumed to produce positive affect as evidenced by smiles and laughter* (Eckland 1981 p.38).

Eckland in her description of learning also introduces the concept of pleasure: *New sections of the environment are continuously being fed into the system and repeated encounters will lead to differentiation, recognition and pleasure as the standards are formed in response to stimulation. Further encounters with the stimulus lead to the establishment of mastery in the sense of effortless, practiced recognition or familiarity. In a later stage larger incongruities are tolerated and eventually enjoyed and novel and more complex stimuli excite interest. As the schemas become elaborated, at least some of them stop growing and become stable and habitual. Their activation no longer claims attention* (Eckland 1981, p.40). Thus Eckland uses schema theory terminology to describe what Csikszentmihalyi (1998) calls 'flow' (see section 2.1).

Eckland also sees affect in the terms of schema:

- motive is the presence of schema disequilibria: whenever an attempt at assimilation fails and corrective attempts are not immediately successful, a motive will originate

- arousal increases with the degree of incongruity in schemas and creates anxiety. High levels of incongruity are innately aversive and therefore avoided by people. Eckland considers that the stronger the schema and the more important it is, the more arousal will be generated when incongruity between that schema and the environment is discovered.

Incongruity of the situation with the existing schema taxes the organism's resources in two ways: it occupies processing capacity and at the same time produces arousal (the greater number of schemata are activated the more attention is divided, the more difficult it gets and the more difficult it is to focus attention on the task at hand. This produces anxiety.

2.2.2.5.4 SPAARS theory

Power and Dalgiesh (1998) consider that conflicts between goals and/or the various problems involved in attaining them are central to a functional understanding of the role of emotions. According to their theory goals can be represented in all three levels of mental representations: analogue (we can crave for some taste or like one colour better than another), propositional (when we can easily articulate our goal easily, for example, I want to catch the bus), or a schematic model, which is more like a projection into the future, (like becoming an actress), which needs knowledge, ability, skill, time and effort to be achieved. The goals sometimes can be in conflict (for example, the goal of becoming an actress can be in conflict with loving chocolate and eating it as much as possible). Power and Dalgliesh consider that *emotions are goal-related processes, which configure the system in order to resolve goal-related issues. The pattern of an individual's virtual goal hierarchy is a reflection of the schematic*

model information 'in place' within the system (Power and Dalgliesh 1998, p.168). This principle is well developed in Maslow's need hierarchy (see section 2.3.4).

Power and Dalgliesh, like Eckland (1981) and Damasio (2000) consider that the highest aim of the human mind is to create a system that will keep the balance between the external and internal world: *The highest goal in the system is to preserve the current configuration of dominant schematic models of world, self and other. These models provide individuals with their sense of self and their sense of reality* (Power and Dalgliesh 1998, p.169). The reaction to a threat to the highest goals creates anxiety that makes us concentrate on the issue in order to resolve the conflict and return to a state of balance. In cases of inability to resolve the conflict depression may set in.

Figure 1 Power and Dagliesh (1998) model of emotion production
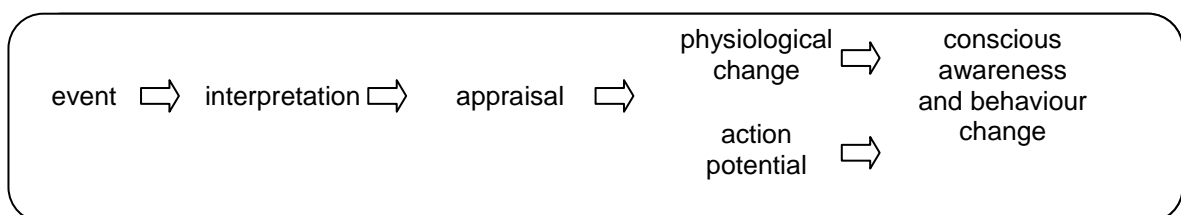
Power and Dalgliesh consider that the components of emotion as a process always involve an event, its interpretation and appraisal according to one's goals, physiological and/or behaviour change and conscious awareness of the feeling (see Figure 1). If explained in the form of mental representations, the perception of an event, for example, seeing a snake and hearing it hiss, happens at the analogue level (visual and auditory information) which is interpreted at the propositional level (snakes are dangerous) and appraised at the schematic level (the goal of survival is threatened), so emotion products and output systems are activated: physiological reaction, action propensity (to run away) and the person becomes aware of being afraid.

In this case the memory of seeing a snake is going to be stored in the schematic representation format: together with the environment, sound and sight; and whenever we find ourselves in a similar situation we will become cautious. On seeing a snake, or even something like a snake, we will not have to go through the stage of interpretation and appraisal, the emotion system will be activated automatically without our conscious awareness. Power and Dalgliesh say that *appraisal needs to have occurred at some time in the emotional history of the individal's experience of that event for the automatisation of the process of emotion generation, so that it appears as if a concurrent process of appraisal is occurring even though it is not and has in fact occurred at some time in the emotional past. In other words, the accessing of the schematic level of meaning is 'short-circuited'* (Power and Dalgliesh 1998, p.179). We think that we have reason to be angry with somebody, but it may well happen that we are just reacting because of some previous situation when we felt hurt. In the example of the snake, however, we might find that we would freeze or recoil without thinking that we had ever met a snake before, because the recent findings in psychology (see Goleman 1995, LeDoux 1999 and Power and Dalgliesh 1998, Damasio 2000) suggest that in some cases it is only necessary for an event to have occurred once in the evolutionary history of the species for it to become activated if necessary. Damasio suggests that human beings are born with memory in the form of what he calls 'dispositions' of which we are not aware, but which can be activated in certain situations (see 'affective schemata' in section 3.3.3.).

LeDoux calls these prepackaged responses automatic and involuntary. He says that many animals get through life mostly on automatic pilot reacting to a situation; humans, however, use the automatic reaction to 'buy time' for thinking and deciding the next step. In this way a human being can stop 'reacting' and start 'acting': take control, think of a plan specially made for the particular situation. However, once we

start thinking, we come up with several plans, and then we have to choose which is the best and think what will happen if it fails. As LeDoux says, *bigger brains allow better plans, but for these you pay in the currency of anxiety* (LeDoux 1998, p.177). Damasio says similarly: *Consciousness and its revelations allow us to create a better life for self and others, but the price we pay for that better life is high. It is not just the risk and danger and pain. Worse even: it is the price of knowing what pleasure is and knowing when it is missing and unattainable. Feeling of what happens is the answer to a question we never asked and it is also the coin in a Faustian bargain that we never negotiated. Nature did it for us* (Damasio 2000, p.313).

Power and Dalgliesh (2000) suggest that even in the automatic generation of emotions via the associative level, when we are retrieving the emotional system products from the longterm memory (having already once solved the situation) we are aware of both physiological change and the content of the associative representation of the interrelationships of the event and the emotion products. This is why they suggest that

Figure 2 SPAARS model of emotions Power and Dalgliesh (1998)
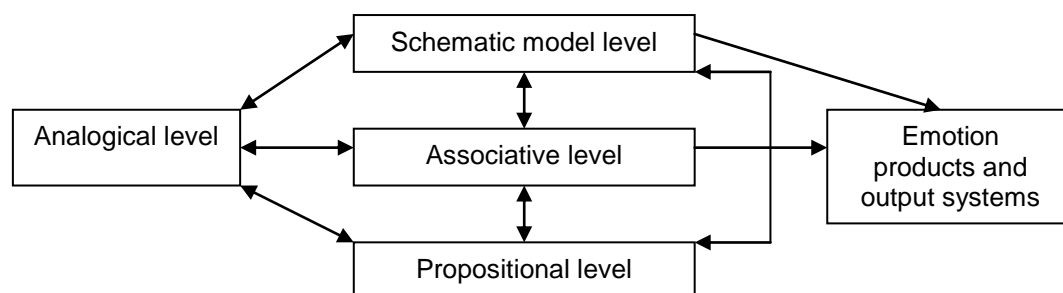
the associative level of representation is a separate level of representation apart from the analogical, propositional and schematic model levels (see Figure 2).

Power and Dalgliesh call this the SPAARS model of emotions (Schematic, Prepositional, Analogical and Associative Representation Systems). According to their model emotions are generated from either the schematic model information level (when we have interpreted and appraised the facts as important for our goals) or the

associative level (when we have already met a similar situation and are assuming that this is a similar case). Emotion generation from the schematic and associative levels is based on information from the analogical and/or propositional level. Emotion cannot be generated from the propositional level of information directly, as it is information in the form of 'facts' that are not in any way connected with our own future plans or goals.

Power and Dalgliesh have also proposed ways in which we can either control or be controlled by emotion. They say that the SPAARS model of emotion generation is organised by the currently dominating schema, therefore the model is self-organising: if the positive self-schema is active, we will concentrate on the positive information available at the analogical and propositional information levels. On the other hand, if we are aware of not only analogical information (images, smells and sounds) and propositional information (thoughts and language), but also the higher order schematic meaning these things evoke, we can restructure and re-organise the information, to inhibit some connections and facilitate some others.

To my mind the SPAARS theory allows us to explain the interaction between cognition and emotion, as it integrates different levels of information processing, both cognitive and emotional into one system. On the other hand, although Power and Dalgliesh refer to the possibility of the self-organising feature of the emotion production process, Lewis (1996) develops the idea further to explicitly show how emotions develop and change. The idea of cycles of appraisal appears also in Power and Dalgliesh, but is mainly developed for the explanation of anxiety and worry escalation, while Lewis has managed to depict the constant change and re-evaluation of situation and its interaction with thinking. The feedback arrow from emotions to propositional information level in Figure 2 is missing denoting the fact that the Power and Dalgliesh's model is not really a self-organising and developmental model, as

emotion appears only in the form of output, which is not fed back into cognition. As a result it denies the self-awareness of a person and the ability to understand and monitor one's own emotions.

This section reviewed cognitive theories on emotion. Table 1 summarises the common features of these theories: the cause is either a stimulus or an event, the end product of the process is a conscious awareness or feeling that some change has occurred within the organism. The ways in which these theories connect the stimulus with the response are different, they all contain information that supplements the others, therefore it is not possible to say that that one theory is better than another. It is clear that many systems interact in emotion production (cognitive, physiological and different nervous systems) and a theory that depicts all the systems' input will not be exclusively a theory of emotion, it will be rather a theory that explains human behaviour (like self-organisation theory, see section 2.3.7).

Table 1 Review of emotion production theories

Theory	Theorists	Components of the theory of the production of emotions			
Drive theory	Cannon 1929	stimulus	response	feedback	feeling
Arousal theory	Yerke-Dodson 1908	stimulus	arousal		feeling
Appraisal theory	Schachter & Singer 1962 Mandler 1984	stimulus	appraisal	cognition	feeling
Unconscious affect	Zajonc 1984	stimulus	unconscious affect		feeling
Schema theory	Bartlett 1932 Piaget 1978 Power and Dalgliesh 1998	event	interpretation	appraisal that leads to physiological changes	conscious awareness and behaviour change

2.2.3 Functions of emotion

The discussion of the functions of emotions is complicated because different researchers treat the term 'emotions' differently: most include motivation within the construct of emotion, but some do not (for example, Bower (1994)). Here I will not differentiate between different affective variables, but will refer to all of them all as 'emotion', just as they were used in literature.

2.2.3.1 Lewis (1996)

Lewis (1996) proposes that different researchers studied different functions of emotion:

1. Frijda (1986) studied the role of emotion in directing our attention and behaviour to where it is needed: cognative elaborations serve action readiness generated by emotions to guide adaptive function

2. Bower (1981) focused on the role of emotion in priming or retrieving of particular memories or stored knowledge

3. Isen (1984) studied the influence of emotion on cognitive organisation, attention and judgements

4. Dodge (1991) found that emotion induced selective interpretation of significant events

5. Oatley and Johnson Laird (1987) found that emotions highlighted critical junctures in plans according to underlying goals and needs.

This list of findings suggests that emotion influences our memory, cognition and perception and thus guides our behaviour to achieve our goals and satisfy our needs in the most efficient way. It suggests that emotion affects all mental processes (abstract thinking and reasoning, planning and enactment of plans as well as executive management of the process). It is interesting that the same processes are claimed by Sparrow and Davies (2000) as part of 'cognition' without any reference to the influence of 'emotion' (see section 2.1.). This creates an impression that cognition is an unconscious victim of the influence of emotion, or rather that the researchers studying cognition consciously avoid the topic of interaction between cognition and emotion (see Skehan 1998 and Purpura 1999).

2.2.3.2 Bower (1994)

Bower (1994) distinguishes between motivation and emotion and also describes their functions in relation to the cognitive system: *Motivation and emotion serve multiple functions in the cognitive system. While motivation mobilizes resources for actions, directs attention and guides execution of plans, emotions serve largely as 'commentators' reacting to the present situation, evaluating the execution of plans and their outcomes* (Bower 1994 p.305). Nevertheless, he considers that the role of emotion in cognition is diverse as emotion participates in input, storing, retrieval and the evaluation of internal and external information:

1. emotion signals to the cognitive system the important discrepancies between actual and expected plans

2. emotion directs attention to the causally significant aspects of situation

3. emotion serves to encode and index an unusual event in the memory

4. emotion promotes the persisting rehearsal of corrective actions

5. emotion arousal retrieves associated thoughts, plans and memories.

2.2.3.3 Coulter (1986) and Pritchard (1976)

Coulter (1986) also sees emotions as part of rationality. He considers that affect and rationality are much more closely interrelated than has been noted in the behavioural sciences. He refers to Pritchard (1976) who says that *our capacity to experience certain emotions is contingent upon learning to make certain kinds of appraisals and evaluations. It is learning to interpret and appraise matters in terms of norms, standards, principles and ends or goals judged desirable or undesirable, appropriate or inappropriate, reasonable or unreasonable* (Pritchard 1976, p.124, quoted in Coulter 1986).

Pritchard points to the connection between cognition, emotion and volition and their interdependence. However, he transforms emotions into purely cognitive procedures: he says that emotion is learning, appraising and evaluating function without any reference to the participation of feeling of involvement (empathy) that is associated with the bodily reaction and which differentiates emotion from non-emotion.

2.2.3.4 Rolls (1990)

Rolls' (1990) list of the functions of emotions is to my mind the most comprehensive as it unites bodily sensations, effects on cognition and memory as well as behavioural effects on an individual and society. He considers that emotion

1. elicits autonomic and endocrine responses preparing the periphery of the body (increasing heart rate, blood circulation, breathing speed)

2. makes the choice of behavioural responses more flexible

3. sends motivating signals (if positive, the organism will work to obtain them, if negative, it will work to avoid them. Emotion will also influence our decision-making mechanisms)

4. communicates emotional state enhancing the stabilty of social groups

5. creates social bonding and all forms of emotional attachment

6. evaluates events and memories

7. affects cognitive evaluation of events or memories

8. facilitates storage of memories   (Rolls 1990 p. 161-167).

2.2.3.5 Caccioppo et al (1999)

A totally different approach is taken by Caccioppo et al (1999) as they have chosen to interpret affect from the point of view of natural sciences. They consider that affective

system has been sculpted by the hammer and chisel of adaptation and natural selection to differentiate hostile from hospitable stimuli and to respond accordingly (Caccioppo et al 1999, p. 839). Therefore they propose that the affect system is composed of evaluative processors, whose main aim is to provide a subjective estimate of the current significance of the stimulus to the current goals. Subjectivity (evaluation of the events as appropriate or inappropriate for achieving one' s own goals) differs affective system from the perceptual system (seeing, hearing) whose aim is to give objective estimate of the stimulus. The task of the affective system is to make a judgement whether we should approach or withdraw from the stimulus, whether the stimulus is positive or negative and then, having made the judgement, the system of affect directs our attention, guides decision-making, stimulates learning and triggers behaviour. In this, Cacioppo et al's (1999) definition of the role of affective system agrees with Arnold's (1960) definition of appraisal as good or bad for me (see section 2.2.2).

According to Cacioppo et al the division between negative and positive affect is accepted in psychology because *the theoretical module in the affect system that computes attitudes, preferences and actions derives input from at least two specialised evaluative channels that process information in parallel: one in which threat-related* (i.e. negative) *information is derived from the flow of sensory inputs and its associations and a second in which safety and appetitive* (i. e. positive) *information is derived* (Cacioppo et al 1999). Psychologists use two arguments to support this proposal: firstly the fact that both approach and withdrawal reflexes are represented as separate motive systems in the brain, and secondly the finding that approach and withdrawal stimuli are processed at least in part in parallel.

Cacioppo et al also discuss the reaction of the affect system to complex situations when the system of affect comes up with both positive and negative evaluations of the

situation: *the antagonistic effects of the activation of both positivity and negativity are integrated into a net affective predisposition or action which can be represented as an overlying bipolar response surface* (Caccioppo et al 1999, p.842). The evaluation processors function constantly and the response surface constantly changes and this change fuels our behaviour and its predispositions. The authors report the results of Goldstein and Strube (1994) who found that exam feedback activated positivity and negativity differently: the students who scored above the mean showed an increase of positive affect, but their level of negative affect remained unchanged; the students who scored below the mean showed an increase of negative affect but no change in positive affect.

This and similar findings of other researchers made Caccioppo et al suggest that we possess a positivity 'offset', that is the motivation to approach is stronger that the motivation to avoid, hence our interest in novel experiences. To protect us in a case of danger there is a negativity bias: they draw parallels between affect and gustatory systems: human taste buds respond to sweet, salty, sour and bitter stimuli. Most can detect sweetness in approximately 1 part in 200, saltiness in 1 part in 400, sourness in 1 in 130 000, and bitterness in 1 in 2.000.000. (…) It may represent differences in the currency functions for positive and negative affects that are so pervasive that it has been termed 'negative bias' (Cacioppo et al 1999, p. 848). Evidently authors consider that the nature has equipped us with more sensitivity to sour and bitter taste as this kind of food could be more dangerous to us than sweet tasting food.

Cacioppo et al refer to studies in politics, sociology and biology where negativity bias has also been discovered. The aim of negativity bias is the survival of the species: species with a positivity offset (motivated to explore) and a negativity bias enjoy the benefits of exploratory behaviour and the self-preservative benefits of a predisposition to avoid or withdraw from threatening events (Cacioppo et al 1999, p. 849). Here we

can see the interaction between the two forces that we have inherited: to move forward to attain positive affect and to retreat from danger to avoid negative affect that warns us of danger.

However, avoidance is not the only effect of negative affect, it serves also as *a call for mental or behavioral adjustment, whereas positive emotion serves as a cue to stay on the course* (Cacioppo et al 1999). In Eckland's terminology: scheme disequilibrum and incongruity create anxiety and make us create new systems that can accommodate new information, thus promoting growth and adaptation to the environment (see section 2.2.2.5.4).

2.2.4 Emotion in the brain

Nearly every researcher who has addressed the issue of emotion has also discussed the connection between emotion and thinking. However, the quality of each theory in social sciences often depends on the researcher's ability to argue his or her point. It should be different in natural sciences. Therefore I will briefly turn to LeDoux's field of research, physiology.

LeDoux proposes that just as we have different systems of memory (conscious and subconscious) we also have different systems of emotion production. Conscious, declarative or explicit memory is mediated by the hippocampus and contains also explicit memory about emotional situations. The implicit, subconscious memory system also contains the emotion memory system; this is mediated by the amygdale and contains implicit emotional memory. In traumatic situations the memory and emotion systems (both explicit and implicit) operate in parallel, as a result some of the stimuli are preserved in the explicit, and some in the implicit memory. Later, when we are exposed to a similar situation, both the systems are reactivated.

The explicit memory system (situated in the hippocampus) will provide the information on where we were, with whom and what we were doing and, that we felt

awful (just as a fact), but the amygdale mediated system will make the muscles tense up, it will change the heart rate and blood pressure and release other bodily and brain responses. The fact that both the systems operate in parallel makes us think that emotions are subversive and interfering (as the bodily reaction interferes with our thinking capacity), although it is only one of the emotional memory systems (that of the amygdale) that can *influence ongoing perceptions (for example, visual), mental imagery, attention, short-term memory, working memory, long-term memory, as well as higher order thought processes* (LeDoux 1999, p. 287). The hippocampal emotional system meanwhile simply provides information on how we felt on our previous encounter with the phenomenon in question.

The power of the amygdale's influence has gradually developed during the evolution of our species to deal with danger and it is activated as soon as we experience fear. It is followed by the intense cortical arousal, which makes us concentrate on the source of danger. The problem, though, is that once the fear system is turned on it is difficult to control and can easily develop into anxiety because the activated amygdale keeps the cortical network in a state of hypersensitivity. This is further escalated by the feedback from the body, which is also sending messages of being in the state of arousal through the steroid and peptide hormones (and not adrenalin, as it was thought earlier) that are released by body organs and travel to the brain via blood system. This changes not only the way we feel, but also the way we look and others perceive us. If previously the researchers were mainly concerned with exploring the way our facial expressions and body postures influence our own feelings and thoughts as well as that of others, now Damasio proposes that our bodily feedback underlies also the so called "gut feelings", which play a crucial role in deciding not only our emotional feelings, but also affects our decisions.

LeDoux concludes that at present *the amygdale has a greater influence on the cortex than the cortex has on the amygdale, allowing emotional arousal to dominate and control thinking. Although thoughts can easily trigger emotions (by activating the amygdale), we are not very effective at wilfully turning off emotions (by deactivating the amygdale)* (LeDoux 1999, p. 303).

However, LeDoux considers that the situation may change. Having compared the connections between the amygdale and the cortex in different species among primates and mammals, he suggests that there is a possibility that the connections between the amygdale and cortex will continue to expand and that as a result humans might in future be better able to control their emotions (see also Damasio 2000). There is, however, another possibility - that there is an increasing number not only of fibres going from the amygdale to the cortex and but also from the cortex to the amygdale. This would increase the interaction between cognition and emotion. LeDoux says: *if these nerve pathways strike a balance, it is possible that the struggle between thought and emotion may ultimately be resolved not by dominance of emotional centres by cortical cognitions, but by a more harmonious integration of reason and passion* (LeDoux 1999 p.303). This to my mind would lead not simply to being able to control anxiety during one's performance but also to a conscious use of the additional arousal-provided energy to improve the results of our performance.

2.3 Models of Interaction between emotion and cognition

It is mostly the researchers of emotion who have investigated the interaction between emotion and cognition, because of the role the emotion plays in cognition. Damasio says *emotion assists reasoning, especially when it comes to personal and social matters involving risk and conflict. Certain levels of emotion processing probably*

point us to the sector of the decision making space where our reason can operate most efficiently (Damasio 2000, p.42).

Oatley and Jenkins (1976) consider that mediating between emotion and cognition is the main function of emotion: *the core of emotion is readiness to act and the prompting of plans: an emotion gives priority for one or a few kinds of action to which it gives urgency – so it can interrupt, or compete with alternative mental processes or actions* (Oatley and Jenkins 1976, p.96). However, the interaction was also examined by cognitive psychologists who provided the area with graphical models. These will be examined in the following section.

2.3.1 Domains of the mind and goals

I will present here those theories of the mind that attempt to connect affect and cognition to see how the subject matter of the representations that people hold might be organised, how these representations are instantiated and in what format they occur.

Power and Dalgliesh (1998) consider that the mind is best conceptualised as a functional, goal-directed system and that it is the role of the emotions to enable a person to switch from one goal to another.

Figure 3 Power and Dalgliesh (1998) Domains of knowledge
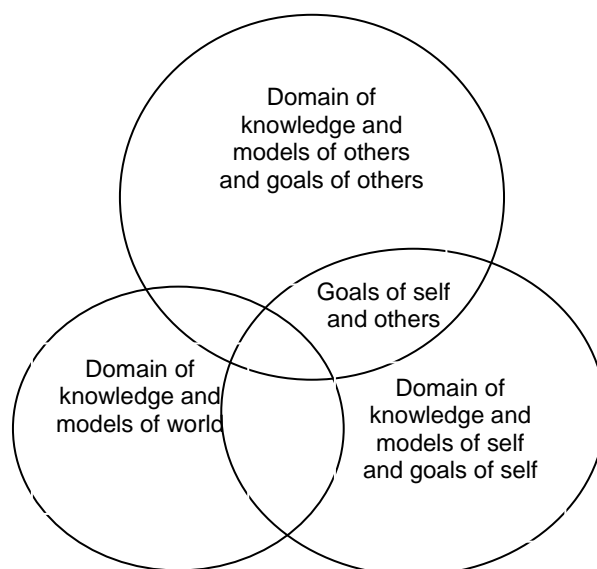
Therefore, before discussing emotions Power and Dalgliesh address the domains of mind. The graphical depiction of their theory of mind can be seen in Figure 3.

The main domains of the mind in their model are:

1. domain of knowledge and models of the world: semantic knowledge about the world, knowledge of the physical world and its objects and their interrelations and the individual's views of the world, such as 'the world is a reasonably safe place', or 'the world is just';

2. domain of knowledge and models of the self: semantic knowledge of ourselves and our capabilities, personal memories and abstracted models of the self such as 'I am a successful person';

3. domain of knowledge and models of others: streightforward semantic knowledge about other people, episodic memories and abstract models of others and stereotypes about other people.

Power and Dalgliesh propose that subsumed within the domains of knowledge and models of self and others is information which they call 'goals'. *Goals are a way of talking about the temporal dimension of representations and plans which the individual operates with* (Power and Dalgliesh 1998 p.159).

Power and Dalgliesh consider that *it is the conflicts between goals and /or the various types of other problems involved in attaining*

them which are central to a functional understanding of the role of emotions (Power and Dalgliesh p.159). May (1979) proposes a similar idea for the basis of anxiety development (see section 4.1.1.).

Power and Dalgliesh's (1998) goal concept differentiates between:

1. goals of self

2. goals of others: essential in the construction of models of others used during social interaction

3. goals of self and others: social standards and values.

The distinction between goals of self and others and common goals is important in a language testing situation: does the test-taker consider that the test-developers have common goals with the test-takers and wish them to demonstrate their best ability to use language, or that their goal is the opposite: to find out how bad they are at the language in question. How this is seen by the test-taker may influence the way the test-taker feels about the test.

If we compare the Power and Dalgliesh model of the domains of mind with Bachman and Palmer's (1996) model of language use (see section 3.2.3.), we notice that both models have features in common, as we cannot talk about production of language without mind. Apart from 'language knowledge' the two models contain some differences that I would like to comment upon:

1. Bachman and Palmer's model has 'personal characteristics' instead of 'knowledge of self'. 'Personal characteristics' contains one's age, sex, cultural background as well as one's type of personality as objective characteristics. What Power and Dalgliesh model suggests is that what matters is not what my age or sex is, but what I know or think of myself, as a person of a definite age or sex or nationality. This approach contains not just the facts, but also one's attitude to the facts and person's interpretation of the facts. The question remains open though, whether our language performance is more affected by our real age or how old we feel.

2. In Power and Dalgliesh's model our goals are part of our knowledge of ourselves which seems to be more credible than goals as a part of general strategic competence.

3. The connection between knowledge of the self and others and knowledge of the world as they are showed in Power and Dalgliesh's model imply a closer connection and even an overlap between the three which seems to be more realistic than Bachman and Palmer's (1996) model's separate free-floating circles of topical knowledge and language knowledge that are not even connected directly by arrows.

4. The connection between goals and planning and their interaction with knowledge is similar in both models, but Bachman and Palmer's model offers a much more developed interaction pattern between the two and uses also the concept of 'assessment strategies'. Power and Dalgliesh say that they do not want to use the terms like 'evaluation' as it seems confusing, they prefer the use of 'interpretation' and 'appraisal' (Power and Dalgliesh 1998, p.50). None of the terms, however, are present in this model as they consider 'appraisal' to be part of the emotion production mechanism (see Table 1).

2.3.2 Conative constructs

Snow and Jackson (1994) propose that the human mind consists of affective, cognitive and conative domains, although they are mostly concerned with conative constructs which are similar to the concept of goals: *that aspect of mental process or behaviour by which it tends to develop* (itself) *into something else; an intrinsic* (inner) *'unrest' of the organism… almost the opposite of homeostasis. A conscious tendency to act, a conscious striving… Impulse, desire, volition, purposive striving all*
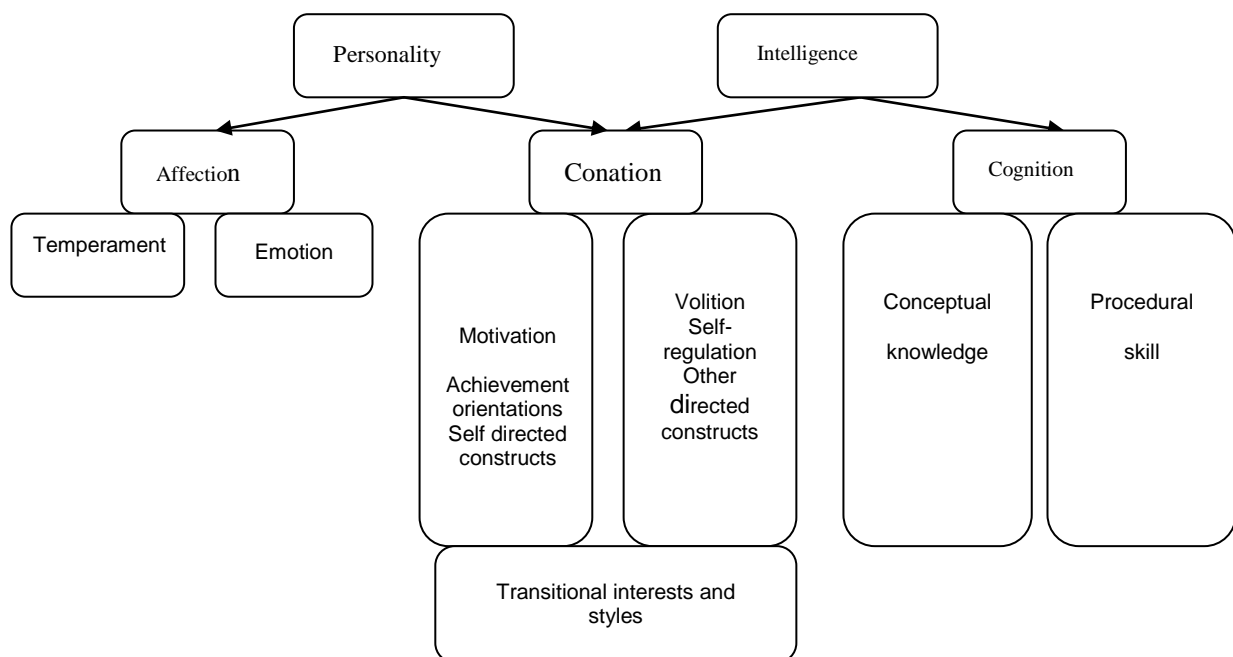
Figure 4 Provisional taxonomy of conative constructs, Snow and Jackson (1994)

emphasize the conative aspect (Snow and Jackson 1994 p.72).

This definition of goals is the opposite of Power and Dalgliesh's view of goals as a part of knowledge. Snow and Jackson describe goals using terms such as 'impulse' and 'desire', which had already been used by Spinoza and Thomas Aquinas to describe the lower levels of emotion. Their view of the concept of goals as a connection between affect and cognition is also depicted in their model of mind (see Figure 4) and their provisional taxonomy of conative constructs. This however, is in contradiction with the concepts they place in the same construct of goals:

1. *several kinds of achievement motivation, distinctions including need for achievement and fear of failure, but also various beliefs about one's own abilities and their use of feelings, of self-esteem, and self-efficacy and attitudes and interests concerning particular subject matter, learning;*

2. *volitional aspects pertaining to persistance, academic work ethic, will to learn, mental effort, investment and mindfulness in learning intentional constructs reflecting control or regulation of actions leading towards chosen goals, attitudes toward the future, and self awareness about proximal and distal goals and consequences;*

3. *and many kinds of learning styles and strategies, hypothesised to influence cognitive processes and outcomes of instruction* (Snow and Jackson 1994, p.72).

For example, the third group of concepts have to do with learning strategies and the second group (see above) which concerns itself with control and regulation is in its sense close to meta-cognitive competence. No wonder Snow and Jackson see conative domain as 'located' in some sense between affect and cognition. It is a kind of commitment pathway from wishes to wants to intentions to goals, which is well explained by the choice of the word 'conative', an adjective form of the word 'conation', which according to the Concise Oxford Dictionary (1987) is *a desire to perform an action, volition, voluntary action.* Snow and Jackson's theory as well as the use of the word 'conation' to my mind indicates the intimate relationship between 'goals' and 'affect'. Focusing mainly on conative area of the mind, Snow and Jackson have only outlined the other domains of mind of their model and it is difficult to know what exactly they meant by 'affection' (see Figure 4), although one could suppose an opposition of 'temperament', a constant feature of a person versus 'emotion' as a reaction to temporary reaction to a situational demands.

2.3.3 Goal setting theory

A totally different approach is developed by Locke and Latham (1994). They do not see goals as a part of knowledge (like Power and Dalgliesh 1998), or as a part of conation (as Snow and Jackson 1994). They see goals as a part of motivation and consider that the basis of 'goal setting theory' is the same as that of motivation theory:

- Drive Theory: Hull (1952) proposed that motivation stemmed from physiological need deprivation which drove organisms to engage in random activity

- Reinforcement theory: Skinner (1953) that eliminated the internal drive state and references to consciousness, but introduced reinforcement control over behaviour

- Subconscious motive theory: McLelland (1961) stressed the subconscious knowledge, values, motives, beliefs, memories, that a person is not aware of. He considered that subconscious motives (such as achievement) regulate human action over the long term.

'Goal setting theory' according to Locke and Latham (1994) is *based on the premise that much human action is purposeful, in that it is directed by conscious goals and that human beings possess the highest form of consciousness, the capacity to reason. They have the power to choose their own goals and pursue long-range purposes* (Locke and Latham 1994, p.14).

Locke and Latham studied the question of why some people perform better than others. If knowledge and ability are controlled for, the answer must lie in motivation. They consider that the difference between the strong and the weak performers can be most directly explained by different attributes of their performance goals: content, intensity, specificity and difficulty. Locke and Latham focused on goal specificity and difficulty in particular because performance is linearly related to the goal level: *given sufficient ability and a high comittment to the goal, the harder the goal the better the*

performance. This is because people adjust their effort to the difficulty of the task undertaken (Locke and Latham 1994, p.15).

Locke and Latham consider that this linear function takes into account our ability to adjust and is more revealing than the achievement motivation interpretation that predicts a performance drop at the highest level of difficulty, the so called inverse U function. The 'goal setting theory' predicts that:

1. *a performance drop at high goal difficulty levels occurs only if there is a large decrease in goal commitment (or a poor strategy use)*

2. *goals that are both specific and difficult lead to higher performance than vague but challenging goals*

3. *goal specificity affects the variability of performance (*Locke and Latham 1994, p.15-16).

The 'goal setting theory' is concerned with the strength of goals and the effort we are prepared to put into achieving them. This makes goal a link between knowledge (or values) and behaviour. If Snow and Jackson considered that 'conation' was a link between 'cognition' and 'affection', then here we have goals as knowledge compared with their impact on our performance. If we want to improve our performance, we have to examine our goals, are they of appropriate level of difficulty, are they specific enough, are we using the right strategies. Therefore 'affect' as a concept is not necessary in 'goal setting theory', as Locke and Latham consider that it is reason that decides which goals will be pursued; this differs radically from Power and Dalgliesh's (1998) view that the main function of emotion is to switch from one goal to another.

Although the 'goal setting theory' connects goals with the strategies we use to achieve them, it does not consider goal-setting as a strategy as Bachman and Palmer (1996) do. Instead, they say that humans possess the 'highest form of consciousness', and

they can choose goals according to their 'content, intensity, specificity and difficulty'. It seems that the goals are in the form of unchanging facts and we always know what goals we are pursuing. This does not agrree with Power and Dalgliesh's (1998) views that we possess a variety of goals, some conscious, some unconscious, some more important than others.

Locke and Latham also contradict Maslow's 'need hierarchy', that proposes that goals are organised in a certain hierarchy. I will discuss this in the next section.

2.3.4 Need hierarchy

The popular theory of 'need hierarchy' was developed by Maslow (1968). He developed a hierarchy of physiological, psychological and social needs.
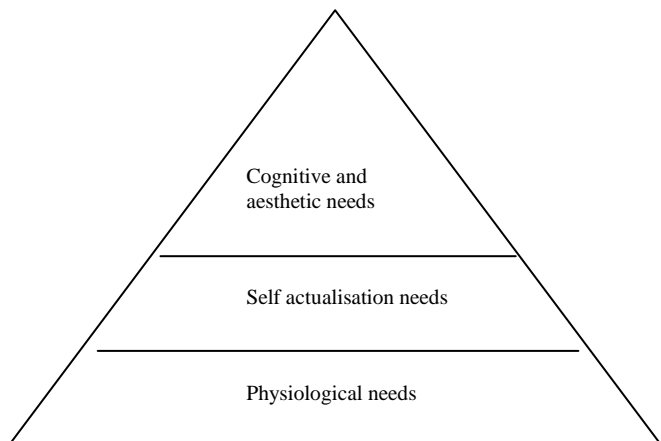
Figure 5 Maslow's (1968) hierarchy of needs

Hierarchy of needs is cumulative - a person who does not possess fundamental physiological needs (food, drink, and shelter) and intermediate or deficiency needs (safety, belongingness, and esteem) cannot possess meta-needs (cognitive, aesthetic and self-actualisation) (see Figure 5).

Translated into 'goal setting theory' it means that we formulate specific goals and spend effort on the goals depending on the level of needs certain goals belong to. If the physiological goals are achieved and we have enough food, drink and shelter, we start valuing social goals and are ready to spend effort on providing safety, belonging and esteem. If, however, the goal belongs to the self-actualisation and the meta-needs are not satisfied, we can formulate it in the most specific way, and even if it is not difficult, we will not be prepared to spend effort on it because Maslow says that the needs are pre-potent: each level is less powerful than its predecessor.

Although there are exceptions when people who do not have enough money for food still buy tickets to go to a concert or buy books, Maslow's theory of a need hierarchy has to be taken into account when explaining how we set our goals and try to achieve them. For example, if the need of prestige or self-esteem is perceived as endangered, as during a test situation, test-takers are unlikely to be preoccupied with their cognitive or self-actualisation needs.

2.3.5 Heider's attribution scheme

Heider's (1958) Attribution scheme theory is different as it does not consider that our success depends on our own characteristics only. He introduces 'environmental force' as an active partner in deciding behavioural outcome. 'Environmental force' is determined by the 'task difficulty' and 'luck'. One may object that it is nobody else but ourselves who chooses whether to take up the task or not, but it is difficult to deny the impact of task difficulty level on the 'behavioural outcome'.

Heider's theory differentiates between 'intention' and 'exertion' which is important

```
┌──────────────────┐                    ┌──────────────────────┐
│  Personal force  │                    │  Environmental force │
└──────────────────┘                    └──────────────────────┘

┌──────────┬──────────┬──────────┐   ┌──────────┬──────────┐
│Intention │ Exertion │ Ability  │   │   Task   │   Luck   │
└──────────┴──────────┴──────────┘   └──────────┴──────────┘
      │ Trying │          │ Can do │
              └──────────┬──────────┘
                  ┌──────────────┐
                  │ Behavioural  │
                  │   outcome    │
                  └──────────────┘
```
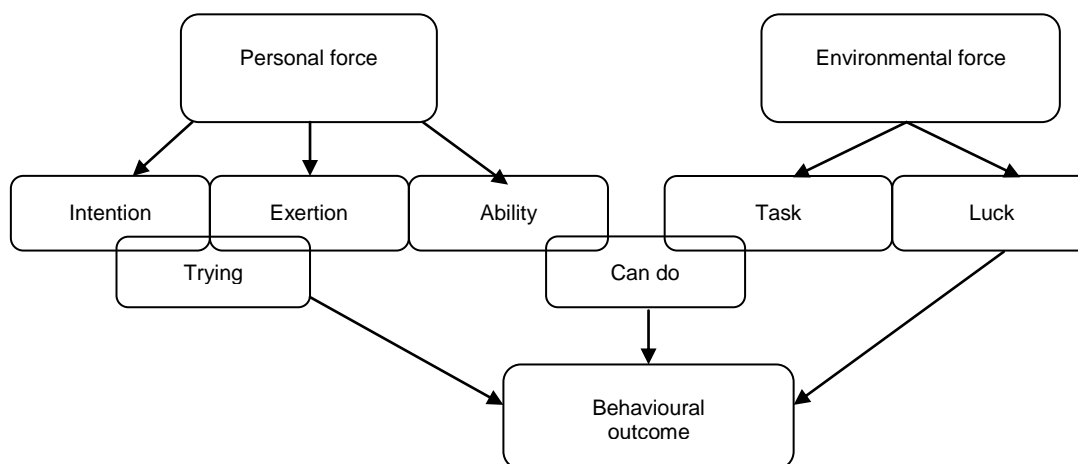
Figure 6 Heider's (1958) Attribution scheme

for language learning as well as language testing: it is not enough to set a goal, one has to exert an effort. He considers that the outcome of any behaviour is attributable to some combination of personal characteristics (ability and trying) and environmental contingencies (task difficulty and luck). This can be seen in his Attribution scheme (see Figure 6).

Heider considers that the relationship between personal and impersonal components is additive: an outcome is due to personal force plus environmental force. The main component of environmental force is what Heider calls 'task difficulty'. If a person succeeds on a very difficult task, he or she must have substantial ability.

Besides the stable property of task difficulty Heider considers that environment includes a variable factor that can affect the outcome, which he calls 'luck'. He says that luck can change the course of action in an unsystematic fashion. This is similar to Bachman's (1990) 'random factors' that affect the reliability of language proficiency measurement.

2.3.6 Self-as-agent and task performance

Combs and Merzano (1990) developed a flowchart that not only incorporates cognition, affect and goals but also shows their interaction with a task (Figure 7). It differentiates between self-goals, interpretation of the goals of others and the
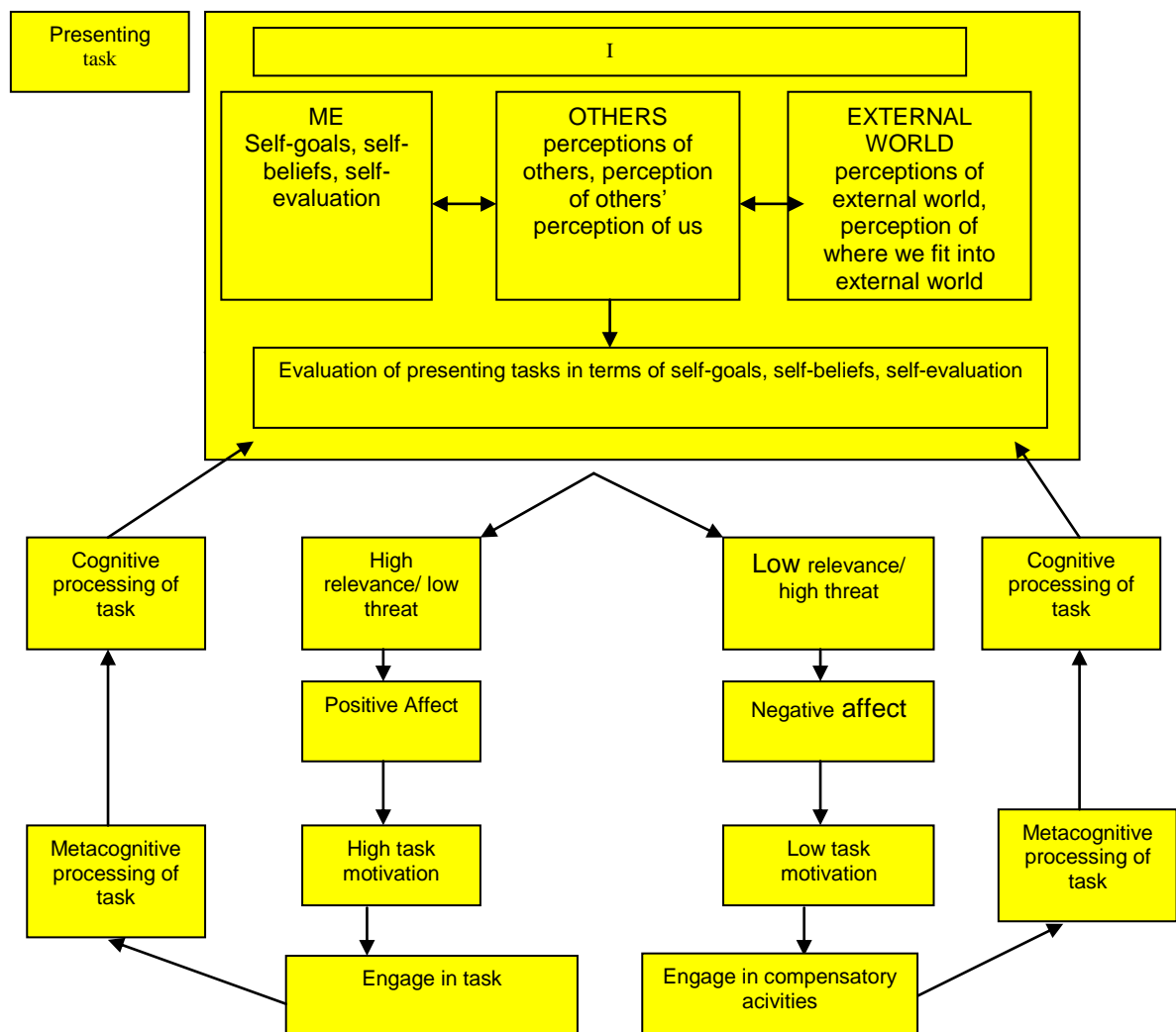
Figure 7 Self-as-agent processing framework (Combs and Merzano 1990)

perceptions of the external world.

Combs and Merzano's objective is to show how a person actively creates and construes his or her personal realities and controls his or her thoughts and cognitons in relation to a task presented taking into account one's goals, beliefs and self-evaluation. It is interesting to note that after the assessment of the match between the task demands and our perception of our own abilities, the next stage is our affective reaction which determines our motivation and once we have accomplished metacognitive processing of the task we can proceed to cognitive processing, which again is evaluated in terms of self goals.

Combs and Merzano differentiate between 'meta-cognition' and 'evaluation of presenting tasks in terms of self-goals, self-beliefs and self-evaluation' and their understanding of 'meta-cognitive processing' differs from Bachman and Palmer's (1996) (see section 3.2.4).

Although Combs and Merzano (1990) do not mention anxiety that is created if an individual has to deal with a task in what they call a 'high threat' situation, they talk about 'negative affect', low task motivation and as a result, engagement in compensatory activities, there is a certain similarity between what Spielberger et al (1978) call 'worry' and cognitions, thoughts about failure, and what Madsen (1982) calls 'facilitating' and 'debilitating' anxiety. What is not clear, however, is:

1. what if the task is of high relevance to me, but at the same time of high threat? will the affect be positive (facilitating) or negative (debilitating)?

2. what about task difficulty and my ablity level, are they of any importance in deciding the positivity or negativity of affect?

The task difficulty issue is adressed partly by Combs and Marzano's view that *evaluation of tasks solely from the level of the cognitive self-system results in aversive*

emotions and unmotivated behaviour (Combs and Marzano 1990, p.59), but evaluation of task difficulty only from the point of affect seems to me to be as biased.

Nevertheless, the fact that this model treats affect as a part of performance, which interacts with motivation and also our cognitive and meta-cognitive processing seems to make this model more comprehensive and systematic than the previous ones. Another important feature of the model is that affect as depicted in this model can also change during our performance from negative into positive, depending on the results of one's progress (see Figure 7). The idea of affect, or emotion, as a process leading to decision-making is further developed in self-organisation theory.

2.3.7 Self-organisation theory

Self-organisation theory started off as an emotion production theory, but gradually developed into a theory that explains the interaction of not only cognitive and emotional, but also motivational, execution and monitoring systems. The self-organisation theory is developing fast and is attracting new theorists (see Lewis and Granic 2000).

Here I will present the views of Lewis and Granic (2000), Izard et al (2000) and Scherer (2000) as their findings on the constant interaction between cognition and emotion offer an explanation for the interaction between the test-takers' perceived level of performance and their level of anxiety.

2.3.7.1 Lewis and Granic's dynamic system's model

Frijda (1993) and Lewis (1996) propose a 'self-organisation theory' that connects cognition and emotion in an ongoing process of interaction.

Lewis considers that when cognitive appraisal elicits emotions they generate further cognitive activity. This cognitive activity is a continuation or elaboration of the initial

appraisal event, and it is guided by emotion toward the most relevant features of the situation. This elaborated appraisal, in turn influences subsequent emotion. *The net result is that cognitive activity continually influences an emotional state, which in turn guides the stream of activity* (Lewis, 1996, p.10).

The self-organising appraisal theory allows us to explain the constant change of the emotional state as well as provides a mechanism, which explains how emotions influence and guide thinking constantly, not just as a one-off event. It also proposes a way in which cognition can influence emotion through appraisals before and after emotion output thus making use of the information acquired through perception as well as through one's own bodily reaction.

Self-organisation theory takes into account the fact that in spite of our attempts to predict and control emotional development *it is indeterminate and malleable at almost any age* (Lewis and Granic 2000, p.1) Lewis and Granic explain this by drawing parallels with systems in nature, which are also characterised by interactions among many components (they can be molecules in fluid, cells in a body or brain or individuals in society). The interactions among the elements are reciprocal and they recur over time as they gradually become more stable.

Lewis and Granic propose that cause-effect relations take the form of feedback loops, effects grow or shrink due to the activity of the system itself, not only because of the information received from the environment. At higher levels new forms of organisation can develop through the spontaneous coordination of system components' interaction. Lewis and Granic call this self-organisation in the service of a particular function, task or goal. Thus personality development can be modelled as the consolidation of cognition-emotion interactions that self organize across occasions. This can explain, for example, the development of test-anxiety or classroom anxiety as a trait of personality.

2.3.7.2 Izard et al's process model of personality development

Izard, Ackerman, Schoff and Fine (2000) also consider that *personality development* emerges through interactions of emotions and cognition and their linkage in affective cognitive structure. They see it as a process consisting of the following steps:

1. recursion among system elements: interactions among the elements of a system in the form of positive and negative feedback

2. emergence of unique forms and patterns: affective-cognitive structures that represent the preferred solutions to environmental or organismic influences, which Izard et al (2000) call 'attractors'

3. consolidation of the forms over repetition and time: the attractors become more elaborate and stronger as similar solutions repeat over time

4. constraints on system performance: the set of the possible attractors is finite, their number is constrained by a person's physiological reactivity, initial organisation of system elements and task demands.

Izard et al consider that the theory of dynamic systems can explain both positive and negative behaviours, thus attractors can be also maladaptive as a result of high stress situations; proneness to negative emotionality can contribute to consolidation of *deviant emotion-cognition-action sequences* (Izard et al 2000, p.17), for example, development of anxiety as a trait.

2.3.7.3 Scherer's model of emotion production

Scherer (2000) proposes a component process model of emotions to account for the changes that emotions introduce in a human being (see Table 2 below).
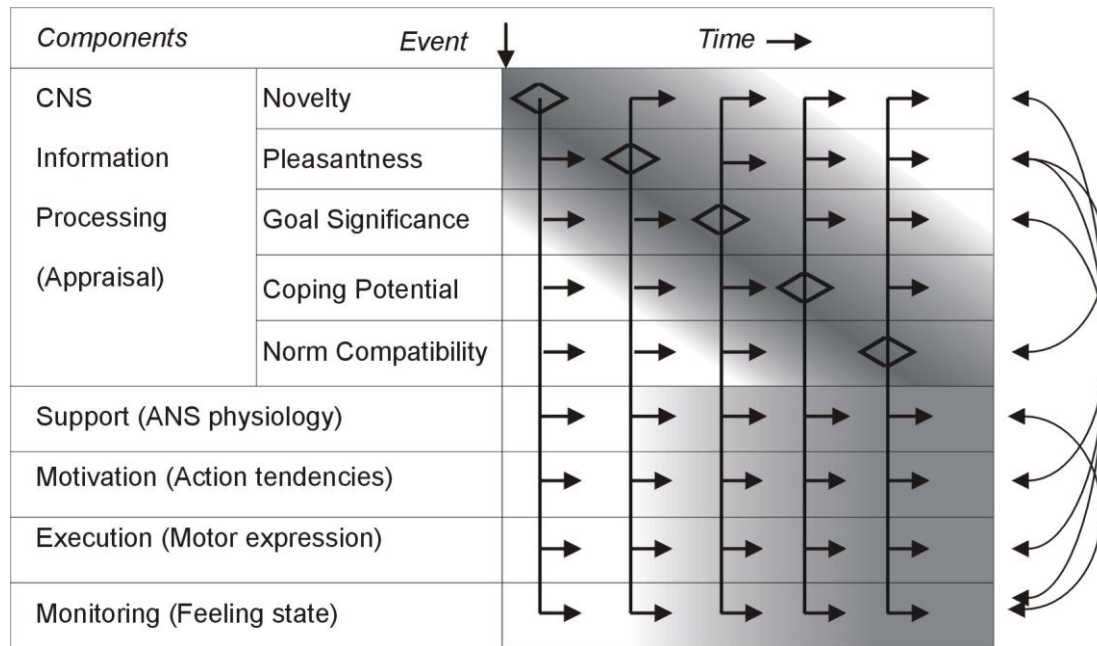
Table 2 The component process model of emotions (Scherer 2000)

This model (Table 2) represents the cognitive system (appraisal that is made up of goal setting and assessment areas), the autonomic nervous system (arousal), the motor system (expression), the motivations system (action tendencies) and also the monitor system (feeling) change in time. Scherer (2000) considers that it is the appraisal that takes place first. The person decides whether the event is a novel experience, whether it is pleasant or not, how it matches the existent goals, whether the coping potential is adequate (c.f. Bachman and Palmer's 1996 'assessment strategies') and how it complies with the existent cultural norms. The information from the central nervous system is reacted to by the autonomic nervous system, it affects action tendencies, facial expression and feeling state when the person becomes aware of the changes that he or she is going through and can start consciously to react. On the other hand, Scherer considers that the appraisal is affected by the feedback from the overall system (physiological changes produce increasing arousal which influences motivational system, affects the deployment of attention, perception and judgement).

Scherer proposes that because different nervous systems (central, somatic and autonomous) do not respond to stimulation in a uniform manner, they are often difficult to register, especially if we use linear Pearson correlations for their detection. He says that complex co-variation should be used to detect lagged synchronisation and non-linearity of the interaction between the systems. He considers that emotion should be studied as processes of self-organisation among neuro-physiological systems that are mapped onto cultural meaning structures (Scherer 2000, p.80). Thus he connects not only different individual systems (cognitive, physiological and affective), but also cultural systems that are all interacting and influencing each other.

2.3.8 Damasio's Behavioural score

Damasio (2000) proposes a process model to describe the interaction between the different systems that are necessary for language production. His model can be compared to a music score written for seven different instruments. The three lower lines, that represent wakefulness, background emotion (well-being or malaise, calm or tension) and low-level attention, are active throughout the day while the person is awake. These represent core consciousness – *the subtle, fleeting feeling of knowing, constructed anew in each pulse* (Damasio 2000, p.196). Without core consciousness extended consciousness is impossible. We can see that emotions for already part of core consciousness (see Figure 8).

Figure 8 Behavioural score (Damasio 2000)

The four upper lines represent focused attention, specific emotions (embarrassment, jealousy, guilt or pride), particular sequences of actions and verbal reports; these form what Damasio calls, extended consciousness. Just like Maslow's hierarchy of needs, Damasio's behavioural score is cumulative. Without focused attention we cannot identify our emotions, choose actions to take and produce verbal report. Damasio says that extended consciousness has become possible to us thanks to our ability to learn and thus retain records of a myriad of experiences and our ability to reactivate these so that they can also create a sense of self-knowing in the form of personal memories. This, however, is only a half of the whole picture, because a similar 'orchestral score' in the form of mental stream of images is present in the private mind, as the person becomes aware of his or her wakefulness, need to pay attention, feelings of emotions and streams of words denoting events (either real or remembered). Apart from the seven lines, there is also the inner sense of self that informs the mind of the existence of the individual. Its presence can be detected by the observer from the influence the person exerts on external behaviours (its appropriacy to the situation and its purposefulness over extended periods of time).

I think that Damasio's model, although complex is also explicit as it makes us aware of the interaction of the different systems, at the same time showing the objective (observable) existence and our own subjective interpretation of the different systems and their products. This is especially important in language production, where we produce language to translate our thoughts into sounds or written symbols that are sometimes interpreted in a totally different and unexpected way. Every language tester can relate a situation when a task designed to evoke one kind of response is reacted to by test-takers in a totally unexpected way.

Although Damasio uses 'attention' instead of 'cognition', I think that his model, just like Scherer's (2000) model and other models of mind presented in this section

suggest an interaction between cognitive and affective variables, especially so in the area of goals :

1. Power and Dalgliesh's (1998) model suggests that what matters is not only what we are, but also what we think we are and what we think our goals are

2. Snow and Jackson's (1994) model suggests that goals should be viewed as mediators between affective and cognitive domain

3. Locke and Latham (1994) propose that goals should be seen as motivation

4. Maslow's (1968) need hierachy suggests that goals differ not only in their quality but also in their strength

5. Heider's (1958) Attribution scheme distinguishes between intention and exertion (our goals versus our actions) as opposed to task difficulty and luck

6. Scherer's (2000) component process model shows how emotions are born as a result of interaction between cognitive, physiological, motivational and motor systems

7. Damasio's (2000) process model shows us the mechanism how the different levels of systems ensure the continuity of the self thus creating an individual being who can control his or her attention, is aware of his or her emotions and plans accordingly his or her behaviour.

I will return to these suggestions after I have examined interaction between affect and metacognition in the research in applied linguistics in Chapter 3.

Chapter 3   Interaction between meta-cognition and affect in language use

Chapter 3 will focus on the concepts of meta-cognition and affect as well as models that depict the interaction between cognitive and affective variables in language use. Applied linguistics uses such concepts as meta-cognition, strategic competence and affect (but not emotion), to research their effect on learner's or test-taker's performance during language acquisition or use. At present in applied linguistics research we can see two opposing views concerning test-taker characteristics: affect as a part of language use framework (Bachman and Palmer (1996) and affect as a source of error (Davies et al, 1999 and Bachman, 1990). I think that the latter view is rooted in the construct of 'individual characteristics' and its effect on language performance. Therefore before examining the role of meta-cognition and affect in language use models I will briefly examine the construct of 'individual characteristics' and the place of affect in it.

3.1   Individual characteristics in language use models

Izard et al (2000) propose that individual differences are rooted in neurohormonal (genetic processes), sensimotor, affective and cognitive systems and by studying these we would be able to understand both similarities and differences between individuals.

 A different approach is proposed by Skehan (1991) who says that an individual can be understood only as an individual *without the straitjacket of other people's categories* (Skehan 1991 p. 293, quoted in Kunnan 1995). I quite agree that it is often the case that an individual as a whole should not be classified and pigeonholed because every individual is unique, but, at the same time, we can consider individual features as categories and try and understand the role of each in the composition of the whole. This hopefully leads to a better understanding of each individual.

If we look at each individual as a being that cannot be classified, then we have to admit that language acquisition is also an individual process. We are each learning language for our own purposes, thus selecting structures and vocabulary that suit our aims, building language potential that is highly individual and developing a language potential that is tailored to satisfy our own needs. Furthermore, we are not only using language knowledge that is taught us, we are also developing new areas of language knowledge, creating new words that are influenced by our mother-tongue, our needs and our language potential. Thus language can be considered as one of individual differences.

According to Bachman (1990) when we test somebody's competence in social sciences, we try to quantify person's mental characteristics, such as aptitude,

Attitudes and orientation → Motivation

Aptitude → Proficiency → Self-identity → Additive bilingualism / Subtractive bilingualism
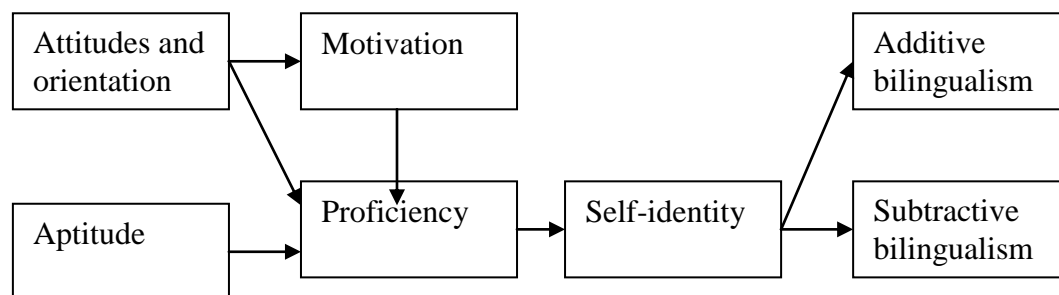
Figure 9 Lambert's (1974) social psychology model of bilingual proficiency (from Bourhis 1990)

intelligence, motivation, attitude, native language, fluency in speaking and achievement in reading comprehension (Bachman 1990, p.19). Here we can see that language skill is treated as one of a person's mental characteristics, therefore I think just as we are studying and categorising test-taker's language ability, psychologists have to study and classify other individual characteristics so that we can understand better the role of each feature in language acquisition and use.

Language proficiency models also suggest an intimate relationship between language proficiency and personal characteristics. Lambert (1974) says that the development of proficiency in a second language has important implications for an individual's self-identity (quoted in Bourhis 1990).

In Lambert's model (see Figure 9) we see how the linguistic and individual features interact and we cannot say that one is a cause and the other an effect, they define each other, they are both the source and the result: attitudes feed into proficiency, which in its turn feeds into self-identity.
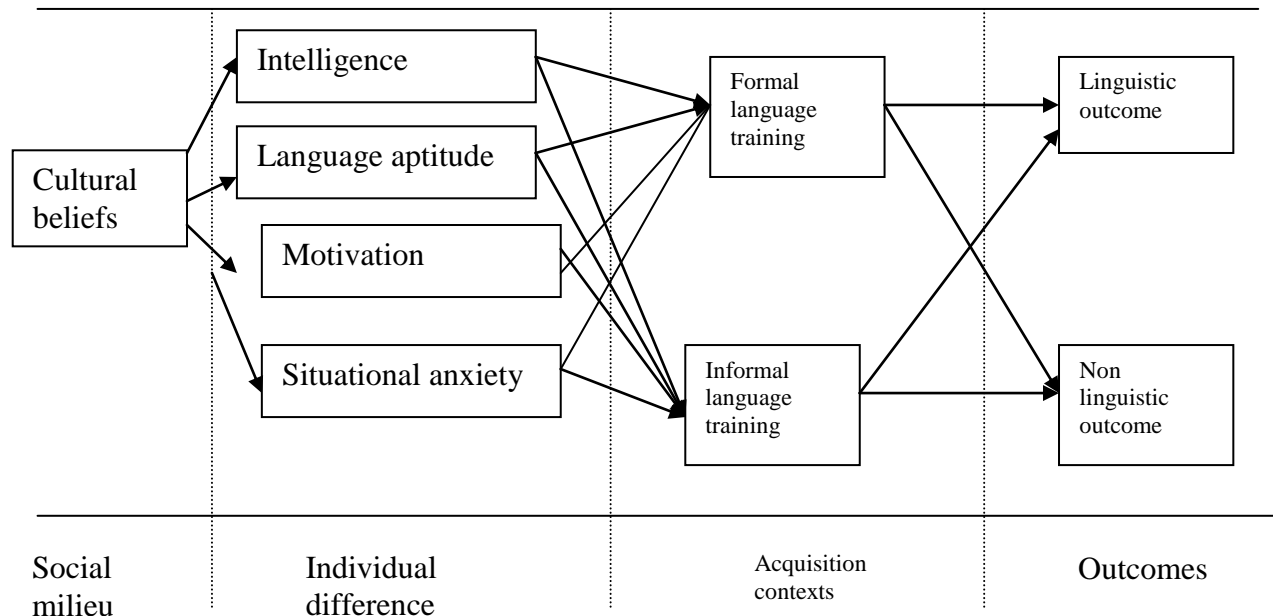
Figure 10 Gardner's (1983) socio-educational model of bilingual proficiency

Although Lambert calls his a 'model of bilingual proficiency', it can also refer to any person who has acquired a foreign language at a high level of proficiency. It is not possible to be able to adequately understand and interact with a speaker of a foreign language without acquiring also the socio-cultural competence, which means not only acquiring new knowledge, but also developing changes in behaviour, and attitudes.

Gardner's (1983) model of bilingual proficiency (see Figure 10) suggests that individual differences are defined by intelligence, language aptitude and two affective variables (motivation and situational anxiety). These affect language acquisition (both formal and informal). Gardner also considers that apart from the linguistic outcome there is also non-linguistic outcome, which in this model unfortunately, is not specified.

However, to my mind the non-linguistic outcomes should be linked to the 'cultural beliefs' to show their growth as a result of the acquisition process. These changes would further affect intelligence, language aptitude and affect and would feed into the process of language learning. I think that while we are using language we are always learning new things, even if we have passed through formal acquisition period. This would also agree with Lewis' (1996) proposed dynamic systems approach to the interaction between cognition and emotion systems.

Apart from affective variables Gardner's (1983) model of individual differences contains also language aptitude and intelligence. Oller (1983) draws parallels between intelligence and a general factor of language proficiency, *the pragmatic mapping of utterance forms into the contexts of experience* (Oller 1983, p.355). Bachman (1990) is of a similar opinion; he considers that language competence contains strategic competence: *a general ability which enables an individual to make the most effective use of available abilities in carrying out a given task, whether the task be related to communicative language use or to non-verbal tasks such as creating a musical composition, painting or solving mathematical equations* (Bachman 1990, p.106). In 1996 Bachman and Palmer reformulate strategic competence as 'meta-cognitive strategies' thus making use of a well-established concept.

3.2   *Meta-cognition*

In linguistics the meaning of the concept of 'meta-cognition' is different from that of social psychology. Meta-cognition is seen as an active force: O'Malley et al had already proposed in 1985 that students without meta-cognitive approaches were essentially learners without direction and without the ability to review their progress, accomplishments and future learning directions (O'Malley et al 1985, p.24). In 1996 Bachman and Palmer's model (1996) extended the notion of meta-cognition to

explicitly include a goal-setting component (as well as planning and assessment components) thus making the difference between the meaning of the concepts in applied linguistics and social psychology explicit. In this section I will review the development of the meta-cognitive strategy notion and its use in communicative language use models.

3.2.1 Definition of the concept

The views of the concept of meta-cognition among theorists in applied linguistics differ, as does the role of meta-cognition in language use. Stevick (1996) defines the term meta-cognition as *a Greco-Latin word that stands for what a person knows about his or her knowing*. He proposes his own interpretation of the term:

1. *we know what kind of knowledge we have or lack*

2. *we know when our attempts to use our knowledge are succeeding or failing*

3. *we know something about how we control or use our knowledge* (Stevick 1996, p.12).

Stevick does not discuss either meta-cognitive strategies or processes. The role he attributes to affect makes the need to operationalize knowledge unnecessary as affect for him forms the frame of reference for knowledge, it controls access to memory and provides evaluation of the performance (Stevick 1999, section 3.3.1).

Wenden (1991) proposes that the use of knowledge has to be viewed as a process:

1. Planning consists of preplanning and planning in action

2. Monitoring involves self assessment during the act

3. Evaluating takes place after the event when the learner has self-assessed the performance (Wenden 1991, p.28)

3.2.2 Operationalization of the concept

O'Malley et al (1985) made the concept fully operational by coupling meta-cognition to the 'learning strategy concept': *meta-cognitive learning strategies are generally applicable to a variety of learning tasks and include knowledge about cognition or applying thoughts about the cognitive operations of oneself or others and regulation of cognition or planning, monitoring and evaluating a learning or problem solving activity* (O'Malley et al 1985, p.25).

The vocabulary used by Purpura (1999) and Cohen (1998) when they talk about meta-cognitive processing (see Table 3) shows the movement from unconscious behaviour to consciously selected behaviour (Cohen 1998).

Table 3 Terms used to describe meta-cognitive processing

	Purpura 1999	Cohen 1998
Processes	The individual stages of mental activity seen in models of human information processing (e.g. comprehending, storing, retrieval)	Unconscious behaviour when learners are not able to identify any strategies
Strategies	Conscious or unconscious techniques, behaviours or activities that an individual invokes in language learning, use or testing	Processes, which are consciously selected by learners and may result in action taken to enhance the learning or use of foreign language through storage, retention, recall.

In 1990 Oxford developed a list of strategies in language learning (SILL) that include also meta-cognitive strategies. She divided SILL into two groups:

1. primary strategies that operate directly on the language itself (previewing, focusing attention, getting the meaning, taking notes, recognising and using

contexts, communicating, practising, learning the rules, reasoning, learning outside the classroom, memory building)

2. support strategies that enhance learning indirectly by creating a good attitude, establishing learning goals, reducing learner's frustration, tension, fatigue or anxiety. They also include planning and goal setting, self-management, social cooperation, creating practice opportunities and culture orientation.

The use of the concept of strategies, however, has its pitfalls. Purpura (1999) suggests that the term 'strategy' is tied to several concerns:

1. What is the precise nature of a strategy, is it an action, activity or behaviour, an operation a technique or tactic? Purpura (1999) defines strategies as specific actions, activities or behaviours that are directly connected to language acquisition, use or testing. This approach to my mind has its own problem, as actions cannot be determined by meta-cognition only, they are rather a product of at least cognition and motivation. This immediately involves affective variables in the definition of a meta-cognitive strategy.

2. Are strategies observable (Oxford (1990) or are they both observable and unobservable?

3. Are strategies conscious (Flavell 1979 and Cohen 1998) or are they both conscious and unconscious (Faerch and Kasper 1983)?

4. Which strategies have direct and which indirect impact on learning? Oxford (1990) thinks that memory, cognitive and compensatory strategies are 'direct' because they provide a direct support to learning, while meta-cognitive, affective and social strategies provide a non-direct support to language learning (Purpura 1999, p.23-24).

Nowadays the term 'meta-cognition' is often used side-by-side with the term 'strategic competence' (see Bachman and Palmer 1996 and Alderson 2000). Schoonen et al (1998), however, in their research on native and foreign language comprehension, use Alexander, Schallert and Hare's (1991) framework of meta-cognitive knowledge, where 'meta-cognition' is divided into four dimensions:

1. self knowledge: perceptions and understanding of oneself as learner and thinker

2. task knowledge: analyses of the cognitive demands of a task

3. strategic knowledge: knowledge of the processes that are effortfully planned and consciously invoked to facilitate the acquisition and utilization of knowledge

4. plans and goals: knowledge of the goals that may be established and the general plans that may be invoked (Alexander, Schallert and Hare 1991, p. 329).

This framework, to my mind, uncovers not only the complexity of the concept, but also allows us to compare the earlier models of language use which utilized the concept of 'strategic competence' with Bachman and Palmer's model which uses 'the meta-cognitive strategy' concept.

It is the third component 'strategic knowledge' or 'strategic competence' that has been part of foreign language performance models since the 1980s when Canale and Swain included it in their communicative competence model. Gradually the use of the notion of 'strategic competence' changed from 'repair strategies' as they were originally introduced by Canale and Swain to 'general language use enabling strategies' (Bachman 1990).

3.2.3 Meta-cognitive variables in language use models

In this section I will examine three models of language use produced by different authors: Canale and Swain (1980), Tarone and Yule (1989) and Bachman and Palmer (1996). The three models to my mind represent the gradual development of the present day view of the role of meta-cognitive strategies in language use. The theoretical models are supported by Purpura's (1999) empirical investigation of meta-cognitive strategies that is also discussed in this section.

3.2.3.1 Canale and Swain's language competence model

Knowledge of language in Canale and Swain's (1980) model consisted of three separate competences: grammatical competence, socio-linguistic competence and strategic competence. 'Ability for use' was not included in the model of competence, but was defined as the realization of all three competencies.

Canale and Swain deliberately excluded 'ability for use' from their model because performance might contain such factors as volition or motivation and they doubted that there was any theory of human action that could adequately explain ability for use; therefore they could not include it in their model. Nevertheless they defined strategic competence as the ability to compensate in performance for incomplete or imperfect linguistic resources in a second language (Canale and Swain, 1980, p.7). This made McNamara conclude that their definition of strategic competence as *a capacity for strategic behaviour in performance is likely to involve non-cognitive issues such as confidence, preparedness to take risks and so on* (McNamara 2000, p.18).

I do not share Canale and Swain's (1980) view that we have to wait until there is a satisfactory human performance model that could be used to develop a language performance model. On the contrary, I think that the data provided by language

testing have to be used also to test and improve the existing theoretical models of human information processing provided by social sciences. Purpura (1999) does this in his research thus making the results of his research important not only for applied linguistics but also for the understanding of human information processes in general. As a result we can understand the human information processing better and are able to adapt the instruments for language proficiency measurement accordingly.

3.2.3.2 Tarone and Yule's approach

Tarone and Yule propose that although repair strategies are important in successful communication, the area of strategic competence that they consider most important is *the overall skill of a learner in successfully transmitting information to a listener, or interpreting the information transmitted* (Tarone and Yule 1989, p.103).

Tarone and Yule consider that strategic competence involves the ability to select an *effective* means of performing rather than a correct one; as a result, the criterion for the evaluation of one's strategic competence is the degree of success in communicating. Its effectiveness would depend on:

1. *the speaker's knowledge of language*

2. *the speaker's knowledge of the world*

3. *the speaker's assessment of the listener's knowledge of the world* (Tarone and Yule 1989, p.106).

Although Tarone and Yule do not speak of affective variables here, these are implied by the speaker's assessment of the listener's knowledge. The speaker's assessment is impossible unless he or she makes the effort to put him or herself into the listener's shoes. This implies empathy with the listener and a consequential adaptation not only

of the content of one's performance, but also one's form of utterance. This reminds me of one of the roles of affect in language use (see Stevick 1999 in section 3.3.1) that of affective feedback which constantly monitors the form and contents of production in language use.

3.2.3.3 Bachman and Palmer's model of language use

In 1990 Bachman proposed his model of communicative language use, where strategic competence was considered as an important part of communicative language use; this also supported Tarone and Yule's approach. Bachman explained the difference that can be caused by one's strategic competence as *willingness to exploit what they knew and their flexibility in doing so* (Bachman 1990, p.105). I see 'willingness' as an indirect reference to affect, which is not present in Bachman's model of communicative language use published in 1990, but appears in Bachman and Palmer's version of the communicative language use model in 1996.

In 1996 Bachman and Palmer reformulated strategic competence as 'meta-cognitive strategies': *We conceive strategic competence as a set of meta-cognitive components or strategies, which can be thought of as a higher order executive process that provides a cognitive management function in language use as well as in other communicative activities. We identify three general areas in which meta-cognitive components operate: goal setting, assessment and planning* (Bachman and Palmer 1996, p.70). I will discuss these in detail in section 3.4.1.

Another innovation of the model is the inclusion of affective schemata in the language use model. The reason for this is that they *are the means by which language users assess the characteristics of the language use task and its environment in terms*

of past emotional experiences in similar contexts (Bachman and Palmer 1996, p.65) (see Figure 11).
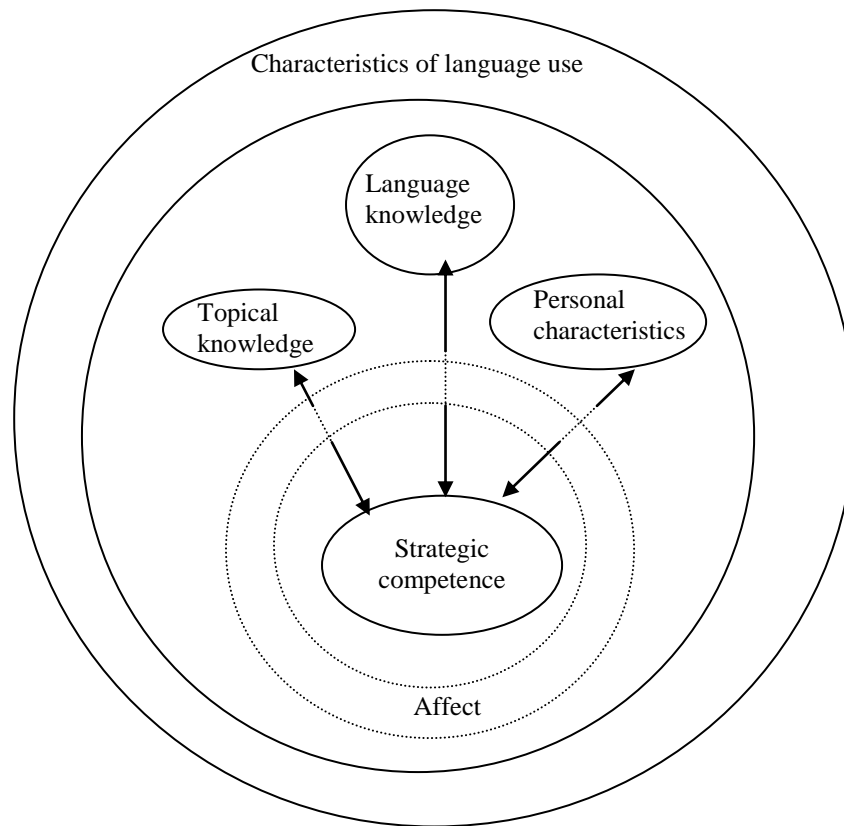
Figure 11 Some components of language use and language test performance, Bachman and Palmer (1996)

Bachman and Palmer consider that only through the interaction between strategic competence, topical knowledge and affective schemata can language users create and interpret discourse (Bachman and Palmer 1996, p.62).

If we compare Bachman and Palmer's (1996) model of language use with the models of interaction between cognition and emotion in psychology (see section 2.3.), we find many points in common:

1. Bachman and Palmer's model agrees with Power and Dalgliesh's (1998) theory and Snow and Jackson's (1994) proposals that affect and goals (one of the strategic competence areas) are directly connected and it is impossible to understand one's affect without understanding person's goals.

2. Similarly, Power and Dalgliesh's (1998) model of Domains of knowledge supports Bachman and Palmer's (1996) view that goals are closely connected with our world knowledge (in Bachman and Palmer's case 'topical knowledge').

3. If we compare Bachman and Palmer's (1998) flowchart with Combs and Merzano (1990) 'Self as agent' processing framework we find that it supports Bachman and Palmer's (1996) suggestion that affective schemata will decide the flexibility with which we respond to a task.

4. Bachman and Palmer's (1996) model also agrees with both Combs and Merzano's (1990) and Heider's Attribution scheme (1958) in including task at hand as separate factor that can affect the quality of our performance.

5. The idea of interaction between cognition and emotion that is at the basis of Scherer's (2000) and Lewis and Granic (2000) self-organisation systems theory, although not explicitly, is also present in Bachman and Palmer's (1996) model.

This allows me to conclude that Bachman and Palmer's model of language use agrees with the theories of psychology on the interaction between cognition and emotion.

3.2.3.4 Purpura's empirical model of strategic competence

Purpura (1999) investigated the nature of cognitive and meta-cognitive strategy use in different groups of language ability and the effects of these on second language test performance. He based his research on results from the First Certificate in English Anchor test. The aim of his research was to investigate empirically the models of strategic competence proposed by Faerch and Kasper (1983), Oxford (1990) and

Wenden (1991) and to explore the influence of strategic competence (both cognitive and meta-cognitive) on language performance, interaction between cognitive characteristics of the test-takers and their performance in a test, so that research on strategy use in human information processing models could be related to second language test performance.

He defined meta-cognitive strategies as *a set of conscious or unconscious mental or behavioural activities, which are directly or indirectly related to some specific stage of the overall process of language acquisition, use or testing* (Purpura 1999, p.6).

Instead of following the tradition of good learner strategy investigation Purpura chose the basic human information processing component model. According to this model the information processing system has two stages:

1. a cognitive stage of input, storage and retrieval

2. a meta-cognitive function of system management (Purpura 1999, p.31).

Having carried out Structural Equation Modelling (SEM) analyses of his data (for more details see section 5.2.1 and 10.3.1), Purpura found that

1. the hypothesis that meta-cognition was a multidimensional construct consisting of on-line assessment and post-assessment strategies was not supported. Instead, he found that meta-cognitive strategy use consists of only one underlying factor represented by general assessment processes with goal setting and planning strategies as special cases of assessment. Purpura says that meta-cognition may also include planning and goal setting strategies. His study, however, supports only the uni-dimensional notion of meta-cognition where the assessment processes are represented by questionnaire items relating to assessing situation, monitoring, self evaluating and self-testing.

2. Meta-cognitive strategy use has a direct positive impact on all three cognitive strategy use variables (comprehension, memory and retrieval) and this supports

the hypothesis that meta-cognitive strategies constitute self-management behaviours that oversee and manage the cognitive behaviours in second language acquisition use and testing.

3. Meta-cognitive strategy use has no direct impact on second language test performance (with the exception of two strategies of self-evaluation and monitoring which had a direct positive effect on lexico-grammatical ability); instead, it invokes retrieval type strategies, which in their turn have a direct impact on language performance. Purpura's research signifies that cognitive strategies function in concert with meta-cognitive strategy use and a student needs to use meta-cognitive and cognitive strategies together to maximise learning performance (Purpura 1999, p.127). This led Purpura to conclude that strategic competence includes cognitive and meta-cognitive strategies.

4. From the point of view of cognitive behaviour there are two main behaviour types during a language test: either 'process' or 'product oriented'. Students who are 'product oriented' view test input simply as a context that leads to the expected response but students who are 'process oriented' view test as a learning opportunity and rely more on memory strategies. The more 'product oriented' the test-takers are, and prefer retrieval instead of memory strategies, the better their performance.

As a result of Purpura's research, Wenden's (1991) hypothesis of meta-cognitive strategies as a process (consisting of pre, while and post processing strategies) was rejected and only one factor of general assessment was found to represent the meta-cognitive strategies. This led Purpura to question the multi-dimensional depiction of meta-cognitive strategies proposed by Faerch and Kasper (1983), O'Malley and Chamot (1990) and Oxford (1990).

Purpura also found that meta-cognitive strategies had a significant indirect impact, and cognitive strategies had a significant direct impact on language use.

3.2.4 Functions of meta-cognition in language use

Skehan (1998) considers that meta-cognitive strategies are concerned with reflection and flexibility:

1. reflection represents the learners' self-awareness in learning, their ability to appreciate their own strengths and weaknesses

2. flexibility organises and gives purpose to the way how cognitive and social-affective strategies are used and increases the likelihood of appropriateness of strategy choice

Converted to Bachman and Palmer's terminology, reflection could be seen as a form of self-assessment, and flexibility as a combination of planning, goal setting and task-assessment strategies.

Wenden (1998) considers that meta-cognition has the following functions in language use:

1. meta-cognitive knowledge constitutes internal feedback, it is like a state of awareness which reveals how well learning is proceeding

2. it is the basis for internal assessment of comprehension and/or progress towards the goal

3. it may suggest the reasons for any problems revealed by the state of awareness

4. it may be drawn upon to guide decision making during the monitoring process phase of

5. it will lead learners to expand their meta-cognitive knowledge

If we compare this list of functions of meta-cognition with Stevick's (1999) list of roles of affect (section 3.3.1.) in language use it is difficult not to notice parallels. For

example: the role of feedback is attributed by Stevick to affect and to meta-cognition by Wenden, similarly, choosing one's goals and selecting the most appropriate plan (flexibility) is attributed to meta-cognition by Skehan and to affect/emotion by Stevick (1999) Power and Dalgliesh (1998), Rolls (1990), Frijda (1986) and Pritchard (1976). This suggests that there is an overlap between the functions of the two notions in psychology as well as linguistics. This provides basis for the analyses of interaction between cognition and emotion and will be discussed further in section 3.3.1.

3.3 Affect

Affect in applied linguistics has been traditionally studied as one of a language user's individual variables (next to language aptitude, attitudes, language anxiety and language learning strategies) (see Gardner 1997) that either promote or hamper language acquisition or use process.

Language acquisition research that focusses on motivation studies affect as a positive force that has to be promoted (see Gardner1990 and Oxford 1990) while classroom anxiety research (Horwitz 1986) and test anxiety research (MacIntyre and Gardner 1991) see affect as interference. Until lately the role affect in language learning as a unified system was not studied in applied linguistics, although Oller remarked that *emotive or affective factors play a greater part in determining success or failure in schools than do factive or cognitive factors*. (Oller 1979, p.105). Oller also said that the problem is how to determine what the affective factors might be.

The situation has improved lately, as we now have an instrument for researching affect (the Affective survey of Ehrman and Oxford 1991) and a theoretical framework of the roles of affect in language acquisition (Stevick 1999). Although Ehrman and Oxford (1991) stll see affect as a composition of three separate variables (motivation, self beliefs and anxiety), the development of a single instrument is an important step

in acknowledging affect as a system that interacts with cognition and shapes our behaviour in general, and language learning and use in particular (see Okada, Oxford and Abo 1996).

Although language testing has been traditionally concerned with affective variables as a source of error in language proficiency measurement, the situation has also changed in language testing, as Bachman and Palmer (1996) propose a theoretical language model that contains affect as a part of language use. Nevertheless, Purpura (1999) and McNamara (1996) consider that it is not clear how affect could be operationalised to describe and measure its influence during a language test.

3.3.1 Functions of affect in language use

Stevick (1999) defines affect functionally: *affect towards a particular thing or action is how this action or situation fits in with one's needs or purposes and its resulting effect on one's emotions* (Stevick 1999, p.44) and thus includes in the term 'affect' both the positive (motivation) and negative evaluation (anxiety).

Stevick seems to be using 'emotion' as a wider term than 'affect' as he says: *The inclusion of emotion along with needs and purposes is not surprising when we consider that emotions are commonly responses to how one's various needs and purposes are or are not being met* (Stevick 1999, p.44).

Stevick considers that affect has at least five roles in learning and memory:

1. affective data are stored in the same memory networks as other kind of data and may even be the kinds of data around which those networks are organised

2. affective data may call up from long-term memory certain other kinds of data and these extra-data may act as a clutter on the worktable, using up processing capacity

3. the affective side of feedback influences the shaping and reshaping of the networks of the long term memory

4. affect is important in initiating voluntary playback of language and plays a part in response to involuntary playback

5. affect may interfere with one's inability to draw on data in the long-term memory.

I will now proceed to compare each of the 5 Stevick's roles of affect in language use with the findings of other researchers.

3.3.1.1 Organisation

The first role that Stevick (1999) sees for affect is that of 'organisation'. He considers that affective data are themselves stored along with other kinds of data in memory. According to Damasio (1994) emotions are not merely parts of the network of memory. He calls the brain *the captive audience of the feelings*. And since what comes first constitutes a frame of reference for what comes after, emotions may actually be the parts around which those networks are organised and they may have a say in how the rest of the brain and cognition in particular go about their business.

Power and Dalgliesh (1999) also discuss the interconnection of the cognitive and emotional variables in schemata, but they consider that the inner goal structures determine the organisation of the cognitive and affective structures. This suggests a more dynamic and interactive model of schema organisation than Stevick's (1999).

3.3.1.2 Interference

The second role that Stevick discusses is that of the interference that affect can cause: affect can interfere with the cognitive processing of the language input. According to Goleman (1995) *the prefrontal cortex is the region of the brain responsible for working memory. But the circuits from the limbic brain* (the supposed seat of the emotion system, LeDoux 1999) *to the prefrontal lobes mean that the signals of strong emotion can create neural static, sabotaging the ability of the pre-frontal lobe to maintain working memory* (Goleman 1995 p.27). Stevick exemplifies this with a metaphor of a cluttered working table that interferes with one's work.

Goleman's (1995) and Stevick's (1999) use of the term of interference describes how language anxiety is created and how it operates. The problem, however, is that Stevick neglects the fact that emotion is also responsible for the support the emotions can give to our processing capacity - 'flow' ; this is discussed by Goleman (1995) and Csikszentmihalyi (1998), (for explanation of 'flow' see section 2.1.).

So I would paraphrase the role of 'interfering' as supporting or hampering the activity depending on the person's previous experience and the level of difficulty of the task. Csikszentmihayli and Nakamura (1989) say that to enter a state of flow the task has to be just above the appropriate level of difficulty: difficult enough to challenge, but not so difficult that it causes anxiety. The instructions have to be straightforward and not ambiguous. So for the task at hand one either can either make use of the energy emotions provided or get sidetracked and think about the disasters of failure.

3.3.1.3 Monitoring

The next role Stevick (1999) discusses is that of monitoring and control. He says that it is through feedback that affect constantly reshapes the networks of the long-term

memory in one's brain. Stevick differentiates between external and internal, cognitive and affective, positive and negative feedback. Cognitive feedback answers the question *How satisfactorily did I get my message through*? Affective feedback answers the question *What kind of feeling did I come away with?* (Stevick 1999, p.51)

- Stevick considers that external cognitive feedback derives its force from the human desire to transmit and receive ideas. When we perceive that the communication has been full and accurate, the external feedback is positive.

- Internal cognitive feedback is made possible by the two-way traffic between working and long-term memory and particularly by the comparisons of different linguistic items and forms that are made in working memory. If the forms do not match, the cognitive feedback is negative, but if we find a form that matches the meaning that we want to express, the cognitive feedback is positive.

- External affective feedback derives its effectiveness from our desire to identify with or disassociate ourselves from a particular group of people. For example, it answers the question 'Do they like me?' If the other person seems to be interested in what I am saying, the external feedback will be positive and will influence my willingness to keep communicating, in spite of my internal cognitive negative feedback (my cognitive awareness that I am making mistakes while talking).

- Internal affective feedback depends on the two-way traffic between the short-term and the long-term memory and relates to evaluating one's own linguistic performance. The question to be answered is: 'Do I sound like a member of the speech community?' The criteria include both message-bearing as well as

non-message-bearing features like minor points in pronunciation and word choice.

Rolls (1990), however, differentiates between two separate functions of emotions: evaluation and communication about the way the person feels. This to my mind brings out the difference between evaluation of one's own actions and perception of the emotions of others. These two functions are actively used both by the test-taker and the interlocutor during a speaking test that involves interaction between the test-taker and interlocutor. If, for example, the external affective feedback from the interlocutor is positive it helps the test-taker to speak better. Negative affective feedback, however, can easily cause anxiety.

The role of feedback or monitoring in Stevick's interpretation can be compared to Bachman and Palmer's (1996) of the area of assessment: *taking stock of what is needed, what one has to work with and how well one has done* (Bachman and Palmer 1996, p. 71). Evidently this may be another area of interaction between affect and meta-cognition.

3.3.1.4 Language playback

Stevick (1999) considers that affect is important in initiating voluntary playback of language (spoken by others) and plays a part in response to involuntary playback. This to my mind is part of the assessment strategy. For example, after an interview that was important to me I would 'go through it' again and again scanning my memories for mistakes or reliving my success. Although Stevick does not talk about it, involuntary replay seems to have an important role in pronunciation and intonation acquisition, when learning a new foreign language.

In a language testing situation this role is used as after the test is over, we replay and examine our own performance once again. The role of playback of affect would be useful to Wenden's (1991) 3rd role of meta-cognition (evaluation after the event): the evaluation of the performance could work in concert with language playback.

3.3.1.5 Control of access to memory

Stevick (1999) considers that affect influences our ability to draw material from the long-term memory. He describes an experiment where he pretended to be either as an interested listener or an indifferent one, in the first case the person responding talks more fluently and has more things to say than in the second case. He explains this by Damasio's (1994) statement that *along with negative body states, the generation of images is slow, their diversity is small, and reasoning inefficient; along with positive body states the generations of images is rapid, their diversity wide, and reasoning may be fast though not necessarily efficient* (Damasio, 1994 p.147).

Stevick's (1999) discussion of the roles of affect in language learning is crucial for an understanding of the interconnection between the cognitive and affective aspects in language use. However, if we compare the roles described by Stevick with the functions of emotions in psychology we find some functions that are unaccounted for:

1. Rolls (1990), considers that emotion has its role in creating the need for social bonding. This may be the basis of language ability development of the human species and the basis of motivation in foreign language learning. It is possible, though, that Stevick has deliberately 'ignored' motivation as it has been adequately discussed by other researchers (for example, Oxford 1990).

2. Damasio (2000) adds one more function of emotion in language production, that of providing speech rhythm and intonation that characterize the speaker in

general as well us informs us of his or her well-being and attitude to the object under discussion. He considers that to produce language we have to be awake (alert) and have background emotions (feel well-being or malaise, calm or tension) and this is reflected in a person's speech. He says that emotional aspects of the communication are separate from the contents of the words and sentences spoken. Words and sentences from the simple 'yes', 'no' and 'hallo' to 'good morning' or 'good-bye' are usually uttered with a background emotional inflection. The inflection is an instance of prosody, the musical, tonal accompaniment to the speech sounds that constitute words. Prosody can express not only background emotions, but also specific emotions (embarrassment, jealousy, guilt, pride), when emotions are expressed explicitly and purposefully.

All the roles or functions of affect discussed here have a direct influence on language test performance: affect mobilizes our energy to do well in the examination, it organizes our memories and controls access to it, helps us choose the task if there is a choice and helps us choose strategies for fulfilling it. Affect determines the rhythm of our speech and gives additional meaning to the content with the help of intonation. Affect also monitors our progress by telling us how well or badly we are doing during an examination and evaluates our performance after it. It warns us if the needs of the task do not match our abilities; it mobilizes more strategies and activates associations with our previous experiences to improve our performance. Affect can also overreact by activating so many strategies and so many previous experiences that we cannot cope with all the activity and give up, thus destroying our own performance.

This description of the functions of affect may seem exaggerated only if we see affect as isolated from cognition, but if we think of affect as basis of consciousness, it is

only natural that it is present in every cognitive operation either as a background emotion or as a secondary or social emotion (see Damasio 2000).

3.3.2 Affective factors as a source of error

The Dictionary of Language Testing (Davies et al 1999) defines affective reaction as *the emotional reaction or engagement of a test-taker to a test. Affective reactions are recognised as influencing the quality of the test performance and as such will contribute to measurement error* (Davies et al 1999, p.4). This definition presents a collective view of language testing: affect is a source of error and should therefore be avoided, in spite of being part of Bachman and Palmer's (1996) model of language use. The dictionary (Davies et al 1999) refers to Porter (1991) as a further source of information on affective factors in language testing.

3.3.2.1 Predictable versus unpredictable factors

Porter (1991) defines affective factors as emotions and attitudes that affect our behaviour. He distinguishes between the effects of predictable and unpredictable factors: *predictable affective factors include age, status, personality type, acquaintance-relationship and the gender of the test-takers. (…) Unpredictable factors include moods that weaken the concentration or associations that for personal reasons affect the test-takers' performance* (Porter (1991) p.33). Although this definition explains sources of possible error, it does not have much in common with the present understanding of affect. For example, it is not clear why it should include gender or age.

However, one has to take into account that this was written in 1991, when affect was not part of Bachman's language use model (see Bachman 1990), but was treated as a test-taker characteristic and a source of error. To control measurement error, test developers are cautioned against using tasks that could involve affective response: *If*

we ask the test-takers to deal with an emotionally charged topic, such as abortion, gun control or national sovereignty, their affective responses to this topic may limit their ability to utilize the full range of language knowledge and meta-cognitive strategies available to them (Bachman and Palmer 1996, p.66). The question remains, however, what to do with the test situation, which is for many people an emotionally charged experience.

Porter (1991) acknowledges the problem that affect causes for language testers saying that *there would not be a lot that we could do about it as we cannot strip the texts of all the emotions because emotional content is the central feature of natural language use. At the same time we have to admit that it is a potential source of unreliability* (Porter 1991, p.34).

This seems to be a problem without a solution only if we treat affect as a unitary concept. If, however, we apply the different functions of affect to this problem we can see that affect has different roles to play in language use. Some functions of affect can interfere with the ability to access language knowledge and therefore can cause error of measurement (especially in interaction with test facets) while others are necessary for us to understand language and react in an appropriate way.

3.3.2.2 Systematic versus unsystematic factors

Bachman (1990) includes test-taker characteristics in his discussion of the sources of error. He differentiates between three different groups of factors other than language ability that affect performance on language tests:

1. test method facets (uniform from one test administration to the next)

2. attributes of the test-taker that are not considered part of the language abilities we want to measure, such as cognitive style and knowledge of particular content areas, sex, race and ethnic background

3. unsystematic random factors (test-takers' mental alertness or emotional state, changes in the test environment, idiosyncratic differences in the way different test-administrators carry out the test instructions).

Bachman (1990) says that there is sufficient evidence about the influence of personal characteristics on language test scores to distinguish them as a separate factor. In addition they can also interact with the test-facets to constitute additional sources of variation. He calls the first two factors systematic and the third unsystematic. Thus according to his classification the affective factors' impact on test performance can be systematic (as a part of strategic competence) and unsystematic (when the test-taker experiences some temporary personal problems or is affected by test-tasks in a way that differs from that of the other test-takers). As a result I do not see in this model any differentiation between test anxiety and any other functions of affective or volitional aspects of language use. One can argue that Bachman's (1990) model of language use does not contain affective schemata, and this is true. But role of the affective schemata in language use in the 1996 model has not been supported by a re-evaluation of their impact on test-score.

Given the nature of affect as a cluster of variables that, according to Oxford (1990), consists of anxiety, motivation and beliefs about oneself, it will be a systematic factor that affects the test-takers' performance, as it will affect performance in every circumstance. Random factors and unsystematic error are caused by the test-takers' inability to react in a regular way because of personal reasons (for example, an illness) or because of ambiguity in the test rubric, test input or setting. Bachman and Palmer (1996) do not examine these factors beyond stating: *The test-takers' assessment of the characteristics of the task will determine their affective response to that task and their flexibility in using his* (I guess the authors meant also ' her') *language ability to complete it. We would thus expect a positive affective response to the test task to make*

the test task relatively more interactive, and a negative affective response to have the opposite effect (Bachman and Palmer 1996 p.145).

Thus according to Bachman and Palmer the crucial difference would lie in the interactiveness of the test. One might oppose here that as proposed by Shohamy (1982), another important feature for the test-takers is that the test-taker recognise the test format ('novelty check' Scherer 2000). So if the test-taker is used to traditional tasks while learning and has suddenly to fulfil an interactive task, it will be the interactiveness of the test that will cause negative reactions. (This was not the case in Latvia's Year 12 exam, though.)

I think that if we continue to treat anxiety and other affective variables as individual characteristics that cause error in our measurement of language performance and therefore strip our tests of any topic that might cause affective reaction in a population of thousands of test-takers, we will be sacrificing language test validity to increase reliability of measurement. I believe that by accepting affect, like meta-cognition, as a part of language competence we have to accept responsibility for the possibility that some topics will be 'emotionally charged' for some test-takers. Consequently we have to develop methods to register test-takers' reactions to our tasks and decide how far our test should go in measuring test-takers' ability to express their emotions and perceive and interpret emotions of others as expressed in a foreign language. This attitude to emotions or affect would to my mind increase foreign language test validity and have a positive washback on language teaching.

3.3.3 Affective schemata

Another stumbling block for a researcher of affect in language use can be created by the term 'affective schemata'. It is not clear how Bachman and Palmer (1996) differentiate between 'affect' and 'affective schemata', because in the model they have

'affect' (see their Figure 9), but in the description of the functions we see the term 'affective schemata'. Also in the model of interaction between the meta-cognitive strategies and affective schemata (see their Figure 10) the authors use the term 'affective schemata'. This inconsistency was noted by McNamara in 1996, but he could give no explanation for it.

Frijda and Mesquita (2000) define affective schemata (also called 'sentiments' in psychology) as *appraisal structures that include concerns of relevance of its object. They consist of the latent representation of some object as being relevant to one's concerns. Sentiments thus are dispositional emotions. They are schemas with the same structure as emotions* (Frijda and Mesquita 2000, p.55).

Affect, in its turn is a much wider concept. Stevick defines affect towards a particular thing or action as to how this action or situation fits in with one's needs or purposes and its resulting effect on one's emotions (Stevick 1999, p.44). The crucial difference between the two terms to my mind is that affect is not just a latent representation of the evaluation of the previous encounter (in the form of memory), it is the on-line evaluation of how this new encounter fits into one's goals and all the resulting emotional reactions that are encountered by the person and which have to be dealt with.

Frijda and Mesquita (2000) remark that sentiment (or affective schemata) can turn into a real emotion if the object is encountered with sufficient urgency or proximity, but it is just a possible development. Affective schemata can be defined as one of the roles of affect (as discussed by Stevick 1999), that of memory function (see section 3.3.1.).

Although one might argue that Power and Dalgliesh (1998) consider that emotion (or affect) production needs a schema (see section 2.2.2.). According to their theory

emotion is born as a reaction to the schema a person has created while imagining how the present situation could influence achievement of his or her long term or short-term goals. However, Power and Dalgliesh (1998) do not call emotions "schemas", because for them a schema is just one phase in development of an emotion, which is followed by bodily reactions. Another argument against using the term is the fact that 'affective schemata' are used for a different concept in psychology (see Frijda and Mesquita above).

If we compare the roles of affect in language learning as described by Stevick (1999, pp.43-57) with Bachman and Palmer's (1996, pp.65-66) description of affective schemata we find that the only role that is missing in Bachman and Palmer's (1996) model is Stevick's initialisation of voluntary or involuntary playback (Role 4). This could be explained by the fact that it is perhaps more important in language acquisition than in testing. Apart from this one function, the two frameworks match almost perfectly if we look at Bachman and Palmer's definition of the role of affective schemata:

Affective schemata can be thought of as the affective or emotional correlates of topical knowledge [see Stevick 1999, p. 47, Role 1]. *These affective schemata provide the basis on which the language users assess, consciously or subconsciously, characteristics of the language use task and its setting in terms of past experiences in similar contexts* [p.50, Role 3]. *The affective schemata (...) can either facilitate or limit the flexibility with which he* (I think the authors mean by 'he' both male and female language users) *responds in a given context* [p.47, Role 2]. *The affective responses of language users may thus influence not only whether they even attempt to use language in a given situation, but also how flexible they are in adapting their language use to variations in the setting* [p.52, Role 5] (Bachman and Palmer 1996, p.65).

Having compared Bachman and Palmer's (1996) description of the role of affective schemata with Stevick's (1999) roles of affect in language use, and Frijda and Mesquita's (2000) definition of affective schemata, I have to conclude that Bachman and Palmer's definition of 'affective schemata' is closer Stevick's definition of 'affect' than Frijda and Mesquita's definition of 'affective schemata'. This allows me to connect Stevick's findings on the functions of affect with Bachman and Palmer's model of language use in one system that will be explored in the next section.

3.4 Relationship between affect and meta-cognitive strategies in language use models

As in psychology, in linguistics the interaction between meta-cognition and motivation has been operationalised, measured and its effect assessed (Okada et al 1996), while the interaction between meta-cognition and affect is described only theoretically [in Bachman and Palmer's (1996) model]. I will start this section with looking at the research on interaction between language competence and individual differences and later review the research on interaction between meta-cognition and affect in language use models.

3.4.1 Bachman and Palmer's model of interaction between affect and meta-cognition

Bachman and Palmer's (1996) model of language use and strategic competence (Figure 11) is supported by a flow-chart demonstrating the interaction between the strategic competence components and affect (see Figure 12). In spite of the fact that in the textual description the role of affect is only sketched in, we can deduce from the flow chart how the interaction happens: the task and the situation is assessed through the interaction between 'affective schemata' and all three meta-cognitive strategy areas. Assessment and planning areas interact also with language knowledge, topical

knowledge and personal characteristics to come up with a plan that is uttered or interpreted and again assessed by the meta-cognitive strategies. I will now examine the interaction between affect and each area of meta-cognition separately.

3.4.1.1 Interaction between goal setting strategies and affect

Goal setting involves identifying the language use tasks or test tasks, choosing one of several tasks, if given a choice, and deciding whether to attempt to complete the tasks. Bachman and Palmer say that the choice is usually much greater in language use situations than language test situations (Bachman and Palmer 1996, p.71).

The goal setting area description is the only meta-cognitive strategy area where
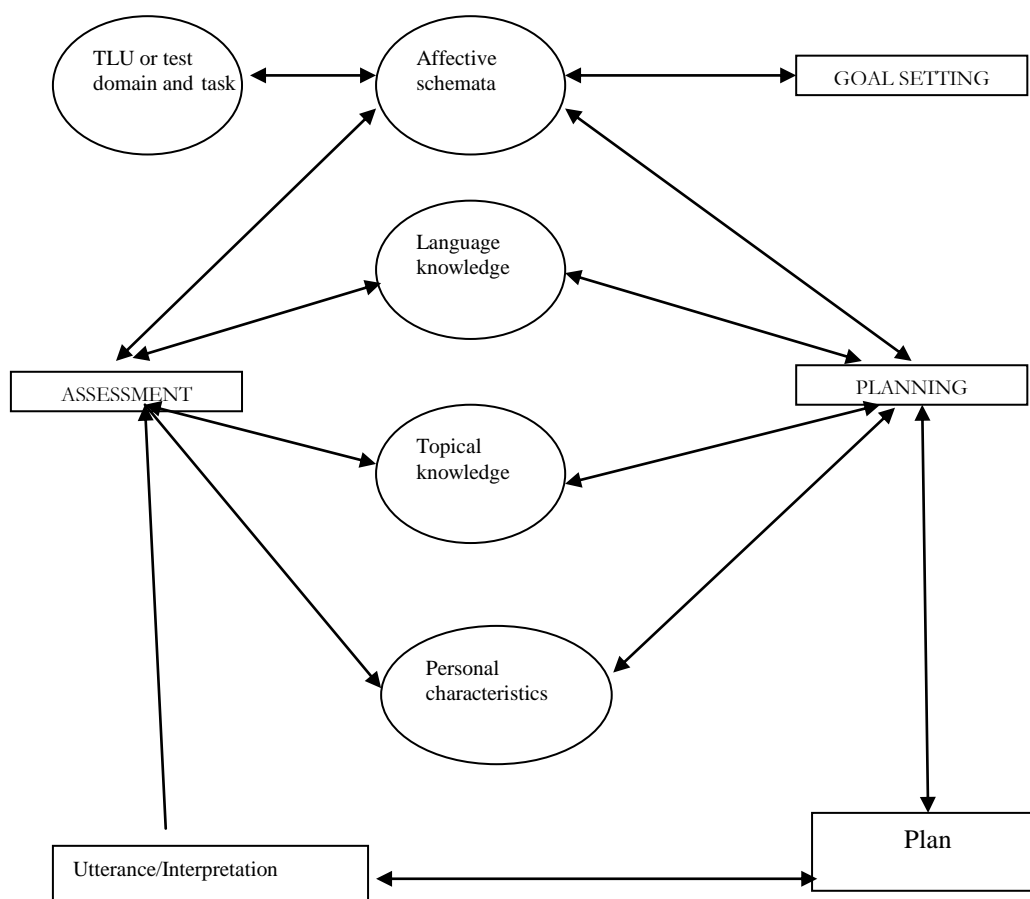
Figure 12 Meta-cognitive strategies in language use and language test performance, Bachman and Palmer (1996)

Bachman and Palmer do not specify the way in which affective schemata interact with this area. It could just be an error, because as we can see in Figure 12 there is a two-way arrow between the goal setting strategy and affective schemata, so an interaction between the two takes place (according to Bachman and Palmer's model). The goal setting strategies are isolated in the flow-chart as 'TLU or test domain and task; they interact only with the 'affective schemata'. They do not relate to the assessment of the tasks, whether these match the goals or not, and there is no arrow between 'goal setting' and 'planning strategy', although it seems improbable that goals would not influence any plans made.

3.4.1.2 Assessment strategies and affect

Assessment strategies take stock of what is needed, what one has to work with and how well one has done (Bachman and Palmer 1996, p. 71). 'Assessment' is the most active strategy use group in their model (see Figure 12), it interacts with affect, language knowledge, topical knowledge and personal characteristics as well as assesses an utterance or an interpretation after execution. Thus it is a strategy that is active throughout the test or language use task.

Research in other areas (test anxiety, see section 4.2.) implies that assessment strategies can also start acting before the test as a motivator. For example, if the test taker has decided that his or her knowledge is inadequate for the needs of the test, the decision supported by the emotional reaction ('worry' see section 4.1.3.) will make the test-taker draw up a plan of how to improve the situation and prepare for the test. Bachman and Palmer, however, are more cautious and just remark that *assessment also takes into consideration the individual's affective responses in the application of assessment strategies* (Bachman and Palmer 1996, p.71), leaving us to interpret it as we like, whether it is the 'lateral representation that is taken into account' (definition

of 'affective schemata' in Frijda and Mesquita 2000) or the active role that also propels the individual into remedial action.

Bachman and Palmer (1996) consider that *assessing the characteristics of the language use or test involves identifying the characteristics of the language use task or test task in order to determine*

1. *the desirability and feasibility of successfully completing the task and*

2. *what elements of topical knowledge this is likely to require* (Bachman and Palmer 1996, p. 72-73).

Here the authors are using assessment strategy as a link between the test task affective evaluation (desirability) and language knowledge.

What worries me is the fact that there is no direct connection between assessment strategy and the test task; it is as if the task is seen only through an affective schemata lens. Although the authors say: *The affective responses of language users may thus influence not only whether they even attempt to use language in a given situation, but also how flexible they are in adapting their language use to variation in the setting* (Bachman and Palmer 1996, p.65), this implies influence, but that it is not the only basis of decision-making as the graphical representation seems to imply. I think that there should be a direct connection between assessment strategies and language use task or test just as there is a direct connection between 'assessment' and the individual's own 'topical knowledge' and 'language knowledge (see Figure 12).

The topical knowledge of the individual would be the one that would contain the affective schemata in Frijda and Mesquita's (2000) sense. Bachman and Palmer (1996) remark that apart from determining the extent to which relevant topical knowledge and areas of language knowledge are available, assessment will also

consider the individual's available affective schemata for coping with the demands of the task. It can happen that the affective schemata (the memories connected with the topic) are so powerful that they can interfere and make the relevant language knowledge unavailable. This effect, however, is not produced by the affective schemata, but by the emotional reaction that turns the latent representation into real emotion with all its consequences (see role of interference, section 3.3.1.).

3.4.1.3 Planning strategies

Planning builds on the results of an assessment of what knowledge will be necessary for the task in question: Planning involves deciding how to utilize language knowledge, topical knowledge and affective schemata to complete the test task successfully (Bachman and Palmer 1996, p. 73). It involves selecting concepts, words, structures and functions, formulating one or more plans and selecting a plan for implementation as a response to the task. It may also involve prioritisation among the elements selected, considering how they will be combined most effectively and will result in a plan whose realisation is a response to the task. The authors give an example of how test takers go about planning depending on their assessment of the task needs and the goals they have set for themselves. Evidently there is an interaction between goal setting and planning, as there is not just the ultimate goal but also the interim goals for each task as a result of assessment. There is also an assessment of the answer provided and this will again produce new goals and new plans.

Although Bachman and Palmer's flow-chart (Figure 12) is clear and explicit in showing interaction between affect and meta-cognitive strategies during language use, there are several points that cause objections:

1. 'TLU or test domain and task' is not connected to any meta-cognitive strategies directly. As a result, the flow chart creates an impression that 'goal

setting', 'planning' and 'assessment' strategies will have access to the 'task' only through 'affective schemata'. As a result the flow-chart disagrees with appraisal theory (for example, Scherer 2000) or, in fact, any cognitive emotion theory that postulates that emotion production is based on cognitive appraisals

2. The meta-cognitive areas (goal-setting, assessment and planning) are not interconnected which suggests that there is no direct interaction between them. This would suggest that our plans do not depend on our goals and the findings of assessment strategies do not influence the new plans or revising of goals directly without mediation of affective schemata.

3. According to the flowchart 'language knowledge' and 'topical knowledge' can be accessed by all three meta-cognitive strategy areas directly, but they will not have a direct access to the 'task' as it can be accessed only through 'affective schemata' (see Figure 11). This disagrees with Bachman and Palmer's definition of the role of assessment strategies: *Assessment provides the means by which the individual relates her topical knowledge and language knowledge to the language use setting and tasks or to the testing situation and tasks* Bachman and Palmer 1996, p. 71). It is not clear how assessment can relate language knowledge to the tasks if there is no direct connection between 'assessment' and 'task'. Bachman and Palmer say that assessment strategies take into account 'individual's affective responses', but to my mind 'taking into account' in the verbal description does not match the overwhelming centrality of the 'affective schemata' in the flowchart.

McNamara says that in Bachman and Palmer's model *For the first time an attempt has been made to deal explicitly in a model of second language communicative ability with the* ability for use*, which relates to affective or volitional factors. This is an important advance. However, Bachman and Palmer, having confidently*

lifted the lid of Pandora's box, shut it again. Their subsequent discussion of the significance of the inclusion of the affective domain in language performance is restricted to some advice about how to 'bias for the best' (McNamara 1996, p.74). It is difficult not to agree with McNamara, both in appreciation of dearing to tackle the subject as well as in regretting avoiding articulation of the consequences of their theoretical findings in language testing practice.

3.4.2 Motivation and meta-cognition in language use

Okada, Oxford and Abo (1996) were also interested in the interaction between meta-cognition and affect, but chose another affective variable, motivation, to investigate the interaction empirically. They used a strategy inventory for language learning (Oxford 1990) and an affective survey (Okada et al 1996) to investigate the interaction between motivation and meta-cognition. Their main hypotheses were:

1. learners of Japanese are more motivated than learners of Spanish

2. learners of Japanese show more frequent use of a wider range of strategies than do learners of Spanish

3. significant correlations exist between motivation and strategy use for each language group.

All the three hypotheses were supported by their research: learners of Japanese turned out to be using more cognitive strategies. The authors quote their students reporting the following strategies*: I look for patterns, I develop my own understanding of how language works and I imitate the talk of native speakers, I find opportunities to practice the language, I try to notice the errors and find out the reasons for them, I look for people with whom to speak the language and affective strategies I pay close attention to the thoughts and feelings of others with whom I interact in the language*. All in all there were thirteen significant differences, which were to the advantage of the learners of Japanese.

The correlation coefficients between total strategy use and total motivation were significant in both language groups. For learners of Japanese it was .56 (p<. 002) and for learners of Spanish it was .58 (p<. 0003). Other significant correlations were also found for the separate strategy groups (see Table 4)

Table 4 Correlations between meta-cognition and affective variables (Okada, Oxford and Abo 1996)

Correlations between…	Japanese group	Spanish group
Meta-cognitive strategies and effort	.74 (p<0001)	.63 (p<0001)
Meta-cognitive strategies and total motivation	.72 (p<0001)	.69 (p<0001)
Meta-cognitive strategies and desire to use the L2	.69 (p<0001)	. 50 (p<0024)
Meta-cognitive strategies and intrinsic motivation	.57 (p<0014)	.62 (p<0001)
Meta-cognitive strategies and extrinsic motivation	.47 (p<0092)	

Okada et al 's (1996) research supports the hypothesis that meta-cognition is closely tied to motivation. They consider that learners with a strong will to pursue their goals will no doubt be active in planning, organising and evaluating their own study.

Chapter 3 explored the notions of affect and meta-cognition separately and then their interaction in the most influential language use models. The results of the examination suggest that both in the definition of functions of affect and meta-cognition and in the existing models of language use there is evidence of interaction between affect and meta-cognition. This provides basis for a supposition that there will be also interaction between foreign language anxiety and meta-cognition. Before examining the interaction between the two, I will first explore the concept of anxiety (see Chapter 4).

Chapter 4 Anxiety as a shadow of intellectual activity

The amount of literature available on anxiety in psychology is overwhelming and it is impossible to review even all the main fields of research here. Therefore I will concentrate only on the findings of those researchers who saw anxiety as interacting with (meta) cognition, like, for example, Lidell (1949) whose idea I have used for the title of this chapter. He says: *I have come to believe that anxiety accompanies intellectual activity as its shadow and that the more we know of the nature of anxiety the more we will know of intellect* (Lidell 1949. p. 184).

4.1 Anxiety and its correlates in psychology

Anxiety is usually used together with terms like fear, worry, and stress. These are sometimes named as sources of anxiety and sometimes as its constituent parts. I will start by describing views of anxiety and then examine also its correlates: fear, worry and stress.

4.1.1 Anxiety

Anxiety according to Eysenck (1992) is a multidimensional construct and it can be studied as a physiological, cognitive or social phenomenon. It can also be studied either as a healthy developmental or as a neurotic phenomenon (May 1979).

I will focus on cognitive approaches to anxiety in general and healthy anxiety in particular as I am mostly interested in the effect of anxiety on language test performance, but I will also briefly address the physiological and behavioural aspects of anxiety as the test-takers have to come to terms with all those aspects of anxiety that are caused by the test situation as well as their previous learning experience and even heredity.

4.1.1.1 Healthy anxiety: a means of development

Kierkegard (1849) considers that the main cause of anxiety is human confrontation with freedom and the possibility of attaining one's goals. As soon as a possibility is visualised, anxiety is experienced. If there is no hope, there is no anxiety, and we experience anger or depression instead. Kierkegard considers that normal anxiety has to be overcome by a person if he or she is to grow and develop; our selfhood depends on our capacity to move ahead despite anxiety. Kierkegard says: *Possibility means* I can. *In a logical system it is convenient enough to say that possibility passes into actuality. In reality it is not so easy and an intermediate determinant is necessary. This intermediate determinant is anxiety* (Kierkegard 1849/1941, p.44).

Kierkegard considers that by avoiding anxiety we lose the most precious opportunities of our own education as human beings. He says that anxiety is a teacher within ourselves; it makes itself known to us as soon as there is the possibility to learn. Our task is to learn to accept anxiety as a teacher; this will not only allow us to realize our potential, but also to lose the dread of anxiety.

A similar view is expressed by Csikszentmihalyi (1998). He considers that anxiety is one of the stages of an individual's development, which has to be gone through in the acquisition of every skill, to obtain every novel experience. Whenever we learn something new, we start off feeling anxiety. It is a feeling that we have to face a challenge that provokes anxiety (see the Figure 13). If and when we decide to move on, we act in a state of arousal until we start experiencing success, which according to Csikszentmihalyi, is the highest experience of joy and fulfillment of human existence. It could also be seen as the opposite of anxiety. If anxiety is the punishment our subconsciousness inflicts for not going ahead, flow is the reward. After flow, we experience the feeling of control, relaxation and finally boredom; once the skill has

been fully automatisized we do not learn anything by exercising it. The next stage is apathy and worry which again turns into anxiety until we take the next step in our development and launch into a new experience that needs the acquisition of a new skill.
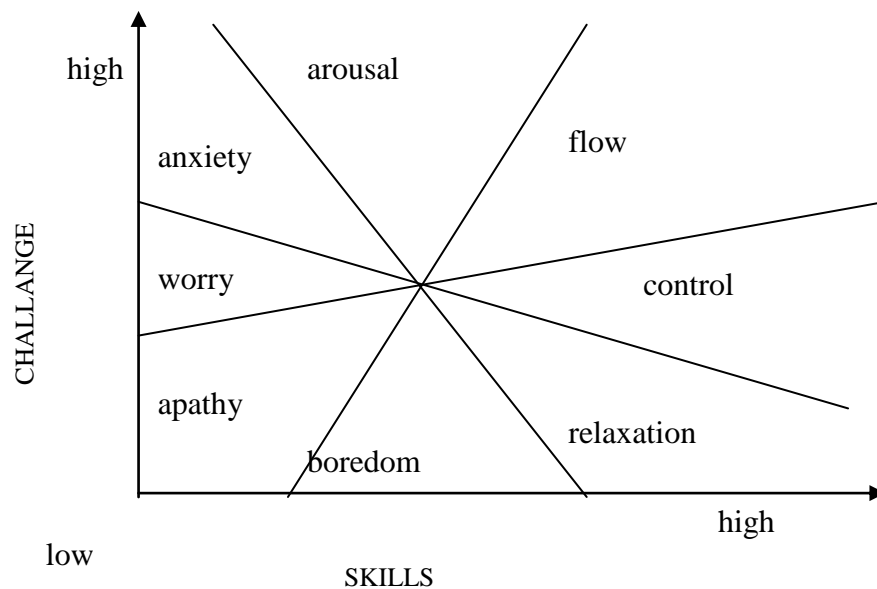
Figure 13 The relationship between challenges and skills (Csikszentmihalyi 1998)

May (1979) also considers that anxiety has a positive role. He says that confrontation with anxiety can relieve us from boredom, sharpen our sensitivity, and assure the presence of the tension that is necessary to preserve the human condition. The presence of anxiety indicates vitality. Like fever it testifies that a struggle is going on within the personality (May 1979, p.xx). May considers that the view that mental health is living without anxiety is illogical. Although he admits that anxiety can be destructive, he considers that the constructive part of anxiety ensures our development, it does not let us relax and be complacent; it is the driving force of our intellect. Only if its message is ignored, does anxiety become destructive. If we still do not take the steps to resolve the problem, depression sets in. May considers that anxiety is the experience of Being affirming itself against Nonbeing (aggression, fatigue, boredom and ultimately death) (May 1979, p.xxi). This may come from

Hamlet's question of 'being or not being', of the choice of suffering in the mind or taking arms against a sea of troubles and by opposing ending them which can be seen as another attempt to ignore anxiety.

May (1979) considers that the results of his studies support the view that personalities of higher intellect and originality experience more anxiety and have to learn to live with it. If they do not manage to develop efficient ways of dealing with their anxiety and start avoid situations causing anxiety they have to accept the impoverishment of their personality.

Liddell (1949) has similar views, he says that the human capacity for imaginative reality-testing, for dealing with symbols and meanings are all processes that are intertwined with our ability to experience anxiety. He proposes that the capacity to experience anxiety and the capacity to plan are two sides of the same coin and that anxiety follows intellect like a shadow. The reference to planning as being part of the anxiety experience has been also investigated by Eysenck (1992), see section (4.1.3). This suggests that anxiety experienced while using a language could also interact with the test-takers' use of planning strategies.

May (1981) considers that normal anxiety is a reaction proportionate to the threat, it does not involve repression and can be confronted constructively at a conscious level. If anxiety is dealt with, its causes are located and the problem solved, anxiety becomes the instrument of growth: *All growth consists of the anxiety-creating surrender of past values as one transforms them into broader ones, it consists of giving up immediate security in terms of more extensive goals* (May 1981, p. 246). According to May it is the soundness of the goal system that ensures the ability to deal with anxiety: mature values transcend the immediate situation and allow us to see it in context and to locate the reasons for anxiety; the more mature the values, the less

it matters whether they are satisfied; it is the holding of the values that produces satisfaction and security.

Mowrer and Ullman (1945) also view anxiety as playing a constructive role in human development. They consider that anxiety is not the cause of personal disorganisation; rather it is the expression of such a state, anxiety is the striving of the total personality to return the repressed and reestablish unity, harmony, oneness and health.

4.1.1.2 Neurotic anxiety: narrowing of awareness

This section will examine the causes and effects of neurotic anxiety and the process of development of the neurotic form of anxiety.

4.1.1.2.1 Causes of neurotic anxiety

Neurotic anxiety according to Kierkegard (1849) is a more constrictive and uncreative form of anxiety. He considers that neurotic anxiety is a result of the individual's failure to move ahead in situations of normal anxiety. We avoid challenge by ignoring it or forgetting it; as a result what we get is 'anxiety without any cause', which sometimes can even get misplaced (we start having irrational fears of height or darkness) which entangles us even further.

Kierkegard's suggestion could be used to explain the general anxiety at schools about the final exams; this could in fact be a misplaced fear of the unclear future (when difficult questions have to be answered, such as: what will I do next, what university shall I enter, where will I find a job).

Lewis (1996) also sees anxiety as a temporary state, whose aim is to develop strategies to reduce or eliminate itself. He sees anxiety development as a process consisting of following steps:

1. anxiety couples with painful expectancies to focus attention on helplessness and vulnerability

2. plans, goals and scripts that facilitate escape are highlighted

3. these produce action tendency to escape from the painful subject

4. if escape is impossible, the system moves to a defensive appraisal (or appraisals) of the immediate situation that might produce positive emotions (for example, don't blame me, it's all your fault!)

5. the system chooses the appraisal that produces the most positive emotion and builds its own script or action plan or schema that guides further actions

6. the initial appraisal is avoided and plays a diminishing role, but the conflict or anxiety now has become an indirect source of positive emotions.

In point 4 we can see explicitly how emotion and cognition interact in deciding on the action plan: should we move forward in spite of threat and hope for success or withdraw and find somebody to blame or try to 'forget' about the problem. The cases when people decide to withdraw and develop a habit of avoiding solution of inner conflicts lead to neurotic anxiety.

It is not always that we do not want to find a solution. It is also possible that a solution is impossible to find. May (1979) considers that the source of neurotic anxiety is a conflict between several threats that are all important and are juxtaposed to each other. As a result the individual cannot deal with one of them without being confronted with another and the individual feels threatened whichever way he or she turns.

LeDoux (1999) explains the mechanism of how our incapability to solve problems develop into neurotic anxiety. He says: *innate emotional reactions occur when the amygdale is turned on. In contrast, for avoidance, the brain has learned some*

response that can be performed in the presence of a learned trigger that short-circuits the innate response. These responses, once learned, prevent emotional arousal. They become habits, ways of automatically responding to stimuli that routinely warn of danger (LeDoux 1999, p. 262).

Thus the fear, instead of being addressed and resolved is just avoided and can develop into an anxiety disorder. Gray (1990) also talks about anxiety as a learned reaction: anxiety is a state elicited by frustrative nonreward, or novel stimuli. Normal reaction to such stimuli consists of an inhibition of ongoing behaviour, an increase in the level of arousal (so that the next occurring response is performed with extra viogour and speed) and increased attention to the environment (stop, look and listen and get ready for vigorous action).

4.1.1.2.2 Effects of neurotic anxiety

Kierkegard (1941) describes neurotic anxiety as an exceedingly painful experience: *No Grand Inquisitor has in readiness such terrible tortures as has anxiety, and no spy knows how to attack more artfully the man he suspects, choosing the instant when he is the weakest, nor knows how to lay traps where he will be caught and ensnared, as anxiety knows how, and no sharp witted judge knows how to interrogate, to examine the accused, as anxiety does, which never lets him escape, neither by diversion nor by noise, neither at work nor at play, neither by day or by night* (Kierkegrad 1844/1944, p. 139). Thus the experience of anxiety has different psychological as well as physiological effects.

Goldstein (1938) suggests that as anxiety increases, the real cause of anxiety retreats, it disintegrates until we start being afraid of everything, we withdraw from the world, and any useful action or perception is suspended. Goldstein says that the relationship between the self and the world is precisely what breaks down in anxiety. As a result,

anxiety appears as an objectless phenomenon that finally leads to the disintegration of self.

Power and Dalgliesh (1998) quote American Psychiatric Association classification system (1994) description of Generalised Anxiety disorder that clearly brings out the effects of anxiety: *Trembling, twitching or feeling shaky, muscle tension, aches or soreness; restlessness; easily fatigued; shortness of breath; palpitations and tachycardia; sweating or cold, clammy hands; dry mouth; dizziness or light-headedness; nausea or abdominal distress; flushes or chills; frequent urination; trouble swallowing (lump in a throat); feeling keyed up or on edge; exaggerated startle response; difficulty concentrating or blank mind; trouble falling or staying asleep; irritability* (Power and Dalgliesh 1998, p.208).

Eysenck considers that when we address the question of learning, the cognitive aspect of anxiety is most important in explaining its changes over time: *An individual's previous experience in stressful situations of a particular type will influence the information which is stored in long term memory. This will in turn have an impact on that individual's susceptibility to anxiety in that type of situation* (Eysenck 1992, p.42).

The attention of the anxious person is focused on the negative aspects of the situation which can be observed in his or her behaviour: *the anxious person is hypervigilant, constantly scanning the environment for signs of impending disaster or personal harm and becoming oblivious to stimuli that indicate that there is no danger* (Beck and Emery 1985, p.31).

4.1.1.3 Trait versus State anxiety

Eysenck (1992) says that most theorists draw a distinction between trait and state anxiety although the manner of differentiation is not the same:

1. the distinction between trait and state anxiety is temporal: states last for relatively short periods of time while traits remain stable considerably longer (Allen and Potkay 1982, quoted in Eysenck 1992)

2. the distinction between state and trait anxiety is the distinction between disposition to a predictable response in appropriate circumstances (trait anxiety) and the occurrence of a single observable event (state anxiety) (Ryle 1949 and Fridhandler 1986, quoted in Eysenck 1992).

3. According to Eysenck (1992) trait anxiety is a personality dimension, whereas state anxiety is subjective, consciously perceived feelings of tension and apprehension experienced in particular circumstances.

Thus Eysenck considers that it is consciousness that distinguishes between state and trait of anxiety. We are conscious of having state anxiety, but trait anxiety is part of our personality, which we might not even be aware of. I will return to state and trait anxiety division in section 4.2. on test anxiety.

4.1.2 Fear

Power and Dalgliesh consider that anxiety is the result of unresolved fear, therefore I will address the question of fear in order to understand the basis of anxiety, in spite of the fact that fear differs from all the other anxiety correlates in that it is usually treated as an instantaneous automatic reaction.

4.1.2.1 Fear as part of the survival kit

LeDoux (1999) considers that the fear system's main result is not the experience of fear, but rather the responses to the demands that maximize the probability of surviving the dangerous situation in the most beneficial way. The system operates independently of consciousness as it represents the operation of brain systems that

have been programmed by evolution to deal with danger in routine ways: when the brain detects danger it also sends messages through the nerves to the autonomic nervous system and to the bodily organs and adjusts the activity of those organs to match the demands of the situation: *Nerves reaching the gut, heart, blood vessels and sweat and salivary glands give rise to the taut stomach, racing heart, high blood pressure, clammy hands and feet and dry mouth that typify fear in humans,* (LeDoux 1999, p.128). LeDoux considers that the first automatic reaction of fear is to buy time so that the brain can start planning the most appropriate action.

4.1.2.2 Fear as a motivator

Power and Dalgliesh (1998) propose that healthy fear is functional: it prepares and empowers the individual for fight or flight in the presence of danger (making it in a way a motivator): in retrospect it also helps people focus on the work at hand, for example, it makes you study harder for an exam. It gives people the push they need to concentrate for sometimes as long as a year before the event. They consider that this motivation is achieved through a complex action:

1. physiological arousal: butterflies in the stomach, tension in the muscles, perspiration, dry mouth etc.

2. thoughts about being in danger (cf. 'cognitions', Spielberger et al 1978 in section 4.2.)

3. behavioural components - avoiding the feared stimulus, running away or fighting. (Power and Dalgiesh 1998, p.200)

According to Power and Dalgliesh, the appraisals that cause fear are usually complex:

1. appraisal of the event as incompatible in some way with the existing goals or models of the self, the world or others

2. the appraisal of threat: there is a chance of future non-completion of values

3. evaluation of whether the threat is unwanted

4. evaluation of whether the threat can be avoided (according to Power and Dalgliesh the fact that we have chosen to take part in a situation provoking anxiety ourselves does not change anything: the reactions are the same as in unwanted fear situations)

5. evaluation of one's own physiological reactions.

Power and Dalgliesh have included the appraisals of the physiological reactions in generation of fear because people react to them in different ways: there are people who interpret the bodily reaction as a norm and forget it during the work so that the bodily reactions do not interfere with the quality of the activity. But there are also cases when people can misinterpret as catastrophic bodily reactions such as sweating, feeling flushed, changes in breathing and heart rate, intestinal discomfort, muscular tension and butterflies in the stomach. According to Clark (1988) these fears can produce panic attacks, which can be called "fear of fear" (Power and Dalgliesh 1998 p.211). This feeling can become a vicious circle: the more afraid one becomes, the more intense the feared feelings as the panicker becomes hypervigilant and repeatedly scans the body for evidence of bodily sensations. On the other hand the panickers develop avoidance strategies and as a result never attempt to go through the feared experience and never persuade themselves that it can be accomplished (for example an exam can be taken and passed) so that the panic is groundless.

Fear as motivator in both its effects as well as causes is close to what May (1979) calls 'healthy anxiety' (see section 4.1.1.1).

4.1.3 Worry

Although worry and fear are similar in that they both prepare us for an unexpected event that can have a negative outcome, there is a crucial difference between the two: fear is treated as an instantaneous automatic reaction, while worry is considered pathological if it continues in an uninterrupted manner for 6 months. The temporal difference between the two determines also the difference in the contents: fear is a reaction to something that has just happened, while worry involves planning and preparing for action, so the cognitive element is more important in worry.

4.1.3.1 Worry as cognition

Burkovec et al (1983) define worry as a chain of thoughts and images, negatively affect-laden and relatively uncontrollable. Worry is an attempt to engage in the mental problem solving in an issue that contains the possibility of negative outcomes. It is a primarily verbal activity and has the function of reducing the generation of threat related imagery and the subsequent psychological activity.

Power and Dalgliesh (1998) say that for an average person worry is a necessary process as it allows us to review options, construct schematic models of possible outcomes and make suitable plans.

So two opposing views of worry coexist in psychology: worry with a positive role (as preparation for something) and worry with a negative role (as an uncontrollable self-evaluative stream of thought that spoils our performance).

4.1.3.2 Worry as trait anxiety

Eysenck (1992) studied worry as part of trait anxiety. According to his findings the correlation between trait anxiety and worry is +.64 and he considers that worry is the principal component of trait anxiety. Having carried out factor analyses of his Worry

questionnaire he found that a factor formed by variables relating to general social-evaluative concerns, personal relationships, financial concerns, personal fulfilment and appearance accounted for 52% of the variance. According to Eysenck (1992) worriers have elevated evidence requirements, that is, in order to make a decision they must spend more time weighing up the relevant information in their memory and in the environment. They must be absolutely sure that they are doing the right thing before the decision can be made.

The theoretical framework of the worry system includes not only hypervigilant scanning of the environment and emotional sensitivity but also increased arousal and self-focused attention. This forms a vicious circle of worrying that leads the individual to construct negative models of future events. This, however, is not all. Eysenck also considers that if we choose the right strategies we can find a way to remove the threat.

4.1.3.3 Functions of worry

Just as worry was defined by different researchers as a positive and as a negative phenomenon, its functions can be also defined both as positive and negative. Eysenck proposes the following positive functions of worry

1. alarm function: introducing threat related information into consciousness

2. prompt function: the threat related information continually represented to awareness so that the cognitive system can somehow resolve it

3. preparation function: the individuals can anticipate future situations and conceptualise resolutions .

Barlow (1988) proposes a negative scenario of the role of worry:

1. Certain situations or unexplained arousal lead to the evocation of anxious propositions which are stored in memory and produce a state of negative affect

2. Negative affect causes an attentional shift from the external environment to self-evaluative focus

3. Self-evaluative focus leads to a state of intense arousal

4. Increased arousal activates an apprehensive cognitive schema which produces a perceived inability to control future situations and leads to attentional narrowing

5. Activation of the apprehensive cognitive schema produces worry

6. Worry leads to a dysfunctional performance.

Eysenck (1992) and Barlow (1988) present two opposing views of worry: as a necessary stage in preparation for an important event that might have a negative outcome, or it is just the opposite: an uncontrollable chain of images of our previous failures that destroys our performance. The difference of the two opposing views is determined by

- the time and place of the worrying experience: if this is before the event when we have enough time to think through previous experiences (preferably positive) and find a solution that will allow us to achieve our goals in the best possible way, then this improves our performance, but if we try to deal with the task and worry at the same time, this can spoil our performance

- the severity of the problem, there are some problems which cannot be solved however much time is available to prepare for them.

4.1.4 Stress

Fear is instantaneous, worry is a long period experience, but they are both more or less straightforward phenomena. Stress differs as it can be not only experienced for both long and short periods, but it can also be caused by a cognitive appraisal of the situation as well as physical aspects of the environment (for example cold or heat). Nevertheless I will review the literature that examines stress because its findings on

the physiological changes we undergo when we find ourselves in a demanding situation has implications for language testing.

4.1.4.1 Approaches to the concept of stress

Most of the theories discussing 'anxiety' refer to 'stress' as its basis: for example, Eysenck (1992) considers that '*high* worriers' are more stressed '*low* worriers'. May (1979), however, considers that the term 'stress' comes from engineering and physics. May says: *It seems to have become popular in psychology because it can be defined readily, handled easily, and ... measured satisfactorily, all of which are difficult with* anxiety...*Anxiety, on the other hand, is uniquely bound up with consciousness and subjectivity.... Anxiety is how the individual relates to stress, accepts it, interprets it.* (May 1979, p.96-9).

According to Cassidy (1999) in the 1960s and 1970s stress was viewed as a stimulus: something which occurs in the environment and makes a demand on the person: change of temperature, lack of food or an aggressive roommate, they were all called stressors. The main focus of research at that time was on identifying and classifying stressors, and discovering what makes a stressor stressful. Later the concept of stressors was developed further to connect it with response: the central aspect is control, if we can predict or control the threat or loss, we will not perceive it as a stressor. On the other hand, stressors that are fully predictable become boring and can become stressful as well. It is the element of personal meaning or cognitive appraisal that has the last say in determining the impact of stress.

In the 1980s the main focus of research was the form of behavioural, emotional or physical responses (irritability, lack of energy, sleeplessness, headaches and digestive problems) to the outer demands. The findings were overwhelming: it is difficult to find a physiological or psychological illness that cannot be linked to stress. This

conclusion is in striking contrast to Caccioppo et al (1999) proposal discussed above that negativity bias is beneficial for the survival of a species.

Recently the approach has shifted to what Cassidy (1999) calls a 'transactional model of stress'. He views stress as a transaction between a person and his or her environment and incorporates both stimulus and response perspectives as a part of the process that has to be viewed as changing in time. The process starts with demands made by the situation, our response and the consequences of the transaction between demands and responses. The person in this model is seen as an active agency in this process. As a result stress is viewed in terms of a fit between the person and his or her world. Levi (1987) says that *when the fit is bad, when needs are not met, or when abilities are over or under taxed, the organism reacts with various pathogenic mechanisms (cognitive, emotional, behavioural and/or physiological)* (Levi, 1987, p.24).

4.1.4.2 General Adaptation Syndrome Theory

Selye's General Adaptation Syndrome (G.A.S.) theory (1956) is based on optimal arousal theory (see section 2.2.2). Selye defines stress as a non-specific response of the body to any demand. All the agents of change increase the demand for adjustment to re-establish normalcy. This rise in requirements is independent of the specific activity that caused the increase and is non-specific (Selye 1956, p.128). To explain this Selye uses the *general adaptation syndrome* model, which consists of three stages:

1. Alarm reaction: organism's reaction (fight or flight) when it is suddenly exposed to diverse stimuli that the organism is not adapted to. This stage consists of two phases:

- Shock phase: the immediate reaction (loss of muscle tone, decreased temperature and blood pressure etc.)

- Counter shock phase: mobilisation of defensive phase during which the adrenal cortex is enlarged and the blood pressure increases. If it is successful the body is restored to homeostasis.

2. Stage of resistance: the full adaptation of the organism and the consequent improvement of symptoms, resistance to most other stimuli. As a result the body copes through the depletion of resources necessary for other bodily functions.

3. Stage of exhaustion: since adaptability is finite, exhaustion inexorably follows if the stressor is sufficiently severe and prolonged. Symptoms reappear.

Selye (1956) suggests that we must learn to recognise overstress (hyperstress), when we exceed the limits of our adaptability and under stress (hypostress) when we suffer from lack of self-realisation or boredom. Selye says that our goal should be to strike a balance between the equally destructive forces of hypo and hyper stress to minimise distress. To succeed, a person has to work for things that strengthen his homeostasis in the unpredictable situations with which life may confront him.

Selye says that although internal and external factors influence or even determine some responses we do have some control over ourselves. It is the exercise of this control, or the lack of it, that can decide whether we are made or broken by the stress of life (Selye 1956, p.143).

According to Cassidy the two main criticisms of Selye's model are that it ignores any psychological element and that it proposes that there is a common physiological response to all stressors.

4.1.4.3 Effects of stress

LeDoux's research in brain systems offers an explanation of how stress affects our performance: *in the face of stress, the amygdala keeps saying 'release the adrenal steroid hormone' and the hyppocampus keeps saying 'slow down'. Through multiple cycles through these loops the concentration of the stress hormones in the blood is delicately matched to the demands of the stressful stuation* (LeDoux 1999, p.240).

In the case of mild stress the efficiency of the memory is enhanced by the facilitatory effect of adrenalin (the so called 'flashbulb effect'). However, if stress persists, the control of hippocampus begins to falter; it starts interfering with the ability of the hyppocampus and this leads to memory failure.

The amygdala's ability to process and store information, however, is not only not interfered with, but can even be enhanced by stress, so it is perfectly possible for a human being not to have explicit memories of a traumatic experience, but to have emotional memories *through amygdala mediated fear conditioning* (LeDoux 1999, p. 245). In later life these can develop into unconscious sources of anxiety that cannot be explained and therefore controlled. This can explain why we sometimes are frightened for no apparent reason; we understand that there is no reason to be afraid and still we cannot help ourselves.

Because stress lowers the threshold of anxiety it can make a mild case of fear turn into a serious problem. For example, we may not normally be afraid of height, but when we are distressed by some unpleasant event, we my project our distress to being afraid of height and it may become a problem. This means that all the anxieties that we have had can be easily triggered when we are under stress.

If we compare the theories on anxiety and its correlates, we find that each of the concepts (fear, worry and stress) add to our understanding of anxiety as an overarching term that incorporates the narrower terms (see Table 5): 'fear' serves as a

starting point, 'worry' is a tool that can resolve a complicated problem, 'stress' is a mental and bodily reaction to the presence of an unresolved problem.

Table 5 Comparison between anxiety and its correlates

Concept	My interpretation
Anxiety	'Healthy anxiety' appears as unrest whenever we visualise a possibility to achieve a goal; it energizes and mobilizes our resources to move forward. 'Neurotic anxiety' (shrinking of consciousness) is a reaction to our inability to overcome an obstacle that keeps us from reaching a goal.
Fear	'Fear' has a clear object, it is short-period phenomenon and prepares and empowers the individual for 'fight or flight' in the presence of danger.
Worry	'Worry' is an attempt to engage in the mental problem solving in an issue that contains the possibility of negative outcomes, it can be a long term activity (up to 6 moths); if unresolved can lead to neurotic anxiety.
Stress	'Stress' is the organism's reaction to a bad fit between the person and his or her world with various pathogenic mechanisms (cognitive, emotional, behavioural and/or physiological).

4.2 Test anxiety

The concept of test anxiety has been developing beside general anxiety research. This branch of research studies the causes of anxiety in testing environment and effects on test performance.

In the 1950s test anxiety was conceptualised as drive oriented: Sarason and Mandler (1952) proposed that the testing situation evoked both learned task drives and learned anxiety drives. He suggested that there was both task relevant anxiety (increasing the level of performance) and task irrelevant anxiety (decreasing the level of performance).

In the 1970s test anxiety theory moved from drive oriented to cognition-oriented theories. Sarason (1960) and Wine (1971) found in their experiments that high test-anxious persons were self-dissatisfied and instead of dealing with the tasks, were spending time on self-ruminations. Spielberger et al (1978) elaborated the approach further by conceptualising test-anxiety as consisting of separate cognitive and physiological variables. Recently, the stress has shifted to the interaction between

learner strategies and the level of test-anxiety (Paulman and Kennely 1985) and competition in general (Reeve and Deci 1996). Here I will concentrate on the two most recent approaches (cognitive and skills deficit approaches) as they explore the concept, the causes and effects of test anxiety.

4.2.1.1 Cognitive interference theory

Sarason (1978) considers that in psychology two approaches in the research of anxiety must be distinguished:

1. study of the observable and recordable events (accelerated heart and breathing rates) that characterise anxiety
2. study of the hypothetical state or a person's interpretation of the situation.

In the case of test-anxiety what matters is a person's perception of the situation. If the test is not important for the test-taker, he or she will not be anxious and the performance will not be affected by anxiety.

Sarason considers that the information provided by the situation is processed by us in distinctive ways that are determined by our cognitive appraisals. This is part of our information processing system. Sarason proposes that cognitive appraisal includes categorising and interpreting events. It involves a sort of memory search and a weighing of alternatives as a result of which a response is selected that best fits the situation.

The most commonly observed ways of handling a situation are the following:

1. a task oriented problem-solving approach
2. avoidance of stressful situations
3. defensive distortion of the situation through projection, rationalization, denial

4. anxiety: a type of cognitive response to perceived inability to handle a challenge, marked by self doubt, feelings of inadequacy and self-blame (Sarason 1978, p.195).

Figure 14 Sarason's (1978) test anxiety development steps

```
┌─────────────────────────────────────────────────────────────────┐
│       Step 1:The situation is seen as difficult, challenging      │
│                         and threatening.                          │
└─────────────────────────────────────────────────────────────────┘
                                  │
                                  ▼
┌─────────────────────────────────────────────────────────────────┐
│   Step 2:The individual sees himself or herself as ineffective in │
│          handling or inadequate for the task at hand              │
└─────────────────────────────────────────────────────────────────┘
                                  │
                                  ▼
┌─────────────────────────────────────────────────────────────────┐
│    Step 3:The individual focuses on undesirable consequences of   │
│            personal inadequacy                                    │
└─────────────────────────────────────────────────────────────────┘
                                  │
                                  ▼
┌─────────────────────────────────────────────────────────────────┐
│   Step 4:Self-deprecatory preoccupations are strong and interfere │
│          or compete with task relevant cognitive activity         │
└─────────────────────────────────────────────────────────────────┘
                                  │
                                  ▼
┌─────────────────────────────────────────────────────────────────┐
│        Step 5:The individual expects and anticipates failure      │
└─────────────────────────────────────────────────────────────────┘
```

Sarason proposes that anxiety is one of the ways of handling a situation, but then goes on to define it as a cognitive response to inability to handle a situation. Sarason also says that anxiety is a cognitive response, without any reference to emotional products (sweating, rising blood pressure and heart beat), which are part of the experience of anxiety and can also be interpreted by the test-taker as a sign of danger. He characterises an anxiety response as a cognitive process that involves evaluation of situation, one's own ability (or inability) to deal with the situation and projections into the future and how it will affect one's goals (see Figure 14). The first two steps are similar to Power and Dalgliesh's (1998) emotion production model, but Steps 3,4 and 5 and represent the reaction to negative evaluation.

Sarason says that *every teacher knows students who, while quite able and bright, are virtually terror stricken at exam time. (...) Whereas most students read test questions*

*and proceed to answer them, highly anxious individuals find themselves thinking about the consequences of failure and how much better prepared the other students are (*Sarason 1978, p.197).

Sarason considers that test could be one of the classes of situations that evoke anxiety and that it may be traceable to early experiences for which the child was not cognitively ready. He sees test-anxiety as a combination of worry and physiological reaction patterns that accompany it and because worry is a cognitively demanding activity it will be expected to interfere with performance on complex tasks when the evaluation dimension is emphasised.

In 1980 Sarason redefined test-anxiety, and connected anxiety to both evaluation and attention. He proposed that proneness to self-preoccupation and, most specifically, to worry over evaluation, is a powerful component of what is referred to as test-anxiety. Since worry interferes with attention to the task performance, the more attention is necessary to accomplish the task successfully, the more performance will be affected by the test anxiety. In such cases it is control over one's thoughts that is the deciding factor in successfully meeting a particular situational challenge.

Eysenck (1992) considers that Sarason exaggerates the importance of self-preoccupation and worry and although his experiments support the view that highly anxious individuals perform a task less well than individuals with low anxiety level, there are also several studies in which highly anxious individuals did not perform less well than individuals with a low level of anxiety. Eysenck quotes Blankstein, Flett, Boase and Toner's (1990) findings where "high test anxious subjects" had more negative thoughts about themselves than "low test-anxious subjects", but where the two groups did not differ in their performance.

Eysenck (1992) also criticises Sarason's oversimplified accounts of the interaction between anxiety and task difficulty. My own feeling is that Eysenck did not go far

enough as he should have also referred to the lack of any acknowledgement of the products of the emotional system in Sarason's discussions of anxiety. Eysenck only talks about appraisal and cognitions and does not connect these to the arousal caused by conscious or unconscious fear (see Power and Dalgliesh 1998).

Eysenck's attitude can be explained by his focusing on anxiety as a conditioned response; he disregards the role of unconscious processing. LeDoux considers that Eysenck felt that in the case of anxiety if we eliminate the symptoms we eliminate the neurosis (LeDoux 1999, p.235).

4.2.1.1.1 Worry and emotionality

Liebert and Morris (1967) conceptualise test anxiety as consisting of two major components: worry and emotionality.

Worry is described as a primarily cognitive concern about the consequences of failure (negative expectations and concerns about oneself, the situation at hand and the potential consequences). Worry is aroused and maintained by situational factors that influence one's cognitive evaluations. Emotionality is defined as consisting of autonomic reactions evoked by evaluative stress. It refers to one's perception of the physiological-affective elements of the anxiety experience (indications of autonomic arousal, nervousness and arousal). It does not decrease significantly from the beginning to the end of the test. It is defined by the setting of the classroom, the administrators passing out tests and student conversations about the test (Morris, Davis and Hutchings 1981, p.542). In 1967 Liebert and Morris published evidence that worry affected performance negatively, while emotionality was unrelated to task performance.

Morris and Fulmer (1976) compared the change of the level of worry in students taking an examination in the usual setting with students receiving an item by item feedback and found that both worry and emotionality levels decreased.

The two components are considered to be conceptually independent because they are aroused and maintained by different situational conditions. Morris, Brown and Halbert (1977) found that each component could be aroused independently if before carrying out an intellectual task students were shown a video-recording of peers exhibiting behaviour and verbalisation that indicated either worry or emotionality. Deffenbacher (1978) found that subjects receiving what he calls "ego-involving" instructions (stressing importance and their personal responsibility for their achievement) had higher worry scores than their peers who received reassuring instructions. The emotionality scores, however, were similar in both cases.

There are also findings that suggest that the two variables in some ways are closely related. Smith and Morris (1976) analysed the effects of music played during a test and found that soothing background music did not reduce either emotionality or worry, whereas lively and stimulating music increased both.

Sassenrath (1964) carried out factor analyses of test-anxiety using a Test Anxiety Questionnaire (Mandler and Sarason 1952) and identified seven factors. Four of them involved confidence (individual intelligence testing, group intelligence testing, before course examinations score and during course examinations score), so Spielberger et al (1978) considers that they can be conceptualised as related to 'worry'. Two of the factors (perspiration and heartbeat during the examination) can be related to emotionality. 'Avoidance of intelligence testing' was also singled out as a separate factor.

Richardson, O'Neil, Whitmore and Judd (1977) had similar results. They used factor analyses to analyse the results on the Test Anxiety Scale (Sarason 1958) and found two factors:

1. cognitive concern and worry about oneself and one's performance

2. variety of physical and emotional consequences of intense worry which Richardson et al interpreted as a distinct set of negative emotional reactions to tests.

4.2.1.1.2 Test Trait and State anxiety

As Eysenck (1992) does with general anxiety, Spielberger et al (1978) conceptualise test anxiety as being formed by state and trait anxieties. State anxiety refers to transitory experiences of tension, apprehension and activation of the autonomous nervous system in specific situations. Trait anxiety refers to a personality variable of anxiety proneness, the tendency to experience state anxiety in a variety of situations.

Spielberger et al propose that test anxiety is a situation specific trait: *although examination situations are stressful and evoke state-anxiety reactions in most students, the magnitude of the state anxiety component will depend on a student's perception of a particular test as personally threatening* (Spielberger et al 1978, p.168).

Spielberger et al connect stress and anxiety. They propose that stress refers to the stimulus properties of situation that are characterised by some degree of objective physical or physiological danger. According to their theory stress leads to a perception of danger and this increases state anxiety.

On the other hand, they consider that irrespective of the presence of real or objective danger (stress) a person who perceives a situation as dangerous or threatening will experience an increase in anxiety. The appraisal of a particular situation as threatening will depend upon

1. the characteristics of the situation

2. the individual's past experience with similar situations

3. memories and thoughts that are evoked by the situation.

Spielberger et al (1978) developed a new instrument for measuring test anxiety which he called the Test Anxiety Inventory (a relatively brief self-report scale) and carried out factor analyses to compare the worry and emotionality components of test anxiety with state and trait anxiety.

After the analyses of the results, their conclusion was: *It is not possible to classify test anxiety scales definitely as either measures of trait or state anxiety, but the bulk of evidence is consistent nevertheless with the assumption that test-anxiety is a situation specific measure of anxiety proneness in test situations* (Spielberger et al 1978, p.186).

Humphrey and Revelle (1984) are also interested in the effects of trait and state anxiety on information processing. They, however, consider that it is not only worry or the cognitive components that affect the performance. They proposed that state anxiety increases the level of avoidance motivation, largely because of worry and other self-concerned thoughts. On task effort and arousal both increase sustained information transfer. At the same time arousal also reduces short-term memory through trait anxiety. Because tasks differ in terms of the degree of involvement of short-term memory, the effects of trait anxiety vary from task to task. As arousal increases there is an increase in sustained information transfer and a decrease in short-term memory and it is difficult to predict whether anxiety will facilitate or impair performance. Eysenck (1992) considers that this theory lacks control and compensatory mechanisms and processes which monitor and adjust the functioning of the information processing system. He considers that anxiety instead of producing

avoidance motivation, as argued by Humphreys and Revelle (1984), motivates students to make an additional effort.

4.2.1.2 Skills deficit approach

The cognitive and attentional approach to test anxiety was followed by the skills deficit approach. Kirkland and Hollandsworth (1980) even proposed substituting the concept of 'test-anxiety' for one of 'ineffective test taking'. Their proposal was based on their own, as well as Culler and Holahan's (1980), research findings that study habits and exam-taking skills of highly test-anxious students are frequently poor and that anxiety reduction seldom improves performance on cognitive tasks (Allen et al 1980). If, however, study skills training is combined with anxiety reduction exercises it is much more successful (DiTomaso 1981, Allen et al 1980). This suggestsa link between cognitive and affective variables.

Paulman and Kennely (1984) however, consider that test anxiety is associated with impairment in information processing capacity that is apparently independent of both ability and exam-taking skill. Their proposal is based on multivariate analyses of several instruments:

1. Sarason's Test Anxiety Scale (1978),
2. Spielberger et al's State trait anxiety inventory (1978),
3. the Raven test (Raven 1965) that requires effortful cognitive processing and
4. the Digit Span test to assess working memory and information processing capacity.

Their main finding was that test-anxiety and exam skills have separate roles in mediating cognitive performance.

4.2.1.3 Effects of test anxiety on performance

Cognitive interference and skills deficit approach theorists agree that the cause effect relationship is interactional: a low level of performance leads to higher level of

anxiety, which in its turn leads to a further escalation of anxiety. Nevertheless, their arguments focus on different aspects of performance: cognitive interference theorists are more concerned with the level of performance, while skill-deficit theorists are more concerned with providing ideas on repairing strategies. Therefore I will include a short review of the findings of both the approaches.

4.2.1.3.1 Findings of Cognitive Interference Theory

Research on Cognitive Interference theory provided evidence that test-takers who were preoccupied with self-evaluation during the examination produced lower level performances:

1. Wine (1982) found negative cognitive-attentional effects on performance

2. Houston 1977 and Morris and Engle (1981) found that performance suffers because of a misdirection of attention from the task at hand

3. Sarason and Stoops (1978) found that self-preoccupation was the most influential element in the experience of test anxiety

4. Hollandsworth et al (1979) observed that the major differences between high and low test-anxious students lay not in their level of physiological arousal during tests but in their cognitive reactions

5. Deffenbacher (1978) found that unlike emotionality, worry scores on both a task generated interference scale and a worry scale were highly negatively correlated to performance

6. Morris and Liebert (1969) selected subjects on the basis of high and low trait worry and trait emotionality levels and exposed them to differing degrees of stress during their performance of Wechsler adult intelligence scale sub-tests. They found that the typical anxiety and stress interaction effect on performance was accounted for by the interaction between stress and worry. Emotionality did not have a contributing effect.

Nevertheless, Morris et al (1981) also admit that there is a problem concerning cause and effect. They say that worry may be a reflection of concern about accurately perceived past and present performance difficulties rather than being a cause of poor performance. Evidently both factors may operate simultaneously, but, according to Morris et al (1981), neither of them involves emotionality directly.

4.2.1.3.2 Findings of skills deficit theory

The same bi-directional view on cause-effect relationship is supported by the 'skills deficit' approach. Paulman and Kennely (1984) propose that low ability leads to achievement anxiety, which then results in poor study habits. Less material encoded during study eventually leads to poor test performance. This is further aggravated by anxiety-induced retrieval difficulties from worry within the test situation.

Test anxiety seems to be interfering with both the encoding and retrieval of information. The exam-skilled, test-anxious students, however, are high achievers with efficient study habits. They presumably have studied hard during the course (this could be interpreted as an indirect reference to fear as motivator (see Eysenck 1992)). The effective strategies for study and exam skills facilitate the control of attention toward the task demands. As the exam approaches, they may experience difficulties in concentration. *During the actual examination they may be unable to recall, organise, express what they have learned. Failure feedback, or even self-perceived failure, may then initiate a positively increasing cycle of worry and performance setback similar to that experienced by poorly skilled students.* (Paulman and Kennelly 1984, p.286).

Paulman and Kennely consider that high-anxious skilled test-takers have to be identified and offered special training that will concentrate on stress reduction techniques as well exam taking skills. Paulman and Kennely concluded from the analyses of their data that:

1. test anxiety is associated with an impairment in information processing capacity that is apparently independent of both ability and exam taking skill

2. both poor exam skills and high anxiety generated cognitive interference during problem solving that was negatively related to performance (see also Eysenck 1979)

3. the highest test-anxiety was experienced by high test-anxious unskilled individuals which manifested itself in cognitive interference at the lowest performance level.

Having examined the concept of anxiety in experimental and social psychology and found that most of the findings suggest that anxiety has a detrimental effect on our performance I will now turn to anxiety in foreign language acquisition and use situation.

4.3 Foreign language anxiety

Traditionally foreign language anxiety was considered as social and situational anxiety (Gardner and MacIntyre, 1991 and Horwitz, 1986) caused by the language users' inability to express themselves adequately. Horwitz et al (1986) consider that the essence of foreign language anxiety is the threat to the individual's self-concept caused by the inherent limitations of communicating in an imperfectly mastered second language. Schlenker and Leary (1985) also consider that the intimate relationship between self-concept and self-expression makes foreign language anxiety distinct from other academic anxieties. Recent research shows that foreign language anxiety can be studied not only as a social and situational anxiety, but also

1. as a linguistic phenomenon: Saito et al (1999) and Cheng et al.(1999) propose that instead of talking about foreign language anxiety we should be talking about foreign language reading, speaking, writing and listening anxiety.

2. as a cognitive phenomenon: Onwuegbuzie, Bailey and Daley (2000) suggest that there is foreign language input, processing and output anxiety.

 Here I will present an overview of the traditional approach as well as the more recent approaches to studying foreign language anxiety.

4.3.1   Situational approach

Situational approach to foreign language anxiety has been popular for a long time and as a result, applied linguists have developed several branches within the field of research of foreign language anxiety. They distinguish between

1. foreign language classroom anxiety and foreign language test anxiety (Horwitz 1986 and Gardner and MacIntyre1991)

2. foreign language test state and trait anxiety (MacIntyre and Gardner 1991)

3. harmful versus helpful anxiety (Madsen 1982).

Research on causes and effects of foreign language anxiety and observable signs of anxiety (Oxford 1999) is also well developed. In this section I will review all these aspects of foreign language anxiety research to lay grounds for my own study.

4.3.1.1 Foreign language classroom versus foreign language test anxiety

Gardner and MacIntyre (1991) developed a study to investigate the constructs of different types of anxiety. They used 23 different sets of questions to measure different state anxieties (for example, French, English and general test anxiety, audience sensitivity, fear of negative evaluation, novel situations anxiety and social anxiety). Having carried out factor analyses they found three distinct factors:

1. a social evaluation anxiety factor based on the questionnaire results that examined the test-taker's anxiety during the test of mathematics, English and two social evaluation anxiety questionnaires,

2. a state anxiety factor based on a state anxiety scale, a novelty scale and a dangerous situation scale

3. a foreign language anxiety factor based on a French classroom facilitating, a debilitating, and a French use scale.

Thus according to Gardner and MacIntyre (1991), their results suggest that foreign language anxiety may be distinguished from other types of anxiety (factor 3) although there is some correlation with general test anxiety and state anxiety.

The interaction between language and test anxiety is discussed also by Horwitz (1986). She used the Foreign Language Classroom Anxiety Scale (FLCAS) to measure language anxiety. She found a correlation between test anxiety and foreign language anxiety of .56. This can be explained by the fact that teachers often use tests in the class thus creating anxiety among the learners; test anxiety combines with what Horwitz (1986) calls a threat to the individual's self-concept. This becomes part of learner's and later test-takers' foreign language use experience that will be activated every time the foreign language is used.

4.3.1.2 State versus trait anxiety

Oxford (1999) says that language anxiety can start in response to a particular situation or event (situational or state anxiety), but that it can also be a major character trait. It can start as transitory episodes of fear in situations where a learner has to perform in the target language, and ideally it diminishes over time. However, if repeated occurrences of fear cause students to associate anxiety with language performance, anxiety becomes a trait rather than a state. Once language anxiety has evolved into a lasting trait, it can have pervasive effects on language learning and language performance.

MacIntyre and Gardner (1991) also see foreign language anxiety development as a process. At the earliest stages of language learning, language anxiety is not present. If there is anxiety, it is either trait anxiety, novelty anxiety or even test anxiety. Besides, anxiety can be different for different learners (see Figure 15).
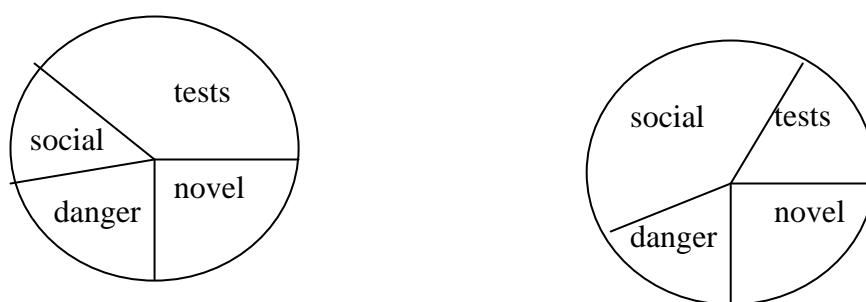
Figure 15 Components of trait anxiety on a fictional scale (MacIntyre and Gardner 1991)

After several classes the student forms attitudes to learning the new language. If these experiences are negative, foreign language anxiety may begin to develop. If they persist, it may become a regular occurrence connected with foreign language use. The student starts to expect to be nervous and to perform poorly. This causes cognitive interference and poor performance. As experience and proficiency increases, anxiety declines in what MacIntyre and Gardner call "a fairly consistent manner" (MacIntyre and Gardner 1991, p.111). In some cases, however, this may not happen. A student may, for example, get into a vicious circle where performance deficits can lead to nervousness and worry, which in their turn cause decreasing levels of performance.

4.3.1.3 Harmful versus helpful anxiety

Oxford (1999) differentiates between harmful and helpful anxiety, although she admits that there is not a common view as to when and how foreign language anxiety can be helpful but there are many research findings about harmful anxiety.

According to Oxford (1999), 'debilitating' anxiety manifests itself in the negative relationship between anxiety and performance. *It harms learners' performance*

indirectly through worry and self-doubt and directly by reducing participation and creating overt avoidance of the language. It can be related to plummeting motivation, negative attitudes and beliefs, and language performance difficulties (Oxford 1999, p.60).

Madsen (1982) differentiates between facilitating and debilitating anxiety impact. He considers that students who are anxiety prone will not be evaluated as accurately on stressful tests as those who are not anxiety prone and therefore he attempts to develop a means of assessing when the anxiety level of a test crosses the debilitating threshold. He measured the level of anxiety of the test takers on scale from 1 (anxiety allaying) to 15 (highly anxiety producing) and proposed that the debilitating threshold is 9 out of 15. He found that on one of the tests, that of Reading, the students consistently scored above 9 and therefore he decided that this test should be excluded from the test battery in order to avoid results that are biased in favour of students who are not anxiety prone (Madsen 1982, 141).

Madsen's proposal that there is an 'anxiety threshold' (9 out of 15 or 60%) above which the test biases against students who are anxiety prone would be useful if his proposal had been accompanied by some arguments as to why the threshold should be exactly 60% and not some other level of anxiety. However, he explains why Reading test should be excluded from test battery:

1. for the Reading test it was the highest as well as the lowest level of proficiency test-takers who marked it as the most anxiety-provoking test

2. the reliability level of the Reading test was lowest for the high anxiety group (.65), while for the low anxiety group it was .89.

On the other hand, Madsen does not comment on the fact that the Listening test reliability for the low anxiety level group was only .57, while for the higher anxiety group it was .70.

4.3.1.4 Signs of foreign language anxiety

Oxford (1999) says that although researchers usually use questionnaires to measure the level of foreign anxiety (for example the Foreign language classroom anxiety scale (Horwitz et al 1986)), filling in the questionnaires is not the only way of determining whether foreign language anxiety is present during a class. Oxford says that foreign language anxiety is readily observable and mentions the following signs of foreign language anxiety:

1. General avoidance (carelessness, cutting classes, arriving unprepared, low levels of verbal production, lack of volunteering in class, inability to answer even the simplest questions)

2. Physical actions (squirming, fidgeting, nervously touching objects, stuttering and stammering, jittery behaviour)

3. Physical symptoms (headache, experiencing tight muscles, unexplained pain or tension in any part of the body)

4. Other signs depending on the culture (over-studying, social avoidance, conversational withdrawal, lack of eye contact, expressive competitiveness, self-criticism, face-saving: laughing and joking).

Oxford does not, however, specify whether these signs can be observed in a classroom or test situation.

4.3.1.5 Causes and effects of foreign language anxiety

Price (1991) suggests that language learning contexts appear to be particularly prone to anxiety arousal. Horwitz, Horwitz and Cope (1986) found that for many students foreign language courses were the most anxiety provoking courses of all the courses they were taking.

Gardner and McIntyre (1991) state that anxiety is the strongest negative correlate with language achievement. Oxford (1999) lists the following studies that have found negative correlations between anxiety and:

1. grades in language courses (Aida 1994, Horwitz et al 1986, Trylong 1987)

2. proficiency test performance (Ganschow, Sparks 1994)

3. performance in speaking and writing tasks (Trylong 1987, Young 1986)

4. self-confidence in language learning (MacIntyre and Gardner 1991, 1993)

5. self-esteem (Horwitz and Cope 1986, Price 1991, Scarcella and Oxford 1992).

Oxford also notes that it is possible that the language/anxiety relationship may be different for the different language skills, with negative correlations between anxiety in one skill, but not another. Ganschow et al (1994) suggest that high anxiety can be the *result* of language learning problems rather than the cause. MacIntyre and Gardner (1991), however, consider that a unidirectional model may be far too simple. They consider that anxiety creates poor performance, which leads to more anxiety and even poorer performance.

According to Gardner and McIntyre (1991) foreign language anxiety is associated with problems in second language learning and they say that it stems primarily from the social and communicative aspects of language learning. It can be considered as social anxiety, and consists of cognitive, affective and behavioural components:

- the affective consequences include feelings of apprehension, uneasiness and fear (MacIntyre, Gardner 1991),

- the cognitive effects are increasing self-related cognition, expectation of failure and decrease in cognitive processing ability (Wine 1980),

The behavioural effects increase sympathetic nervous system arousal, inhibited actions and attempts to escape the situation.

According to Madsen et al (1991) the students' level of anxiety depends on several intervening variables:

1. the subject tested

2. the student's level of intelligence

3. the difficulty of the skill

4. the degree of familiarity with the task

5. perceptions of item difficulty

6. time limitations

7. ambiguity in item stems

8. low quality of recordings of listening tests

9. impact of exam form

10. the amount of time spent on instruction in EFL (Madsen, Brown, Jones 1991)

The reasons mentioned by Madsen et al are mostly concerned with the instrument of measurement or test facets. Only one of them stems from test-taker characteristics (level of intelligence). One would expect that language proficiency level, attitude to language learning, motivation, language anxiety and general state and trait anxiety level also have an influence on test-anxiety.

However, the crucial question for a langugae tester is whether foreign language anxiety undermines test validity which is the case if we are measuring test-taker's ability to control foreign language test anxiety next to language use ability.

I think that if we agree with Bachman and Palmer's (1996) language use model and consider affect (anxiety and motivation as well as other variables) as part of language use, then the popular complaint that the test anxiety produced by test situation interferes with test validity is ungrounded. If we treat affect, just like strategic competence as a part of language use framework, test-taker's ability to control his or

her emotion could be legitimately considered as a part of communicative competence that can be taught as well as tested.

4.3.2 Language Skill approach

Saito, Garza and Horwitz (1999) and Cheng, Horwitz and Schallert (1999) suggest that what we should be talking about is not general foreign language anxiety, but foreign language reading, writing, speaking and listening anxiety. Recent research on second language anxiety appears to support the existence of language-skill specific anxiety (Cheng et al. 1999, p.439). They say that this approach to language anxiety as a skill-specific anxiety would explain the contradictory results of many studies, where anxiety leaves a positive impact on one skill, but a negative one on another.

4.3.2.1 Foreign language speaking anxiety

The concern with foreign language anxiety started off by mainly focusing on oral skills (66% of the most widely used instruments for measuring foreign language classroom anxiety deal with speaking, Cheng et al 1999). Aida (1994) says that her validation of the study of Horwitz, Horwitz and Cope (1986) Foreign language classroom anxiety survey (FLCAS) suggests that there is only one meaningful factor, that of speech anxiety and fear of evaluation.

Cheng et al (1999) also say that the results of their analyses of the FLCAS show high correlations between classroom anxiety and speaking performance. They consider that what is often considered to be classroom anxiety is in fact speaking anxiety.

Hilleson (1996) used diaries, interviews, observations and questionnaires to research the causes of debilitating anxiety during foreign language study. He grouped his findings about foreign language anxiety according to the language skills. He found that his students mentioned the following causes of speaking anxiety:

1. type of personality: some people just do not like talking in a foreign language

2. tiredness: accent becomes more pronounced as a person gets tired and as the person registers it, this becomes an additional cause of anxiety

3. many people participating in a discussion impose a speed of interaction which is difficult to comply with

4. an attempt to speak correctly interferes with speech content.

Hilleson also found that speaking anxiety was not a constant phenomenon, as the students were anxious on one day and satisfied with their performance on another day. Hilleson suggested that role-play liberates the students from their anxieties.

4.3.2.2 Foreign language listening anxiety

Listening, according to Vogley (1998), is the most frequently used language skill in the foreign language classroom and contributes more to language learning success than any other skill. Therefore listening anxiety is especially harmful. Vogley considers that teachers often prepare special tasks for training the speaking skill, but they forget that speaking cannot be trained without listening: before we decide what we are going to say, we need to understand what has been said.

Vogley (1998) considers that listening anxiety springs from the learner's false impression that they need to understand every word that is said. In her research (using open-ended questions addressed to 140 students) she found that the learners mentioned the following reasons for listening anxiety :

1. 51% of the respondents blamed input for their listening anxiety: nature of speech (28%), level of difficulty (11%), lack of clarity (5%), lack of visual support (4%), repetition of input (3%)

2. 30% of the respondents were concerned with the process itself: inappropriate strategies (trying to translate word for word) (24%), lack of time to process (3%), cannot study for listening tests (2%) and cannot check answers (1%)

3. 6% of the students blamed instructional factors: lack of listening practice (3%), 'the test thing' (2%), uncomfortable environment (too small a group, feeling hot or cold) (1%)

4. 13% of respondents blamed personal factors: fear of failure (10%), nerves (2%), the instructor's personality (1%).

Vogley suggests that to remedy listening problems, it is not enough to practise and hope that listening will improve with time. If teachers want to improve their students' performance in listening, they should also actively teach cognitive and meta-cognitive strategies.

4.3.2.3 Foreign language reading anxiety

Saito et al (1999) propose that because reading is not a social skill as it is done privately and has an unlimited opportunity for reflection and reconsideration, it does not depend on a dynamic construction of meaning by two or more speakers. At the same time there are reading specific factors that can make reading skill more anxiety provoking than any other language skill: unfamiliar scripts and writing systems and unfamiliar cultural material used to illustrate the text. Saito et al hypothesize that anxiety about the text will arise at the point when the words the student has decoded do not constitute a comprehensible message. This anxiety should be separated from the other language skill anxieties. The results of an analysis of nearly four hundred students' responses to questionnaires suggest that:

1. reading anxiety is a separate construct (as all the questions aimed at reading anxiety had good internal reliability) which has a significant correlation with overall foreign language classroom anxiety (.64)

2. reading anxiety had a significant negative effect on language performance

3. reading anxiety level depends on the type of target language: in their study American learners of Japanese were the most anxious (mean of 56%), followed by learners of French (53%) and the learners of Russian (47%)

4. the more difficult the readers found the reading process, the more anxious they were.

Although Saito et al say that it is difficult to determine the cause effect relationship in reading anxiety, they suggest that in their study reading anxiety was experienced as a result of actual reading difficulties in text processing rather than reading difficulties stemming from the student's level of anxiety. This to my mind suggests that during the reading process the learners are constantly evaluating the quality of their performance and that anxiety is caused by a negative evaluation.

4.3.2.4 Foreign language writing anxiety

Cheng et al (1999) propose that the results of research over the period since the 1970 suggest that native language writing apprehension has a negative impact on

1. the quality of the message encoded

2. the individual's writing behaviour

3. writing performance

4. willingness to write or take writing courses.

The authors constructed a special questionnaire to address writing anxiety and compared its results with the FLCAS (Horwitz 1986) results. Their findings suggest that:

1. the correlation between the FLCAS and writing anxiety is significant (.65) but is not sufficient to suggest that the constructs are identical

2. factor analyses suggest that writing anxiety consists of three separate factors: low confidence in writing English, level of enjoyment in writing in English and fear of evaluation

3. the comparison between the speaking and writing anxiety constructs as interpreted by the questionnaire suggest that the writing anxiety variables were more highly associated with writing achievement than with speaking achievement

4. writing anxiety scores had a significant predictive ability in relation to writing skill, but not in relation to other skills

5. the correlation between the learners perceived achievement and their level of anxiety was higher than the correlations between their actual achievement and their level of anxiety.

Cheng et al consider that their findings suggest the existence of separate language skill anxieties. Point 5 to my mind suggests once again that there is an intimate relationship between the assessment strategies and level of anxiety, rather than the actual performance level.

4.3.3 Cognitive approach

Onwuegbuzie et al (2000) see foreign language anxiety as *among the most important affective predictors of foreign language achievement* (Onwuegbuzie et al 2000, p. 88). They aim at validating MacIntyre and Gardner's (1994) questionnaire that investigates anxiety as a process consisting of three stages: input, processing and output anxieties that sometimes overlap, but can still be differentiated. This MacIntyre and Gardner's framework reminds me of Wenden's (1998) theory of meta-cognition as a process: *meta-cognitive strategies are general skills through which learners manage, direct, regulate, guide their learning i.e. planning, monitoring and evaluating* (Wenden 1998). The similar stages of meta-cognition (Wenden 1998) and anxiety (MacIntyre and Gardner 1994) suggest that both the processes are parallel, or as Liddell (1950) says anxiety follows intellect like a shadow.

Having investigated the reliability and validity of MacIntyre and Gardner's questionnaire Onwuegbuzie et al (2000) found that:

1. there were significant correlations between the input, processing and output anxiety scales (MacIntyre and Gardner 1994) and FLCAS (Horwitz 1986) suggesting that they are all measuring foreign language anxiety

2. the factor analysis of the input, processing and output anxiety questionnaire suggested that there was an interdependence between the three processing stage anxieties and Onwuegbuzie et al rejected a three factor model. However, they consider that these could be caused by the size of the sample, as large sample (more than 200) analyses tend to lead to a rejection of the underlying models.

I will now examine each of the three processing stage anxieties separately.

4.3.3.1 Foreign language input anxiety

Following Gardner, Onwuegbuzie et al define input anxiety as the fear experienced by the foreign language students when they are initially presented with a new word, phrase or sentence in the foreign language; and it is connected with the student's ability to receive, concentrate on and encode external stimuli. The questionnaire questions that loaded highest in the factor analyses relating to the input anxiety were:

1. I get flustered unless French is spoken very slowly and deliberately (.77)
2. I get upset when French is spoken quickly (.78)
3. I get upset when I read in French because I must read things again and again (.57).

The analyses suggest that input anxiety is mostly closely related to the global foreign language anxiety. If the student's ability to attend to material diminishes, in-put

anxiety can appear: the student starts attending to task irrelevant information thus reducing further the capacity to absorb the input. Students with input anxiety tend to ask for repetitions or they reread the same text several times to compensate for the inadequate input.

4.3.3.2 Foreign language processing anxiety

According to Onwuegbuzie et al (2000) processing anxiety appears when the students are attempting to organize and store input. The amount of anxiety at this stage depends on the difficulty of the material that is being stored, on the level of organisation of the material and the extent to which memory is relied upon.

Questions from MacIntyre and Gardner's questionnaire that represented the language processing anxiety factor were following:

1. I do not worry when I hear new or unfamiliar words, I am confident that I can understand them (.72)

2. Learning new French vocabulary does not worry me, I ca acquire it in no time (.68)

3. I am anxious with French because I have trouble understanding it (.66).

Anxiety at this stage of processing can reduce the student's ability to understand messages or to learn new vocabulary (Onwuegbuzie et al 2000, p. 90).

4.3.3.3 Foreign language output anxiety

Anxiety at the output stage can appear when students are asked to demonstrate their ability to produce previously learned material, that is after the material has been processed and before the production process has been finished. This is how anxiety hinders the retrieval of previously learned material and this hinders student's ability to speak or write in the foreign language. The questions that loaded highest were:

1. I may know the proper French expression, but when I am nervous it just won't come out (.69)

2. When I become nervous during a French test, I cannot remember anything I studied (.63).

Output anxiety levels were higher than input and processing level anxiety and explain more than 40% variance of overall foreign language anxiety. This suggests that there is a closeness between the foreign language anxiety and evaluation anxiety, because it is the output that is evaluated.

The cognitive approach to foreign language anxiety suggest that not only the level of anxiety may change during the processing of language, but also that the quality of anxiety may differ at different language processing stages.

If we bring the findings of the cognitive approach to foreign language anxiety together with the findings of the language skill approach (discussed in section 4.1.7.2.) and situational approach (4.1.7.1) we find a plethora of different anxieties (test, classroom, reading, listening, speaking, writing, input, processing and output anxiety) that appear as separate factors in factor analyses. To my mind this supports Lidell's (1949) suggestion that anxiety follows intelligence as a shadow; and as soon as we have a separate skill, or a situation that demands a special skill, we can also have a separate factor of anxiety.

I will now explore methods that have been used to investigate the interaction between affect and cognition in psychology and linguistics.

Chapter 5   Methods of research of the interaction between meta-cognition and affect

Cognition and affect are popular topics of research, therefore the research methods used range from the statistical analyses of large data sets (Purpura 1999) to impression and theory driven observations and interviews (Oxford 1990).

Researchers have also different ways of starting their research: Purpura (1999) starts with a theory and then elaborates it (the theory, then research approach (Long 1985), Oxford (1990) starts with a thorough review of previous research and ends up with a new theory. Thus their end products are as different as their data: Oxford investigates the process (of language acquisition), but Purpura examines the product (effect of meta-cognitive strategy use on language performance).

Between the two extremes there are other researchers whose methods are also described in this section.

Price (1991) proposes that foreign language anxiety researchers make use of both quantitative and qualitative research methods:

- quantitative research methods:

    1. re-examining the anxiety-proficiency relationship Madsen (1982)

    2. focusing on the relationship between anxiety and learner variables (Horwitz 1986)

    3. examining the effects of anxiety on the language learner or test-taker (Shohamy 1982)

- qualitative methods have been used by fewer researchers:

    1. to identify sources of anxiety (Madsen 1982, Brown 1993)

    2. to develop a theory of foreign language anxiety (Horwitz, Horwitz and Cope 1986).

5.1 Qualitative research methods

Methods of evaluating impact of learner variables have recently moved from observation methods to self-report and from quantitative to qualitative research methods. Purpura (1999) considers that in the 1980s the most popular methods of investigation of cognitive processes and performances were

3. observation (Rubin 1981)

4. structured interviews (Wenden 1998)

5. questionnaires (Bialystok 1978)

6. verbal report protocols (Cohen 1984).

5.1.1 Observational research

Observation of the English language examination administration procedure is used to explore the existing examination system, when a change of the system is envisaged (see Alderson and Szollàs, 2000). Dyer (1997) says that observation is used to examine a socially meaningful situation and it requires a researcher to enter a situation where some behaviour of interest is likely to take place, to watch the nature and frequency with which particular forms of behaviour occur and make a record of what is observed (so called 'field notes'). The researcher has a choice between

1. overt (when everybody knows that they are observed) and covert observation (so that those observed are unaware of being observed)

2. participant observation (for example, when the researcher takes the examination together with all the other test-takers) and non-participant observation, when a researcher simply observes the procedure

3. structured (using a checklist based on a particular theory) or unstructured observation (noting down everything relevant to one's research)

4. continuous observation (field is observed non-stop for a given period of time and all instances recorded) or time-point sampling (to provide a snapshot of behaviour at successive time periods).

5.1.2 Interviews

Madsen et al (1991) remarked that *the source of frustration is of the utmost importance and this also needs to be systematically investigated. ... There is no need to restrict studies to classical empirical research. For example, personal interviews following the exam may provide excellent insights not otherwise available* (Madsen 1991 p.142).

Brown (1993) and Madsen et al (1991) found that questionnaires and interviews led to genuine communication with test-takers and enabled researchers to get to know not only the product (test-takers' performance), but also to evaluate the process and locate the factors within the test, the test-taker characteristics or the test-environment that had influenced the students' performance.

In addition, the qualitative research methods, by involving the test-takers in test method evaluation introduce aspects that are important to test-takers themselves and as a result lead to a better understanding of the processes the test-takers are going through. Smith et al (1995) considers that the qualitative analysis is not driven by prior theory, but instead seeks to adopt a bottom-up approach, in which the data themselves suggest theoretical insights. Although quantitative research methods can also provide unexpected findings, there the researcher is alone with the data, while in qualitative research, the interviewees provide their own interpretations of researchers findings that add to the researcher's interpretation. I think that the recent change of vocabulary that can be seen in psychology [from 'subjects of the experiment' (May

1979) to 'participants of the experiment' (LeDoux 1999)] also suggest the possibility of the change of role of test-taker in language testing research.

5.1.3 Verbalization

Brown (1987) differentiates between the verbalization of stable knowledge and reports on states that occur during problem solving. Such reports can cover predictive verbalisations about possible performance before the event, concurrent verbalisations during the actual performance and retrospective verbalisations after the performance. Verbalisation report methods have been criticized by many authors:

1. predictive verbalisation: Ericsson and Simon (1980), think that the effect of verbalisation method on learning can be both positive and negative depending on the function of the verbal report

2. concurrent verbalisation: Piaget (1978) considers that it distorts the process

3. retrospective verbalisation: Sternberg et al (1982) (quoted in Brown 1987) found that their subjects were consistently describing the strategies they had been trained to use instead of the strategies they had actually used.

Flannagan (1954) however, considered that verbalisation could be reliably used if one followed the so-called critical incident technique, and asked informants only about very specific incidents and did not put any general questions concerning the evaluation of their performance.

Damasio (2000) also considers that verbal reports have their role in the investigation of consciousness and proposes that we must use both external and internal reports. He considers that we can treat subjective phenomena scientifically, as the human mind itself is a subjective phenomenon and it is the task of science *to verify objectively the consistency of many individual subjectivities* (Damasio 2000, p.83). It can be done through establishing a three-way link between:

1. certain external manifestations (for example, wakefulness, background emotions, attention, specific behaviours)

2. the corresponding internal manifestations of reported human behaviour

3. the internal manifestations that we as observers can verify in ourselves when we are in equivalent circumstances.

The overt use of our own mental states and their evaluation (point 3 above) has not often been accepted as a method of research, but Damasio argues that theorizing constantly about the state of mind of others from observations of behaviours, reports of mental states and counterchecking their agreement is a natural part of human activity. If we admit the fact that this is the way we perceive other people and their states of mind, it makes us more responsible in our choice of method of research and also more watchful of our own judgement and the conclusions we make. The admission of the subjectivity of our methods and judgements is, to my mind, a necessary precondition for objectivity in our research.

McLaughlin (1990) says that the most frequently used methods for studying meta-cognitive strategies are qualitative: strategy lists, observations, interviews, diaries, note-taking and structured self-reports. He especially singles out Oxford's (1990) research into second language learning strategies for the thoroughness of the research. He says that Oxford based her research on a review of the literature on learning strategies and from this produced a list of second language learning strategies. Her provisional list of strategies was submitted to extensive field-testing and revision. She interviewed teachers, students and consulted fellow researchers for contributions to the production of the final product. She ended up with one of the most extensive lists of strategies.

Purpura (1999) considers that Oxford's (1990) methodology for the analyses of the data in the development of the strategy list was more thorough than that of others, as

she used combined qualitative and quantitative analyses; she carried out exploratory factor analyses to identify eight factor clusters of the 64 strategies she had included in her list of Strategies In Language Learning (SILL).

5.2 Quantitative methods

I believe that a survey of the literature and an investigation using qualitative research methods can provide new insights into the nature of affect. However, the disadvantage of a qualitative research is that it does not enable the researcher to make any generalisations about the interaction between different affective variables and language performance.

Chomsky (1980) says that we can profitably study motivation, contingencies that guide action, drives and many similar topics, but the freedom of choice remains inexplicable. He compares the complexity of the human mind to the complexity of universe and suggests following the example of physicists who approach their inquiry in 'Galilean style', making abstract mathematical models. Chomsky says: *To what extent and in what ways can inquiry in something like 'Galilean style' yield insight and understanding of the roots of human nature in the cognitive domain? Can we move beyond the superficiality by a readiness to undertake perhaps far-reaching idealization and to construct abstract models, (...), by a readiness to tolerate unexplained phenomena, much as Galileo did not abandon his enterprise because he was unable to give a coherent explanation for the fact that objects do not fly off Earth's surface?* (Chomsky 1980, p.9).

5.2.1 Structural Equation Modelling

Research in language testing has used modelling to investigate the interaction between questionnaire results and language performance test results, in order to develop a

model of the variables and depict the connection between them. Kunnan (1995) investigated the interconnection between test-taker characteristics and language performance and Purpura (1999) analysed the interaction between learner strategies and language performance using Structural Equation Modelling (SEM). It seems to me that the interaction between affect and test-performance should be investigated in a similar way.

Purpura (1999) says that SEM is used to investigate the relationships between background variables and criterion variables, between observed and latent variables based on a substantive theory and empirical research.

According to Purpura (1999), Purcell (1983) was the first to use SEM to investigate second language performance. He based his research on 11 observed variables measuring pronunciation accuracy and expected that pronunciation would form a single latent variable. However, he found that factor analysis with a two factor orthogonal model fitted the data better.

Gardner (1985) used the SEM approach to investigate the role of motivation in second language performance and developed his socio-educational model that showed that language aptitude and motivation, consisting of attitudinal variables, had a direct impact on second language performance.

I will discuss the research of Kunan (1995) and Purpura (1999) in detail because they both investigated the influence of different variables on language performance and because their models fulfil the demands of Caccioppo et al (1999), which are:

What are needed are psychological models of the affect system that do not merely speculate about meditating psychological processes but that instead specify them in detailed, empirically meaningful ways (Cacciopo et al 1999, p. 850).

5.2.1.1 Kunan (1995)

Kunnan (1995) investigated the relationships between test-taker characteristics and test performance in different groups (non-Indo-European versus Indo-European) classified according to their method of training (formal versus informal). He used Structural Equation Modelling (SEM) to develop a language use model that would take into account language ability and the language learning variables that affect it.

SEM, according to Kunan, involves the formulation of models by positing relationships among constructs, followed by an evaluation of these models and, if the models do not adequately explain relationships, the researcher proposes alternative substantive models.

Kunnan discusses the difference between a strong program of construct validation, which involves formal hypothesis testing and a weak program which involves widespread support for explanations from many perspectives that would focus not on confirming or rejecting a set of hypotheses but would instead seek explanations for the phenomena that are being investigated.

Kunnan developed two different models to depict how the learner characteristics influence language performance:

1. The equal influence model, that posited equal, direct influences of all test-taker characteristic factors (home country formal, home country informal, English speaking country instructions and monitoring)

2. Gardner's intervening factors model that posited that exposure variables (home country formal, home country informal, and English speaking country instruction) all influence monitoring, which in its turn influences test performance factors.

Kunnan found that Model 2 brought out more interesting factors than Model 1, which just showed that monitoring had a strong positive impact on one test component (Reading Writing 1) and a negative impact on another (Reading Writing 2).

Model 2 proposed that students who had had formal instruction strive for correctness and monitor themselves more than those who had had more informal training. Also English-speaking country instruction seemed to have inhibited monitoring, while formal home country instruction showed a strong positive relationship with monitoring.

5.2.1.2 Purpura (1999)

Purpura (1999) used Structural Equation Modelling to analyse the interaction between learner strategies (cognitive and meta-cognitive) and language performance. He based his research into meta-cognitive strategies on Bachman and Palmer's model of language use and examined the impact of meta-cognitive strategies on second language test performance.

Purpura followed several procedures in the development of his model:

1. data preparation (scoring and data inputting)

2. descriptive statistics (examining central tendencies, checking for normality)

3. reliability analyses (examining the homogeneity of scales)

4. exploratory factor analyses (examining factor clusters, forming composite variables)

5. single group Structural Equation Modelling (examining the measurement models, examining the structural models)

6. multi-group SEM (performing separate analyses for each group, simultaneous analyses for both groups, testing for cross group variance) (Purpura 1999, p.59).

Purpura (1999) also describes how he analysed his questionnaire reliability. The results were scanned in from the questionnaire forms, and the means were calculated. They ranged from 2.89 to 4.34 (minimum 0, meaning *never*; maximum 5, meaning *always*), which was a much smaller range than that of the cognitive strategies (1.68 to 3.69). The range of the standard deviations was also more varied for meta-cognitive strategies (0.97 to 1.64) than for cognitive strategies (1.12 to 1.63). The values for skewness and kurtosis were within acceptable limits and indicated that the items were reasonably well distributed, but in spite of this, Purpura decided to use the robust instead of the maximum likelihood estimation method (see section 10.3). The strategy type reliabilities ranged from 0.4 for the learning to learn scale to 0.82 for the evaluating scale. The lowest reliability figures were for the items examining learning to learn, goal-setting and planning strategies

Exploratory factor analysis yielded a four-factor oblimin solution. Items that did not cluster together in the exploratory factor analysis were removed, and as a result the whole group of questions investigating goal-setting strategies was deleted. Purpura hypothesized that this could be explained by the fact that when a person sets goals, he or she is also making a series of assessments and plans. The assessment strategies, however, differed according to whether they were carried out before the event, which could be called assessing the situation, during the event (monitoring the performance), or after the language learning or use event, when they were used for self testing. This supported Wenden's (1998) findings that the strategies can be classified according to when they were used. When Purpura regrouped the meta-cognitive strategies according to when they were applied he got the following matrix (see Table 6).

Table 6 Analysis of Purpura's (1999) questionnaire

	Strategies	Nr of items	Reliability
Online assessment processes	assessing situation	8	.72
	monitoring	4	.60
Post-assessment processes	self evaluating	10	.73
	self testing	8	.79

This suggests that his questionnaire reliably investigates the use of meta-cognitive strategies and also agrees with the existing theories.

I will return to Purpura's research methods in section 10.3.

5.2.2 The use of correlation coefficients

Scherer (2000) considers that our present conceptual and methodological tool kits are not adapted to deal with systems that are as complex as emotion processes. He says *we need a complete revolution in our thinking about the nature of emotion, comparable to other shifts in the history of science. In particular we need to move from thinking in terms of discrete boxes, labels or even neural programs to a nonlinear dynamic systems perspective of emotion* (Scherer 2000, p.80). Scherer invites us to turn to the tools that are being elaborated in the domain of chaos theory and catastrophe theory to depict the synchronization of the subsystems that constitute an emotion episode. This would finally resolve the difficulty of those who try to uncover the interactions between the different nervous systems using simple linear Pearson correlations. Linear correlations cannot be used as the nervous systems do not respond in a uniform manner to stimulation.

However, if we look at the graph of interdependence of performance and motivation the Yerke-Dodson law (for example in Fransson 1984, p.88), we see a curve; this has not stopped the researchers using linear Pearson inter-correlations when examining the interaction between meta-cognitive or affective variables and language

performance level. The results, not surprisingly, show both positive and negative relationships. Unfortunately regression method used in equation modeling also uses linear correlation methods, and again the interaction between meta-cognitive variables and language performance is sometimes positive and sometimes negative.

The analyses of the interaction by separate language skills (Cheng et al 1999) does not seem to solve the problem either. Evidently we have to wait until the new instruments (Scherer 2000) are developed before we can understand the true nature of interaction between language use and cognitive and affective variables. Meanwhile, I intend to use both quantitative and qualitative methods to answer my questions, which will be the focus of the next chapter.

Chapter 6   Research questions

My study started off with a measurement of the level of anxiety and its impact on language performance. This led to an exploration of the role of anxiety, its causes and effects on language performance. Later, when the theoretical findings suggested the interaction between cognition and affect as a precondition of understanding of affective variables, the concept of meta-cognition was explored and used to research the basis of foreign language anxiety. Thus I could say that the theoretical findings led to research questions, which were addressed by the practical research. The results of the analyses of the data provided suggested new research questions that were again explored in theory and practice. This explains the number of research questions. I have grouped them in three separate sections (exploring meta-cognition, anxiety and the interaction between the two).

6.1   *What is the role of meta-cognitive strategies in language use?*

This group of research questions is going to focus on meta-cognition, the kinds of strategies test-takers use, the frequency of their use. It will also examine the interaction between the areas of meta-cognition and the interaction between language proficiency and meta-cognition.

6.1.1 What areas of meta-cognitive competence are used by test-takers?

Bachman and Palmer (1996) suggest that meta-cognition should be seen as a set of meta-cognitive strategies that provide a cognitive management of language use, as well as in other cognitive activities. They identify three areas of meta-cognitive strategy use: goal setting, assessment and planning.

Purpura (1999), however, suggests that meta-cognition, as measured by his Meta-cognitive strategy questionnaire is a uni-dimensional construct consisting of a single set of assessment processes.

I used Purpura's (1999) meta-cognitive strategy questionnaire and examined the frequency of use of all the three meta-cognitive strategy areas to find out whether all three area strategies were utilised by the test-takers. Then I carried out factor analyses and SEM to examine the question whether meta-cognition is a uni-dimensional construct as suggested by Purpura (1999) or a multi-dimensional as suggested by Bachman and Palmer (1996).

6.1.2 What is the interaction between different areas of meta-cognitive areas?

If meta-cognitive competence were a multi-dimensional construct as suggested by Bachman and Palmer (1996) then the next questions would be of the interaction between the different areas of meta-cognitive competence. Is the interaction bi-directional as predicted by Bachman (1990) or is one of the areas, namely assessment area, dominating goal setting and planning (as found by Purpura 1999)?

6.1.3 How often are meta-cognitive strategies used in different groups of population?

One of the ways of examining the role of meta-cognitive competence on language use is comparing how meta-cognitive areas are used in different groups of proficiency (see Purpura 1999). As my focus is foreign language anxiety, I also examined the use of meta-cognitive strategies is groups of different levels of anxiety.

6.1.4 What is the interaction between meta-cognition and language proficiency?

The interaction between meta-cognition and language use has been examined by Kunnan (1995) and Purpura (1999), their findings, however, suggest that meta-cognitive competence can have diverse effects on language performance. For example, Purpura (1999) says that his study suggests that not only meta-cognitive, but also cognitive strategies were used by high ability level group less often and comes to the conclusion that to achieve good results, it is not enough to use many strategies, one has to use them effectively.

I used correlations as well as SEM to examine the interaction between the different areas of meta-cognitive strategy use and language proficiency (see sections 10.2 and 10.3).

6.2 What is the role of anxiety in language use?

The question of the role of anxiety may seem unexpected, as usually anxiety is equalled with just one role, that of interference. However, already Madsen (1982) suggested that we should be talking of both debilitating and facilitating anxiety. If we take this into account, we have admitted that anxiety can have different roles. This section of research questions is going to examine the concept of anxiety, its level, signs, its causes and effects before examining the interaction of anxiety and meta-cognition.

6.2.1 What is the level of anxiety during the Year 12 English language examination*?*

Measurement of the level of anxiety is a popular topic and many scales for measuring test anxiety have been developed (see Chapter 4), but there is no common point of reference that we can refer to. Madsen's (1982) research is an exception: he tried to

establish the level of anxiety at which a test crosses the debilitating threshold (9 points out of 15 in the Alpert-Haber Achievement Anxiety scale). From this we can conclude that the level of test anxiety differs according to the test. This anxiety can be measured and must be measured if we do not want to produce tests that are biased against some part of the test-taking population.

So my first research question is: how much test anxiety is caused by the Year 12 examination in Latvia? Is the anxiety level caused by the Year 12 examination similar to other tests? To answer this question, anxiety levels for each task and each skill were measured using a questionnaire that grades the level of anxiety from 1 to 4 (see Appendices 3 and 5). The analyses of the results can be found in Chapter 10.

6.2.2 What signs of anxiety can be observed during a written and oral test?

Oxford (1999) mentions four groups of signs of anxiety that can be observed during the language acquisition phase (general avoidance, physical signs, physical symptoms and other cultural signs). Which of these can be observed during the test situation, written and oral part? Are test-takers aware of them, do they affect the test-taker's performance? These questions will be addressed with the help of observation and interview methods (see Chapters 8 and 9).

Although the amount of literature available on the impact of test anxiety in educational testing is huge (see section 4.2) and although there is also a considerable amount of literature available on the impact of test anxiety in language testing, it is rare to find literature on the causes of anxiety. The complexity of this question is caused by the fact that it is necessary to examine not only the test construction principles but also the construct of anxiety and the interaction of the testing and psychological or affective environment of the test-takers.

6.2.3 What types of anxiety can be distinguished?

I intend to investigate the existence of state and trait anxiety, test and classroom anxiety as well as different foreign language skill anxieties.

6.2.3.1 What is the role of each language skill in foreign language test state anxiety?

Is foreign language test anxiety a unitary concept, or does it consist of four separate language skill anxieties as Saito et al (1999) predicts? What is the interaction between the language skill anxieties like?

6.2.3.2 What is the relationship between foreign language test state and test trait anxiety?

Spielberger et al (1978) proposed that test anxiety is formed by state and trait anxieties, Gardner (1983) proposed that language anxiety can start as a state but if this state occurs repeatedly it can develop into trait anxiety. Is it possible that a person experiences both test and trait anxiety? What is their interaction like?

6.2.3.3 What is the relationship between foreign language test state and classroom anxiety?

Horwitz (1986) suggests that test anxiety is part of classroom anxiety as teachers often use tests in their classes. She reports high correlations (.5) between the two. My question is: is classroom anxiety also part of test state anxiety and what is the interaction between the two? Does classroom anxiety act as a cause of test anxiety? What is the impact of test anxiety on language performance?

6.2.4 What are the self-reported causes of test anxiety?

Bachman differentiates between four categories of influence on language test scores: communicative language ability, test method facets, personal characteristics and random measurement error. In addition to acting as individual influences on test performance, he suggests that components of communicative language ability, test method facets and personal attributes may interact with each other, constituting additional sources of variation (Bachman 1990 p.348). If it is the test score that creates anxiety among the test-takers all the four categories should appear as the causes of anxiety if the test-takers are asked what causes anxiety during a language test. To answer the question 'What causes anxiety during the Year 12 examination?' qualitative research method should be more appropriate than quantitative.

The same questionnaire will be used and students will be encouraged to give their reasons for anxiety in each task (see Appendix 5). The results of the questionnaire will be coded and the frequency for each cause of anxiety will be calculated. The results are discussed in Chapter 10.

6.2.4.1 General state anxiety caused by evaluative situation

Becker (1982) examined both the levels and the causes of test anxiety. He proposed that the psychological situation of the examinee was determined by:

1. the person's estimation of his or her level of competence, particularly in comparison to that of the other examinees

2. the student's level of aspiration

3. an estimation of the difficulty of the examination

4. the student's expectation of success or failure

5. an estimation of the grade he or she would attain

6. an evaluation of the importance to the student of reaching his or her goal. (Becker 1982, p.277).

This list differs from Madsen's (1991) lists of causes of anxiety (see section 4.3) as it takes into account the test-takers' affective variables. The word 'estimation' occurs 3 times in this list, next to words like 'expectation' and 'aspiration', which suggest activation of goals and assessment of the situation and one's own performance as the motives for anxiety. The question that remains unclear is what the test-takers themselves will suggest as reasons and how important this cause is.

6.2.4.2 Difficulty level of the task

Although Madsen (1995) does not research the reasons for test-anxiety, he indirectly suggests that it is the level of difficulty of the examination that causes the anxiety: *the anxiety generated by the Reading test appears to stem largely from the complexity and difficulty of the items. On the three-question state anxiety questionnaire, it was the difficulty of the item that registered the strongest negative reaction* (Madsen et al, 1991, p.140).

If it is the difficulty level that decides the level of anxiety, then the next question is what is the optimum level of difficulty for a task? By optimum I mean the level at which it is not boringly easy nor is it so difficult that it causes too much worry. The test-takers may acknowledge that the task is worrying, but may give a different reason why this is so.

6.2.4.3 Foreign language anxiety

Recent research carried out in language acquisition has singled out reading and listening, writing and speaking apprehension as separate phenomena (see section 4.3.2).

One question that arises within the testing environment is what happens to language anxiety in a test. Madsen (1982) having researched the results of a reading test, suggested not using this test for evaluation because it was more frustrating than the other parts of the examination, and therefore was biased against students who were anxiety prone. A question that remains unanswered is whether it is reading that causes most of the test anxiety or whether the test method is inappropriate.

6.2.4.4 Test-taker characteristics

Test-taker characteristics have been acknowledged as a source of variance in language testing (see Chapter 3):

1. Madsen (1982) in his research found that sex and language have significant effects.

2. Shohamy (1982) found that a student's attitude towards a cloze test was directly related to achievement on the test.

3. Bachman (1990) calls the interaction between test facets and test-taker characteristics (age, nationality, sex, attitudes) the source of error in measurement

4. Heinrich and Spielberger (1982) suggest that *for subjects with superior intelligence, high anxiety will facilitate performance on most learning tasks. For low intelligence subjects, high anxiety may facilitate performance on simple tasks that have been mastered. However, performance decrements will generally be*

associated with high anxiety on difficult tasks, especially in the early stages of learning (Heinrich and Spielberger (1982) p.147).

Although Heinrich and Spielberger do not provide any empirical data to support their predictions, and they are not talking about language proficiency, nevertheless, the idea is worth exploring here as the other test-taker characteristics have received more attention than the level of intelligence. I did not though, look at the impact of the level of intelligence on anxiety, but instead focused on the impact of the language proficiency and level of state and trait anxiety on the level of anxiety.

6.2.4.5 Problems with test design, administration and personal problems

According to Bachman (1990) language test scores are influenced by communicative language ability, test method facets, personal characteristics and random measurement error. Problems with test design and administration to my mind lead to measurement error. However, there is a difference between test-design and administration problems. The test-design problems influence all the test-takers equally and according to Bachman (1990) this leads to systematic error. Specific administration problems such as the invigilator standing at one's table during the test and interfering with one's concentration lead to unsystematic error and will influence only those who had that particular problem. The common feature of all these problems is that they do not have anything to do with language competence. As such, they deserve careful attention as they can suggest ways we can improve test-validity.

All the research findings to date suggest that the question of the causes of anxiety during the test will have a complex answer that will incorporate both the internal and external environment, the previous experience of the test-takers and their experience while taking the test.

6.2.5 What are the effects of anxiety?

The cause-effect relationship between test-anxiety and performance is problematic (see Horwitz 1986) and it is difficult to say whether a weak performance leads to higher anxiety or higher anxiety leads to weaker performance (see section 4.3.3.1). As one of the functions of affect is to self-assess one's performance and give both cognitive and affective feedback, affect can influence performance through feedback (see Stevick 1999 in section 3.3) by evoking many task irrelevant memories that interfere with performance. Gardner (1991) considers this a reciprocal cause-effect relationship when the test-taker is drawn into a vicious circle where the anxiety influences performance which causes more anxiety and still worse performance. I decided to use SEM to examine the effects of anxiety on language use (see section 10.3).

6.2.5.1 What is the effect of foreign language test state anxiety on foreign language performance?

Madsen (1982) and Brown (1980) report negative correlations between foreign language test state anxiety and performance. The question remains, which of the two acts as a cause and which effect, or is the relationship bi-directional as Gardner (1990) predicts?

6.2.5.2 What is the effect of foreign language test state anxiety on the use of meta-cognitive strategies?

Bachaman and Palmer's (1996) theoretical model of language use suggests interaction between meta-cognition and affect. Anxiety being part of affective schemata should also interact with meta-cognitive strategies. What is this interaction like? Are both the variables equally strong, or is one variable acting as a cause and the other as an effect?

6.3 What is the interaction between language proficiency, meta-cognition and anxiety?

Bachman and Palmer's (1996) model does not suggest direct interaction between affective schemata and language use characteristics (see section 3.4). Is this true? Do affective variables affect language use through meta-cognitive strategies, or is there a direct influence of anxiety on language performance? I used SEM as well as correlation coefficients to explore the interaction and produce models of language use (see section 10.3).

I will now proceed to the practical part of my research, which describes the research that was carried out in 1999 and 2000: the description of the instruments, the procedures and the results of the studies: observation and interview results that were analysed with the help of theoretical frameworks (Chapters 8 and 9); questionnaire data that were analysed with the help of statistical procedures (sections 10.1 and 10.2); models that were produced Structural Equation Modelling programme (section 10.3).

I will start, however, with the description of the English language examination (Chapter 7), as it was administered in Latvia: its contents, administration and marking procedures, statistical analyses and an attempt to validate the examination.

Chapter 7   Description of the Year 12 Examination

This chapter describes the environment of the study: the purpose of the examination, its contents, its marking procedures and statistical analyses of the results. The section also contains an attempt to validate the examination.

7.1   Purpose of the Examination

The purpose of a test can be found in the test specifications. Alderson et al (1995) differentiate between test specification and the syllabus. In Latvia, however, there is only one document that is meant for both test-users and test-developers. The document is revised and published once a year alongside all the subject test specifications.

The Year 12 examination in English for upper-secondary school-leavers is, according to the specifications, a test battery designed to assess the English language proficiency of upper-secondary school-leavers. According to Alderson et al (1995) *proficiency tests are not based on a particular language program. They are designed to test the ability of students with different language training backgrounds* (Alderson et al 1995, p.12).

The Year 12 examination, however, also reflects the secondary education course objectives, which are:

1. to develop the ability to use the foreign language effectively for the purpose of natural communication, learning grammar and lexis, practising the use of language elements to achieve a level that would allow survival and independence in an English language environment

2. to develop students' personal development, encouraging the acquisition of learning strategies and study skills, the development of different categories of thinking, imagination, creativity and cultural awareness.

This double nature of the examination, which is on the one hand, an achievement test, on the other hand a proficiency test, can be explained by the specific situation in Latvia: the students have studied English for different periods of time and according to different text-books. There is no common syllabus for all schools as the Education Law states that all schools and even all teachers have to develop their own syllabuses according to their learners' needs. At the same time they also have to follow the National Curriculum and achieve the stated objectives. This is why the Year 12 examination has to be a proficiency test.

The Year 12 examination result is used for decision-making at two levels: to determine the students' level of English language proficiency when leaving the secondary school and to determine the students' suitability to enter tertiary education. The test results are also used by the Ministry of Education and Science to analyse the foreign language teaching situation in different parts of the country. Teachers use the results to assess their own efficiency since they receive feedback about their students' performance in each skill.

However, it is also possible to take a university entrance examination before entering university if the test-taker considers that his or her score on the Year 12 examination is too low and does not represent his or her level of proficiency in English. A transition period started in 1998 when the first centralised examination in English was organised in Latvia and this will last until 2004, when all the universities stop organising their own entrance examinations.

The number of students who choose the centralised Year 12 exam in English has steadily increased from 300 in 1995 to 5184 in 2000 to 12000 in 2001 when the examination was compulsory for all students. The test-takers are 17 to 18 year old Latvian and Russian secondary school graduates.

7.2 Administration of the Examination

Test administration involves the delivery of a set of tasks to a group of test-takers under specified conditions (Davies et al, 1999 p. 4). Since 1997, rigorous measures have been taken to ensure adherence to instructions issued by the examination board.

The test-takers take the test either in their own or in a neighbouring school. If there are 10 or more test-takers in a school, they take the test in their own school. Specially appointed and trained invigilators who are usually teachers from neighbouring schools administer the test. The examination is taken at the same time in all schools all over the country: usually between 9.00-12.00 a.m. The test-takers are already familiar with the types of test-task and with the structure of the test, as the test specifications and training handbooks have previously been sent to all schools.

They are also informed about the criteria that will be used to evaluate their performance and the length of their response. All the test-takers take all the tasks.

However, it has become apparent that the test administration is responsible for several sources of unreliability:

1. students of small schools may have to spend several hours travelling to the school where the test is administered

2. the level of noise can be quite considerable as the test is administered during the school semester

3. the acoustics in big halls can vary according to where the candidates sit.

These issues have been discussed but nothing has as yet been done to counteract them.

7.3 Description of the Target Language Domains

The National Curriculum defines two target language domains:

1. visiting English language speaking countries

2. studying in national and foreign universities with English as the language of instruction.

The present official curriculum has only one level of attainment, which is close to the Threshold Level (Van Ek, 1990). However, as a result of the introduction of centralised marking the officials have had to accept the fact that the students graduate from secondary school with different levels of proficiency. Therefore it has became evident that the present curriculum is not suitable. According to the draft curriculum that is now being prepared, there are three main attainment targets:

1. learners of English as a third foreign language, who have studied English for three years are expected to have reached Waystage level (A2).

2. learners of English as a second foreign language (7 years of study) should reach Threshold Level (B1)

3. learners of English as a first foreign language (12 years) should reach Vantage level (C1).

All three groups take the same examination as this allows the universities to compare the test-takers' performance regardless of the number of years they have studied the language. With the exception of the Year 12 examination title page the language of the examination is English.

Bachman (1990) considers that *the ability to use language communicatively involves both knowledge of and competence in the language and the capacity for implementing or using this competence*. (Bachman 1990, p.81) This view has been theoretically accepted in Latvia since the 1980s, but there used to be many teachers who, until the centralised English language examination was introduced, did not adhere to it in practice. Now the English language examination tests both organisational and pragmatic competence (see Bachman 1990 p.87). Grammatical competence (morphology and syntax) is tested in all the tests, textual competence (cohesion and coherence) is tested mainly in the Writing test, illocutionary competence (ideational and manipulatory functions) in the Speaking test and socio-linguistic competence

(sensitivity to register, dialect and naturalness) in the Listening, Reading and Speaking tests (for sample materials see Appendices 1 and 2). Apart from testing language competence the examination also demands the ability to use knowledge of the world and an understanding of the context of the situation.

The choice of topics is also determined by the curriculum, but as the curriculum is vague, the test specification describes topics more explicitly. The Threshold level specification (Van Ek 1990) was used as the basis for the development of the 16 topic areas to be tested.

7.4 The Contents of the Examination

The Year 12 examination [see the examination materials in Appendix 1 (1999) and Appendix 2 (2000)] of English tests candidates' knowledge of English used for communicative purposes in an English-speaking environment. The students are expected to:

- Understand and appreciate authentic spoken and written English from a variety of sources in colloquial, informative and literary registers
- Seek and convey information, express ideas and feelings in conversation, personal and analytical writing.

The examination consists of 5 tests separately testing the four language skills: Reading, Listening, Writing and Speaking and also Language Use. Each test consists of 3 tasks of increasing difficulty proceeding from more familiar to less familiar situations. Equal weight is given to each section.

7.4.1 The Reading Test

The Reading test consists of three texts, of approximately 1000 words in total. Each text is followed by a set of questions checking reading comprehension. The test aims to assess the following target skills:

1. to extract key information on specific points by scanning the text

2. to understand the gist by skimming the text

3. to deduce the meaning of unfamiliar lexical items through the context

4. to understand the conceptual meaning of the notions of quantity, time, location and direction

5. to understand the communicative functions of sentences

6. to be able to appreciate emphatic as well as explicit facts and events

7. to identify the main idea.

The reading texts can be in the form of a narration, description, announcement, comment, anecdote, report or summary and come from brochures, catalogues, guides, directories, letters, postcards, diaries, public notices, signs or articles from newspapers, magazines or passages of fiction.

The reading comprehension tasks are different every year. They are selected from the list of the task types in the specifications and may consist of multiple choice or open-ended questions, matching, true/false, sentence completion, rearrangement, gap-filling and information transfer tasks.

The Reading test takes 40 minutes. This is considered to be an adequate time limit although some test-takers complain about the lack of time. For a sample paper see Appendix 2.

7.4.2 The Listening Test

The Listening test consists of 2 or 3 texts of varying length and nature. The texts are selected to test the test-takers' ability to understand authentic spoken English. The texts are delivered on a tape and repeated twice. The test-takers are given sufficient time to become acquainted with the task. They are allowed to start answering the questions while listening and during the pauses.

The test aims to assess the test-takers' ability to

1. identify the text-type, obtaining the gist

2. identify the main points from the supporting material

3. understand specific details

4. draw conclusions from and identify relationships between the ideas within the test

5. understand a variety of registers

6. understand the speakers' emotions and attitude towards the listener and topic of utterance conveyed by the speakers' intonation.

Each text has one or more native speakers speaking at a speed appropriate to the content and the text type. The text types may be announcements given through the public address systems, radio and TV news, advertisements, reports, reviews, requests and routine commands, telephone information, interviews, discussions or conversations. The recordings are authentic and may contain traces of regional accents (English or American), natural hesitations, spontaneous repetition or rephrasing of utterances and a limited amount of background noise. The vocabulary may be both formal and informal.

7.4.3 The Writing Test

The Writing test consists of three guided writing tasks requiring about 500 words in total. The test tests the test-takers' ability to

1. write correct and appropriate sentences

2. use conventions peculiar to written language

3. think creatively and develop thought excluding irrelevant information

4. manipulate sentences and paragraphs to use language effectively

5. write in an appropriate manner with a particular audience in mind,

6. organise and order the written material.

The test-takers demonstrate their ability to do three of the following: to write a letter, a postcard, a set of instructions, a report or a guided essay or fill in forms. In all cases test-takers are asked to respond to leaflets, notices, announcements, personal notes and messages, directions, tables and graphs.

7.4.4 The Speaking Test

In the Speaking test, the test-taker is expected to seek and impart information, express ideas, opinions and advice related to a variety of contexts and to discuss these with the interlocutor (their teacher). The test aims at assessing both routine (informational and interactional) and improvisation (negotiation of meaning, interaction management) skills (Weir 1995).

The test-takers are expected to demonstrate their ability to use these skills while imparting and seeking factual information, or expressing and discovering intellectual, emotional and moral attitudes, or getting things done and socialising (Van Ek, 1990).

The test contains three tasks. It starts with an introduction, where the interlocutor introduces the assessor, a teacher from a different school, and tries to set the student at ease. This section is not assessed. The first task for the test-taker is to find out the differences between two pictures, the test-taker's and the interlocutor's, these differences are not seen by the test-taker. Both the pictures depict the same situation, but in a different setting.

The second task is to answer questions on a specific topic, for example, 'environmental problems' and describe specific situations from their own experience. The first question is given to the student 1 minute before the student has to the question answer; the others are read out by the interlocutor from a card.

The third task is a role-play. The test-taker is expected to play a role that might be expected of 17 to 18 year olds in the real world. The final phase is a round-up when

the interlocutor creates a sense of accomplishment in the test-taker. The Speaking test lasts for 15 minutes (see Appendix 4).

7.4.5 The Language Use Test

The Language Use test was introduced after the first trials showed that neither the universities nor the specialised language schools were going to take an examination seriously if it did not have an explicit grammar component.

The Language Use test follows the guidelines presented in the Curriculum. It tests the students' ability to recognize appropriate grammatical forms and structures and also tests their ability to produce correct forms of language.

The test consists of three different tasks, which will test recognition or production ability. The tasks are of a growing level of difficulty

The test aims at assessing the following target skills:

- control of elements of the language in context

- ability to refine and proof-read samples of written English

- ability to recognise the structural features of the noun and verb phrase

- ability to provide the form of the word, phrase or sentence which conforms precisely to the grammatical constraints of the context

- editing skills.

Students have to recognize or produce correct forms and structures of language within the grammatical areas listed in the Curriculum (for example, the Noun, the Article, the Verb, the English tense system, Active/Passive voice).

The texts that are used for the development of the tasks are articles from newspapers and magazines, and forms of imaginative writing (fiction) that are considered to be within the experience of 17-18 year olds.

The tasks are chosen from the following selection:

1. multiple choice, rearrangement, broken sentences, completion items (gap filling), error-recognition and editing

2. a non-random cloze test (200 words) in which 15 words have been omitted (articles, prepositions, auxiliary verbs, cohesive devices, pronouns, determiners)

3. an editing task (200 words) containing a set number of grammatical errors of subject - verb agreement, tense, verb formation, omission or incorrect inclusion of articles and prepositions

4. a short text including information in tabular or diagrammatic form followed by an incomplete text providing the same information.

7.5 *Marking the Year 12 examination*

The contents of the test demand both objective (for the Reading and Listening Tests) and subjective marking (for the Speaking and Writing tests). The procedure demands approximately 300 examiners; it is expensive, time and effort consuming. The English language examination materials are marked simultaneously with the other subject examinations so the whole process has to be carefully planned and prepared.

7.5.1 Objective marking

The objectively marked Reading, Listening and Language Use tests are marked by specially trained groups of markers. Alderson et al (1995) consider *that there is a need to monitor the marking of objective tests, but this simply means checking that the examiners have applied the marking key or mark scheme properly and that their arithmetic is accurate* (Alderson et al 1995, p.128). Having done the checking for several years we have found that the mistakes in arithmetic and misuse of the marking key are so frequent that it is necessary to routinely double mark the objective as well as the subjective papers.

Another common problem with the objective papers is cheating; therefore we use tasks where test-takers have to write a word instead of choosing a multiple-choice letter.

The third problem is that it is often the case that the item statistics in the actual test differ from those of the pre-test. Therefore it is necessary to carry out item analyses before the routine marking. The first three hundred papers are marked immediately after the test and the results are fed into ITEMAN (see Appendix 9 (year 1999) and Appendix 10 (year 2000). The ITEMAN results are discussed among the markers and an answer key is developed and distributed to all the markers. The papers are then distributed to the markers for marking at home.

All the papers are double-marked and the results of each item are fed into the computer. All the results are processed by ITEMAN and SPSS.

7.5.2 Marking the Writing Test

Scoring of subjective papers has always been considered difficult because of the judgement that is made by the markers: *Any scoring procedure, which involves the exercise of judgement by the scorer, is called subjective scoring. In tests where the results hold serious consequences for the test-taker, multiple ratings by trained raters contribute to improved reliability* (Davies et al, 1999, p.191).

A standardisation meeting is held after the examination to establish agreement and understanding of the marking scales. Each year new marking scales (see samples in Appendix 1) are developed for each task. The marking scales are analytical and the markers mark Contents, Organisation, Grammar, Vocabulary and Spelling (see Appendix 1). All the criteria have equal weight.

During the meeting the standardisation packs are distributed; the markers discuss some scripts, which have already been marked and then mark some more scripts together. After the markers have discussed the regular script samples, some

problematic script samples are distributed and discussed. After the meeting each maker receives a pack of scripts for marking at home (on a special form). After a week the markers exchange the packs and mark the second time. The results of the first and the second marker's scores are compared and any scripts that differ in the first and second marking by more than 7 points are selected for a third marking.

Alderson et al (1995) say that correlations will indicate whether the examiner has ranked the scripts in the same order as the other markers: *a reasonable correlation to aim for would be .8* (Alderson et al, 1995 p.132). Therefore, if the correlation is less than .8, we give the script out for a third marking. After the third marking the two closest results are averaged.

7.5.3   Marking the Speaking Test

Before the examination, all the regional group leaders, who are responsible for the appointment of the interlocutors in their region, receive the recordings of the previous years and listen to them together and discuss the scores according to the marking scales. They assess interaction, task achievement, accuracy, fluency and pronunciation. Then they return to their regions and run similar standardisation meetings with the assessors.

During the Speaking Test, while the interlocutor interacts with the test-taker and records the test, the assessor marks the test-takers' presentation. The marks together with the recordings are sent to the examination centre and data-processed. A standardisation meeting is held for the second marking by representatives from each region of the country. After discussion of sample audio-recordings the markers receive a pack of recordings and a form to fill in for marking at home. After a week they return the recordings and the results of their marking are entered into computer. The results of the two markings are compared and problematic recordings (where the markers have given scores with the difference of more than 7 points) are marked a

third time. The two closest scores are averaged to get the final evaluation of the test-taker's performance. The correlation between the first and the second marking was (.59) in 1999.

7.5.4 Assigning cut-off scores

After the introduction of the Year 12 English examination band descriptors were prepared when it became clear what kind of population was taking the test. The English Speaking Union and European Council level descriptors were used as the basis for the development of these bands (see Table 5) Level 10 is Vantage Level, 8 is Threshold and Level 4 is Waystage.

Table 7 Level description

Grade	Level description
10	Free and rich use of language. Even in the most complicated language situations of spoken or written language the use is close to that of a native speaker.
9	Generally free use of language. Even in complicated language situations language is fluent, well organised and mostly comprehensible.
8	The language has been acquired at the level required by the Curriculum of the secondary education of the foreign language. Even in complicated situations language is usually appropriate and accurate. Mistakes are rare and usually not substantial.
7	Adequate language use. In familiar language situations uses language with confidence. Language mistakes do not interfere with communication.
6	In familiar situations understands and uses most of the sentence structures. Adequate use of vocabulary. Sometimes repetition might be necessary.
5	Free use of simple sentence structures and basic vocabulary. Influence of the native language can be observed in language production even in familiar language situations.
4	Limited use of language appropriate for everyday needs. Understands slow speech and simple texts.
3	Can communicate in simple and familiar language situations. Language structures usually contain mistakes, vocabulary can be inadequate.
2	Language use adequate for some familiar everyday situations.
1	Recognises the language. Uses some words and social formulas.

After each test has been marked the result of each of the tests is weighted so that each test has equal weighting: 20%. Inter correlations between the different parts of the tests are calculated and reliability of the objective parts of the tests is checked using ITEMAN (Assessment Systems Corporation 1993). The faulty items are excluded. All the test-takers' results are then placed on the same scale, the mean and the standard deviation are calculated. These are used to categorise the test-takers according to their

score. The grading is made on the curve [see Appendix 7 (histogram of 1999) and Appendix 8 (histogram of 2000)]. Test-takers who are two standard deviations *above* the mean are considered excellent; those who are two standard deviations *below* the mean are considered exceptionally weak (see Alderson et al 1995).

Then the test-development team examines the results of students they know and compares these with the band descriptors to see whether the test-taker's band agrees with his or her actual language competence. If necessary the boundaries are adjusted.

7.6 *Statistical analyses of the examination*

7.6.1 Difficulty level of the examination

Each year's exam paper has to match the specifications as closely as possible. This allows the stakeholders to trust the results and make use of them. Therefore every year during the trialing the test-developers are careful to develop tasks that not only match the test specifications but also are of approximately the same level of difficulty as the previous year. In addition, the difficulty levels of the tasks have to progress from easy to difficult. Since the target population consists of students who have studied English from 3 to 12 years, the difference in the difficulty levels is expected to be great, but the test mean is expected to be 60% in order for the distribution to be as wide as possible (see Alderson et al, 1995).

The 1999 Year 12 exam partly fulfilled these requirements. In the Reading test the first and the third tasks were of the appropriate level of difficulty, the means dropping from 69% to 52% and the first task was the easiest. However, the second task was too difficult: the mean was 32% (see Table 8), so the task was misplaced: it should have been the last task. The mean of the whole reading test was only 49%.

The Listening test tasks were in the right order: the means shrank from 80% to 67% to 60%. This test, however, was too easy, as the total mean for the Listening test was 71%, which was 10% above the desired 60%.

The Language Use and Writing tests were also too easy, the means were 67% and 68% respectively, although the tasks were ordered according to increasing level of difficulty. Nevertheless, the mean of the whole test was 63%, which was close to what was intended. With the overall standard deviation 28% and a minimum score of 13% and a maximum score of 96%, the population was spread adequately across the levels. Distribution took the form of a normal curve for all the tasks together (see Appendix 7, Figure 1), although it was slightly negatively skewed. The histogram of the reading test was positively skewed, whereas writing, speaking, and listening test histograms were negatively skewed. The distribution of the whole test was normal (see Appendix 7, Figures 3,5 and 6).

The histogram of the Language Use test, however, differs: the distribution is bimodal and there seems to have been a split in the population. This peculiarity of the Language Use test can be explained by the security problems of the Year 12 examination in 1999. After the examination took place the Examination Centre received information from two cities that some examination material had been seen before the examination date. It is believed that it was the Language Use test that had suffered most in this security leak. This was the reason why the results of the Language Use test were removed before the scores were calculated.

The histogram of the Language Use test (Appendix 7) reveals the two populations of the Language Use in 1999: the one that had been informed beforehand (mostly in Riga and Daugavpils) and the test-takers from other regions where the test was secure. The data of year 2000 used for Study 4 did not have any security problems.

Table 8 Difficulty level according to the tasks and skills

Task	Min. score %	Max. score %	Mode %	Median %	Mean %	St. dev. %
Reading 37 items	**0**	**90**	**38**	**46**	**49**	**22**
Matching	0	100	91	73	69	22
Gap-filling	0	82	20	35	36	24
Multiple-choice	0	100	33	50	52	27
Listening 40 items	**9**	**100**	**80**	**73**	**71**	**12**
Gap-filling	17	100	88	82	80	13
Multiple-choice	8	100	45	67	67	18
True/false	7	100	73	64	62	16
Language Use 51 items	**1**	**100**	**78**	**59**	**67**	**22**
Multiple choice	0	100	89	68	78	21
editing	0	100	81	56	64	23
Gap-filling	0	100	75	50	57	27
Writing 60 items	**0**	**100**	**70**	**67**	**68**	**15**
Postcard 15	0	100	80	80	77	14
Letter 20	0	100	75	70	71	16
Report 25	0	100	60	60	60	20
Speaking 25 points	**20**	**100**	**80**	**74**	**71**	**17**
Overall	**13**	**96**	**69**	**64**	**63**	**28**

7.6.2 Reliability of the examination

According to Bachman and Palmer 1996, *a reliable test score will be consistent across different characteristics of the testing situation* (Bachman and Palmer 1996, p.19). Alderson et al (1995) say that reliability is relative to the candidates taking the test: a test may be reliable with one population, but not with another. Reliability also depends on the homogeneity of the items: *if the test consists of the same type of items measuring the same skill then the inter-correlations will be high and the reliability index high. If the test contains sections testing different skills in different ways, these sections will not correlate with each other and the reliability will be lower* (Alderson et al 1995 p.89).

The Year 12 examination consists of different types of items measuring different language skills [see correlation matrices in Appendix 13 (1999) and Appendix 14 (2000)]. It is also taken by a population with a wide range of language proficiency levels and this explains why the reliability of the objectively marked tests is so low

(.70). This can be seen in Table 9, which shows reliability indices (Alpha for the objectively marked tests and correlation with the total result for all the tests.

Reliability for Writing and Speaking test in 1999 caused a problem, as I received the data after they had already been processed and for writing there was just the result after the averaging, so I could not see the correlation between the first and the second marking in 1999, but in 2000, I had the full data (see Chapter 10). In Speaking test the correlation between the first and the second marking was .572, which is not sufficiently high either. The examinations centre has decided now to renew the marker training sessions, so hopefully the reliability of marking will improve.

After the first marking correlation between the Speaking test results and the total results was 0.769; after the second marking (of the recorded version) it fell to 0.608, but after the third marking and averaging it was 0.823. This suggests that the marking done from the tapes is not as reliable, or measures a different ability from that of the live marking. The reliability of the objective part of the test was .947, which was satisfactory. The correlations between the results of each part and the total range from .542 (the first writing task) to .858 for the very difficult reading task. This suggests that all the tests had impact on the final score as the correlations are below .9. At the same time the correlations are not too low to imply that all the tests were measuring language skill.

Table 9 Reliability of the Year 12 examination in 1999

Task	points	SEM	Alpha	Correlation with the total***
Reading	**37**	**2.4**	**0.895**	**0.885**
Matching	11	1.3	0.711	0.640
Gap-filling	20	1.7	0.866	0.858
Multiple-choice	6	1.0	0.622	0.643
Listening	**40**	**2.3**	**0.776**	**0.812**
Gap-filling	17	1.2	0.708	0.683
Multiple-choice	12	1.4	0.555	0.683
true/false	11	1.3	0.421	0.558
Language Use	**51**	**2.9**	**0.932**	**0.700**
Multiple choice	19	1.7	0.815	0.642
Editing	16	1.7	0.799	0.608

Task	points	SEM	Alpha	Correlation with the total***
Gap-filling	16	1.5	0.875	0.671
Writing	**60**	**1.3**	**No data***	**0.761**
Postcard	15	0.2	No data	0.542
Letter	20	0.5	No data	0.670
Report	25	0.7	No data	0.725
Speaking	**25**	**0.6**	**.572****	**0.823**
Total	100%	0.2	Not appropriate	Not appropriate

* In 1999 CEC data package did not contain the results of the first and second marking

**Correlation between the first and the second marking

***See Intercorrelations between the different tests in Appendix 12

7.7 Validity of the examination

Messick (1989) presents validation as *an integrated evaluative judgement of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of inferences and actions based on test scores* (quoted in Bachman 1990, p.236). Although tests are as different as the people who construct them, there are certain criteria that have to be examined before one can consider a test appropriate for a given situation. Every test is only as good as it is appropriate for the purpose.

Although a thorough validation of the examination would require a whole study of its own, Table 10 gives an overview of an attempt to validate the Year 12 examination using the Alderson et al (1995) description of different types of validity.

Table 10 Validation of the English language examination

Validation	Method	Evidence
1. Face validity	The investigation of the acceptability of the examination to users	A compulsory examination in 2001. In addition the Tertiary Education Council has accepted that from 2004 the examination should be an entrance examination for all universities. The Council of Ministers has included the centralised examination in the government's plans and as a result a state budget is allotted to it. The French Embassy and the Goethe Institute financed groups of French and German teachers to be trained so that they can produce similar examinations
2. Content validity	Comparison of tasks with the test specifications to ensure representativeness	Each year before the exam Ministry of Education experts, the administration of the Curriculum and Examinations Centre, carry out a check of the content, as well as secondary school teachers after the examination.

Validation	Method	Evidence
3. Response validity	Co-operation on the part of the test taker: ability to understand instructions, motivation, willingness to obey all the conditions of the examination	Established by questionnaires to the students (see Chapter 10)
4. Concurrent validity	Comparison of the test scores with some other measure for the same candidates taken at roughly the same time (parallel versions, self assessment, teachers' ratings etc)	Self-assessment is collected yearly and compared with performance on the examination. The universities are organising their entrance examinations following the same guidelines as the Year 12 examination. The data have not been used to compare the results of the examinations, but, unofficially the Year 12 examination certificate has become a reference point for comparison with different university entrance examinations.
5. Predictive validity	Follow-up studies of the test-takers	The follow up of the test-takers with high and low scores in tertiary education has not been done officially, but because they teach at the university, the members of the examination development team have been able to follow the progress of the test-takers. The fact that many universities have volunteered to accept the certificates year after year suggests that they are satisfied with the results
6. Construct validity	Collection of evidence that underlying theoretical constructs that are being measured in a test are themselves valid.	External validation of the examination for the British Council in 1996 and 1997 found that the Year 12 examination is measuring the test-takers' language proficiency in accordance to the test specifications and the established procedures for ensuring reliability of the test are satisfactory.
7. Reliability	Training and monitoring of the examiners and correlations between the first and the second marking for each marker of the subjective tests, Item analyses and pre-tests for objective tests.	Carried out yearly by the examination centre (see Appendix 10).

The attempt to validate the test has in fact uncovered both problems and assurances that the test is measuring what it is intended to be measuring. Nevertheless there are reassuring findings: the external validator's conclusion in 1997 was that the test conforms to language testing standards. The face validity of the test is satisfactory, the test-takers' teachers rigorously examine the content validity each year as each examination material is scrutinised by the teachers who are preparing the students for the examination.

The greatest problem, the threat to the reliability of the test, which was caused by security problems, was publicly acknowledged in 1999 and did not repeat either in 2000 when the third study took place or in 2001. Another less publicised problem is the fact that there are not enough specialists in the country to ensure the examination's content validity and although there are many bodies responsible for this, some decisions are made by officials without consulting anybody. For example, Writing task 1 (an email) was included in examination without any pre-testing or moderation simply because the administration decided that the previously prepared task 1 was too easy. Hopefully, as the level of expertise and the number of experts grows, such cases will become impossible.

Chapter 8   Study 1 - Observation

If Chapter 7 introduced the language performance measurement instrument then Chapter 8 presents the test-takers and the environment of my research as well as that of the Year 12 language examination.

8.1   The aims

The aim of the observation study was to

- observe the examination procedure

- observe the filling in of the questionnaires of the study

- observe signs of anxiety in the behaviour of the test takers during the Year 12 examination.

I will use Oxford's framework of language anxiety signs to analyze my observations. Oxford (1999) proposed that certain aspects of foreign language anxiety could be observed during language acquisition process. These aspects take the form of general avoidance, physical actions, physical symptoms and other signs, which depend on the culture of the country (see section 4.3.).

8.2   The site of the study

I chose two different schools for this observation: School 1 is a traditional large bilingual country school, which accepts all the students (Russian and Latvian) from the region. As it is situated not far from the country's capital, the best students traditionally leave it for better schools. I observed both the written and the spoken part of the examination there as well as the filling in of the questionnaire for the main study.

School 2 is one of the most popular Riga grammar schools. It is difficult to enter (students have to compete to enter this school). The school is monolingual as there are only Latvian students there. I observed the speaking test on day two and interviewed three test takers on the next day.

I was granted permission to observe the examination by Ministry of Education and was accompanied by two observers from the Hungarian Ministry of Education.

8.3 The written part of the examination

The examination started at 10 o'clock. We arrived at the school at 9 o'clock, so that we could introduce ourselves to the school administrators and the exam personnel.

8.3.1 Personnel

The examination personnel consisted of 2 people for each group of test takers: an administrator of the examination and an observer from a different school (appointed by the Local school board). The observer had to fill in a special checklist to guarantee that the correct examination procedure had been followed.

8.3.2 Test takers

There were 40 test takers in School 1. They were divided into 2 groups and took the exam in 2 separate rooms. When we asked before the exam how they felt and whether they were nervous about the examination, they said they were not, but one could feel the presence of tension: as they all had pens with them they were all hitting them against the palms of their other hands and four test takers were also tapping their feet against the floor. Some were telling jokes, and one could hear constant nervous laughter from the waiting hall.

8.3.3 The site

The room was big, light, temperature and humidity in the room were normal, some talking was heard outside the window, but this soon stopped.

8.3.4 Data collection

The observation process was overt, as all the observers were introduced. I took notes looking for the signs of anxiety throughout the test (thus it was a non-participant observation). I was given all the instructions for the administrators of the exam and examination material and could easily follow the procedure. All the observers were seated in front of the classroom at the board (where the floor was slightly higher, so that I could see all the test-takers).

In the field notes (Table 11) I recorded the examination procedure (including the background noises, as I noticed that the test-takers were affected by them) test-takers' and test administrator's actions as well as all the physical and emotional signs of anxiety that I could observe.

Table 11 Field notes of the observation of the written part of the examination,

Test and Time	Administrator's actions	Test-takers' actions and reactions
9.30		The school administrators arrive with the examination material and ask the test takers to examine whether it is intact. At first nobody volunteers to examine the pack, then one test-taker volunteers and others following the first student's example join in. They announce the pack to be in perfect order and return it to the school director who opens it and divides it between the exam administrators responsible for each room.
9.45	The administrator introduces herself, asks whether everybody is well and they are ready for the examination. The administrator explains the timing of the examination that is written on the board.	The test takers are let into the room after the administrator has examined their passports. They leave all their belongings in one corner of the room, find their seats marked with their names and code numbers on little stickers and sit down.
9.55	The administrator explains that the Ministry of	After some hesitation 3 students volunteer, then 2

	Education has given them some anxiety questionnaires and asks whether anybody would want to fill them in.	more. The ones who volunteered are the ones who seem the least bothered by anxiety. They look through the questions before the examination and one starts filling in the second part of the questionnaire.
		One boy, who did not take the questionnaire, and sits at the first desk, right in front of the administrator, is blushed, continuously stretches his neck and sweats: his shirt collar is open.
Reading test 10.10-10.50	The administrator distributes the Reading test papers.	The test-takers start working and the bouncing of pens and rubbing of the hands stops. One can hear the teacher's voice from the neighbouring room.
10.20		Two test-takers are already doing the second task. Two boys are still red; one is nibbling his pencil.
10.30		Eight test-takers are still working on the second task, one does not work, wriggles, looks out of the window, the bell rings, there is noise upstairs.
10.35		One test-taker seems to have finished, looks nervous (he is grazing his pencil), but 4 are still doing task 2.
10.40		The test-takers are finishing one after another, they again start rubbing their hands and bouncing their pens, one starts tapping feet and constantly checks the time and fills in the questionnaire.
Listening test 10.50	Administrator distributes the Listening test and collects the Reading test. The recording is good quality, the local school board observer is listening at the back of the room, the administrator is so careful as seems to have stopped breathing. However, this does not interfere with her looking constantly over Sergey's shoulder at what he is writing. The administrator is still standing at Sergey's desk and once in a while looks at what he is writing.	All the students receive their Listening tests, write the code number and the listening test starts with the recorded explanations, adjusting the sound and a pause for reading through the questions for the task 1. During the pause one test taker is rubbing his hands, whispering to the others, puts on sunglasses and at last concentrates and starts reading. The signal sounds and all jump.
		After the second task there is a pause and one girl fills in the questionnaire, looks at the boys and giggles.
11.10		Listening test ends, everybody looks relieved, most of the test takers are sighing, straightening shoulders; two are filling in the questionnaire.
Language Use 11.15	LU test is distributed and Listening test collected.	
		Everybody is still writing the 1st task. The school bell rings, and is ignored.
11.30		Six test takers are filling in the 3rd task, tiredness seems to have set in.
11.40 11.45		The bell rings, everybody raises head, looks at the watch, 4 test takers go on writing, the break is announced. Sergey is still Writing task 2, everybody leaves the room.
Break 11.45 11.55		All the administrators and test-takers leave the room. Test-takers return to their places.

Writing test 12.00	Writing test is distributed	Everybody is writing, Sergey is looking at the paper and not writing
12.15		Sergey is still bright red and looking at the Writing task 1 and not writing.
12.17		Sergey starts the 1st task.
12.20		Everybody is contemplating 2nd task, shuffling the paper.
12.30		Three test takers are writing 3rd task, gradually finishing and filling in the questionnaires.
	The administrator takes the exam papers and questionnaires. She looks through the questionnaires with interest.	The first 3 test takers are leaving the classroom.
13.10		
13.15		All the test takers leave.

The analyses of the results of the written part of the examination can be found in section 8.6, together with the Speaking test observation result analyses.

8.4 Speaking test, Day 1

8.4.1 Test format

The test was prepared according to the test specifications and consisted of 3 parts: picture description, questions/answers and role-play. There were 6 papers (Nr.1-Nr. 6) for the first day.

8.4.2 Test takers

There were six test takers in this group (four boys and two girls), all at an intermediate level of language and all were taking their Speaking test after the written part of the test and the break (1,5 hour long). Table 12 shows the code numbers of the test-takers, their language proficiency and the paper they drew out of the whole lot (see Marking Scale in Appendix 1).

We can see that 3 students out of 6 had paper number 2. The test-takers evidently chose the most crumpled papers that their friends had taken before them and evidently had told them about (see field notes in Table 13).

Table 12 The results of the evaluation of test-takers performance in School 1

Code	Paper	Speaking proficiency (my marking)					
		Commun. Strat.	Task ach.	Accuracy	Fl.uency	Pron.	Speak. Total %
001	3	4	5	2	3	4	60
002	4	3	4	3	4	4	60
004	2	5	5	3	4	4	70
005	2	3	3	3	2	3	47
006	5	4	5	2	3	4	60
007	2	2	3	1	2	2	33

8.4.3 Personnel

The interlocutor was a young American volunteer from the Peace Corps; he had been teaching the test takers. The assessor was a teacher from a different school who had been selected by the local school board. There were also 2 observers: me, and one representative of the Hungarian Ministry of Education.

8.4.4 Data collection

All the observers were placed at the very end of the same room where the written part of the examination had taken place; the assessor was also placed inconspicuously at her desk out of the test taker's sight. The interlocutor placed himself at the desk, arranged the papers on a separate desk, checked the tape recorder and invited the first test taker in. We were given the list of the test-takers, their code numbers and Speaking test (both interlocutor's and test-taker's) papers. I took notes (see Table 13) throughout the interviews, recorded the test-takers code number and the paper they chose as well as marked their performance using the same marking scale as the assessor.

Table 13 Field notes, Speaking test, day 1, School1

Student's code Nr.	
001	Sergey comes in, his teacher (now calm) ceremoniously shakes his hand and introduces him to the assessor, tells him that they would proceed exactly as they did when preparing and asks him to choose one of the 6 papers. He chooses a paper and they both started laughing and decide that it is quite similar to the last years' paper they used for preparing for the exam. Nevertheless Sergey is tense, his face is as red as in the morning.
	The tape recorder is switched on, and one minute's pause sets in. After a minute he starts talking, his speech slow, but firm, he describes the picture, haltingly asks questions, the teacher answers the questions about his picture and asks the questions for task 2. Sergey speaks slowly with unnatural pauses and evident effort, makes many mistakes in accuracy. The teacher is very calm, understanding and supportive and they both manage also task 3 (role play). They finish; teacher thanks him and wishes him to have a good rest. Sergey leaves, relief evident on his face.
	The teacher puts the student's paper back. Because of the test-taker's sweating palms and fidgeting the paper has become all wrinkled.
002	The next test taker enters. The same procedure is followed; the interlocutor treats the test-taker with invariable support, uses the same reassuring phrases, the same shaking of hands and supporting attitude. The student is tense; his hands are slightly shaking when taking the paper but otherwise he seems to be in control.
004	The next test taker is a girl; she is also tense. The teacher's manner does not change. Her language is very laborious and slow; she manages to do approximately half of the tasks and is very slow and inactive in responding. She remains tense throughout the interview.
005	The next test taker is more relaxed. When asked to choose the paper he takes the most crumpled and damp one. It seems that the test takers have been exchanging information behind the doors, and the earlier test-takers have told the later test-takers about the papers' contents.
006	The next girl performs well the first task in spite of her problems with language, but during the pause before the second task starts laughing without any evident reason and cannot stop for several minutes. In the end it is difficult to say whether she is laughing or crying. With the teacher's unchanging support she manages to calm down.
007	The next boy again and seems to be perfectly calm although his language is worse than any of his predecessors (accuracy 1, fluency 2, pronunciation 2). He is often not responding to the teacher's questions and never initiates himself (communication strategies 2, task achievement 3). The teacher's support is invariable.

8.5 Speaking test, Day 2

The instructions of the Curriculum and Examination Centre (CEC) envisage that the big schools that have more than 20 test-takers may hold the Speaking test two days, but they have to use different examination papers (Nr.7-Nr.12) that are prepared for the second day. So the test-takers cannot be informed by the test-takers who took the exam the first day.

8.5.1 Test takers

There were 11 test takers in the group (5 boys and 6 girls) that I observed in School 2. The test-takers were all casual; nobody was dressed up. They had all studied in this school, knew each other and they knew the premises. Their language proficiency was higher than in School 1 (see Table 14).

Table 14 The results of the evaluation of the test-takers' performance in School 2

Code	Paper	Speaking proficiency (my marking)					
		Commun. Strat.	Task ach.	Accuracy	Fl.uency	Pron.	Speak. Total %
045	10	6	6	4	6	6	93
023	8	6	5	5	5	5	87
013	6	3	4	3	3	4	57
030		4	5	5	4	4	73
031	6	5	5	5	5	5	83
047	9	4	5	4	4	4	70
035	10	4	5	4	4	5	73
021	8	6	6	5	6	6	97
028	9	5	5	4	6	4	80
025	10	6	6	5	6	5	93
015	6	5	6	4	5	4	80

8.5.2 Personnel

The personnel consisted of two people: the interlocutor, who was the student's teacher and the assessor, who was a teacher from a different school. There were also two observers present: a Hungarian Ministry observer and myself.

8.5.3 Data collection

In School 2 on the second day of the English examination there were 2 teams of examiners. The assessors and interlocutors received their papers half an hour before the examination started and read them in a separate room.

We proceeded to a room far away from the school noise, the room was quite small, it was rather dark, but not too dark for reading. The interlocutor arranged all the papers on a desk (they were all placed in plastic pockets). The first test taker was invited in at

9 o'clock. I took notes of both the interlocutor's and test-takers' actions (see Table 15) and marked the test-takers' language performance throughout the examination.

Table 15 Field notes: Speaking test, day 2, School 2

Test-taker's code number	The test-takers' actions	The interlocutor's actions
045	The girl is very calm, talks fluently, is very responsive, initiates often although she makes some grammar mistakes. The interview ends as abruptly as it started, without any warmth from the interlocutor.	The interlocutor's manner is absolutely formal. She responds to her student's greeting, but does not smile and does not say anything to welcome her. The interlocutor records her code number and number of the paper and the examination starts. There is no warm up phase.
023	The student is absolutely in control and calm. During the second task the teacher is evidently not listening as she reads a wrong response, the test taker tries to repair this and somehow to tie the inappropriate response to some context. When the girl leaves she is evidently relieved to have finished.	The next girl is treated in the same way. The teacher responds even more abruptly, there is practically no eye contact. Having asked the question from the paper, the interlocutor looks out of the window
013	The test taker's language is weaker and he is more under stress than the previous test-takers. The boy is just struggling to do the task. The cassette ends during task 2, the test taker starts, blushes, the interlocutor puts in a new cassette and resumes questioning the test taker in falling tones. All the discourse is loose, not connected; in task 3 neither the student nor the interlocutor can follow their roles naturally.	The next test taker is a boy; the interlocutor meets him with a feeble smile and seems to be more relaxed during picture description. It seems that she does not notice what state the boy is in. When the interviewer gets to Task 2, she again reads out the questions of task mechanically without checking whether the student has finished his answer on the previous question and interrupts his thought.
030	Test taker has to fight against the quick questioning to finish his thought, but he is not worried by the attitude and wins the fight	Interlocutor reads out her cues from her booklet in a very formal manner.
031	The next test taker is a girl she seems to be very anxious from the very beginning, does not look at the interlocutor. Again the person is being interrupted when she is thinking during the task 2.	The interlocutor does not look at the test-taker. The personal questions about the meaning of home to a person in this situation seem surrealistic.
047	The next girl, has the same problem with getting the eye contact with the teacher, speaks with effort, dry mouth but copes well	Attitude as before
035	The girl arrives already red in the face, keeps wringing her hands during the interview, her voice shaking, but performs well in spite of the interlocutor's looking out of the window	Attitude neutral, a few signs of support (for example, sometimes nodding her head).
021	The next test taker seems to be stressed, but in fact is in control of himself and seeks the teacher's eye contact very persistently, he sighs after every two sentences and gradually starts stuttering.	The interlocutor's reading from the paper becomes unintelligible. The teacher covers her face with her hands in reaction to the-test-taker's persistent attempts to get an eye contact with her.
028	The next speaker talks very quickly, seems not to be bothered by the tasks or the interlocutor.	

025	The test-taker seems to be extremely worried, his hands trembling, his voice shaking, he strokes his neck, sighing aloud during the pause for reading, the breath comes with pressure during the presentation and he stutters once in a while in spite of his good language	After this test-taker's presentation the assessor asks the interlocutor to be more supportive.
015	The last test-taker has difficulty with concentration, cannot remember the word "chair" in English, cannot understand what the teacher says, she is constantly looking around in spite of her good language	Attitude as before.

8.6 Findings of Study 1

Oxford (1999) mentions four categories of signs that signal the presence of foreign language anxiety during foreign language acquisition (general avoidance, physical actions, physical symptoms and signs depending on culture). All of them could be observed during the Year 12 examination (Tables 11, 13 and 15).

8.6.1 General avoidance

General avoidance could be observed in the case of one student, who during the written part spent 15 minutes looking at Writing test task 1 without writing a single word, just looking at the paper where he only had to write 30 words for that task (Table 11).

'Lack of volunteering' (Oxford 1999) could be observed before the examination when the school administration asked the test takers to examine the test materials package to see whether it was intact, there was a pause as none of the test takers would take the package. When at last one of the boys took it all the other test takers were very interested and examined it carefully (Table 11).

Another case of lack of volunteering could be seen when the anxiety questionnaires were offered, there was a pause again until one test taker who was brave enough to

take one and then others volunteered. The students who seemed most anxious about the test did not take a copy of the questionnaire (Table 11).

8.6.2 Physical actions

The physical actions category was the easiest to observe before the written part of the examination: out of fourteen test-takers present, ten were hitting their pens either against the palm of the other hand or just shaking their pens; four test-takers were tapping their feet constantly as well as shaking their pens. During pauses in the written part of the examination, I observed squirming, fidgeting, the nervous touching objects on the table and otherwise jittery behaviour practically from all the test takers. During the spoken part of examination I observed two cases of stammering and when I spoke to the test-takers afterwards, neither of the boys stammered in everyday situations. One boy stroked his neck during the performance. I also observed some test-takers whose hands sweated so much that the test papers were practically ruined (School 1) and I also observed that some test-takers had trembling hands (School 2).

One girl (School 2), while taking the Speaking test, was constantly squeezing her hands in a way that was painful to observe. When I asked her afterwards, she said she had never noticed it. She said " I must have been nervous".

8.6.3 Physical symptoms

Physical symptoms could also be observed: one student (School 1) was tense through out the written part of the examination; he was constantly rubbing his neck and moving his head. His face was also dark red throughout the written and spoken part of the examination. One student in School 2 had difficulty in breathing: his breath came with pressure, by fits and starts. He and another student also sighed constantly.

8.6.4 Other signs

Other signs depending on culture is a wide category. Oxford (1999) gives an example of conversational withdrawal: this could be observed in the case of the student 11 in school Nr.2 during the speaking test: the girl did not listen to the interlocutor and spent a lot of the time looking around. Lack of eye contact was a sign of anxiety for this girl, but it could have been just a reaction to the interlocutor's behaviour.

8.7   Conclusion

Having observed the procedure of the Year 12 examination in Latvia, I came to the following conclusions:

1. in the schools where I observed the examination, it proceeded according to the instructions issued by the Curriculum and Examinations' centre

2. the test-takers were placed in appropriate setting, they worked independently and were not disturbed by exceeding background noise

3. the interlocutors' conduct varied from supportive to hostile

4. all the signs of foreign language acquisition anxiety listed in Oxford's (1999) framework could be observed during the foreign language examination (both in its written or spoken part)

5. the test-takers interviewed in a supportive manner exhibited as many signs of anxiety as did the ones who were interviewed in a hostile manner.

I will discuss this further in the next chapter (see section 9.2), where I describe the results of the interview with the test-takers in School 2.

Chapter 9   Study 2 - Interviews

I carried out two interviews, one after the examination in 1999 and another in 2000. The first interview was free, the second semi-structured, but the focus of both the interviews was the causes of anxiety in the Speaking test.

9.1   Pilot interview in year 1999

This section presents the pilot interview of 1999, the interviewees, their experiences during the Speaking test and analyses of their comments. Full transcript of the interview can be found in Appendix 3.

9.1.1   Aims

As the Speaking test was taken by students at different times, each of them arriving and leaving separately and it was complicated to organise the collection of the questionnaires after the Speaking test, I decided to use a different research format, an interview, to find out whether anxiety was perceived during the Speaking test and if so, what its causes were. The main aim of this interview, however, was to learn how to use interview as a research method. I decided to include the analyses of this interview in my thesis only after interviewing the test-takers of Year 12 in 2000, because I found that their reactions were so different (see the discussion in section 9.4.).

9.1.2   Test format

The Speaking test is prepared according to the test specifications (see section 7.4.4 and Appendix 1 for test materials) and consists of three tasks: picture description, questions/answers and a role-play.

9.1.3 Procedure

After the examination I took the tape of the recording of the Speaking test and played it to three test takers and asked the interviewees to comment while listening. The interview took place in 5 days after the examination, it was conducted in Latvian and was recorded. Afterwards I transcribed and translated it (see the translation of the transcript in Appendix 3).

9.1.4 Site

The experiment had to take place in the publishing house where I worked during the day, so there were telephones ringing and people arriving to ask questions. Unfortunately I could not find a better place and time that would suit all of us.

9.1.5 The interviewees

All the girls (Ieva, Dagmara and Karina) were from the same form, they were 17 to 18 year old, and their language level was upper intermediate. I knew them all well as they were friends of my daughter.

9.1.6 Results of the 1999 interview

The interviewees did not consider the Speaking test to be more anxiety provoking than the other tests.

Dagmara said: *When you have done Grammar, the exam is over. Speaking is not a problem.*

Nevertheless, there were some aspects that did cause anxiety. I will first look at anxiety and its causes (foreign language anxiety and test anxiety) and then examine its interaction with other affective variables (see Table 16).

Table 16 Test-takers' comments on the causes of anxiety in 1999

Test-taker	Comment	Cause
Ieva	*I did not understand the word "itinerary", I asked him what it is and I was so stupid, I did not understand his explanation; and the more you worry the more puzzled you become	*lack of vocabulary
	*I was also worried about the picture and forgot to ask questions	*task format
	*How do you understand it (the phrase we had to comment on) "At a dinner party one should eat wisely but not too well and talk well, but not too wisely"? We had one minute, but it was not enough really	*test anxiety
	*I was puzzled by those sandwiches, I thought maybe this was his own idea? *Sometimes I need a word, I remember it in German and forget it in English, and you start making funny constructions. It means we do not have enough practice in English. *The formality at the beginning before the dialogue, everybody listening, you get afraid of your own voice.	*foreign language anxiety
Dagmara	*I did not know how to say 'Midsummer night' or 'St. John's'? *I had seen the word 'similarity', but never used it and did not know how to pronounce it. *And the same with those "museums", I did not know about the stress, because in other languages it is different.	*vocabulary *pronunciation
Karina	I was not afraid, but I was totally tongue tied. I came in and thought, why aren't I worried?	*test anxiety *self-evaluation

9.1.6.1 Foreign language anxiety

As we can see in the Table 16, the most common cause of anxiety during the Speaking test is vocabulary (as with the written part of the examination, see section 10.1.5 and 10.2.5); two of the three interviewees commented on the problems of vocabulary in the tasks. There were also other problems of language: pronunciation, grammatical constructions, but they were mentioned only once. In addition Dagmara said that she suddenly remembered the word in German and could not remember it in English. According to Stevick (1999) this relates to one of the roles anxiety: that of control of access to memory.

9.1.6.2 Test anxiety

In Task 2, the test-takers have to comment on a statement, for example *At a dinner party one should eat wisely but not too well and talk well, but not too wisely*. Although they are given one minute to think it over, it seems that the test takers have difficulty focusing during this task.

Another problem for a test-taker is time:

Karina says: *I was playing for time.*

Helen says: *I tried to think of something to say.*

It seems that test anxiety interferes with attention and the ability to concentrate, and as a result the test takers need more time. This is what Stevick (1999) calls Role 2, 'interference with thinking', because too many areas are kept active (test situation, perceiving, reacting and thinking).

Although the aim of the interview was to examine the causes of anxiety, the test takers, while commenting on their experiences during the test, were providing evidence of the other affective variables: their attitude to the interlocutor and their goals.

9.1.6.3 Interaction between test anxiety and motivation of the test takers

The first role of affect (Stevick 1999) is organisation of our attention and behaviour where it is most needed according to our goals and needs (see also Power and Dalgliesh in section 2.2.2). This can very well be seen as Dagmara says:

I was not worried at all, I had a horrible cold and could not breathe, but I did not mind at all.

For her the goal is genuine interaction during the examination and she says about the interlocutor: *He is that sort of a person you want to talk to, even though I do not know him.* We can see from this that Dagmara was intrinsically motivated to interact with the interlocutor during the Speaking test simply to get to know him better.

During the examination the test takers have to obey the rules of the examination, and they want to be sure that they are doing everything in the correct way. During the

written part of the examination you are on your own and have to decide by yourself. In the Speaking test if you are not sure about the interlocutor, it can add to your test anxiety. This however was not a problem for Dagmara as she compared the interlocutor with their regular English teacher (the interlocutor in School 2, see section 8.5):

I think that in this exam even the two-hour pause will be recorded…. With the other teacher I don't know, maybe she did not understand something about the task. She always understands everything in a peculiar way and cannot find a way out if there is a problem.

9.1.6.4 Interaction of the affective and meta-cognitive strategies

Because the goal of the test-takers is to have a normal conversation and because of the feeling of security, a whole chain of interaction strategies is activated. Instead of just reacting to the test cues the test takers start initiating and eventually it is difficult to separate the examination cues from the real interaction:

Ieva says: *Suddenly I though of asking him something personal. He got confused, but was not shocked. I also got confused, I thought that he decided on his own to remind me to take the sandwiches.* (They were doing a role-play where the students were getting ready to make trip in the mountains).

All the three test takers say that they had no lack of ideas about what to speak:

Dagmara says, *I wanted to speak more and more.*

Ieva says: *Sometimes I think I have* already *talked too much. The questions were such that I could talk and talk.*

Karina says: *I had a feeling that I could have had 5 more tasks.*

This is an example for what Damasio (2000) calls a positive body state (the generation of images is rapid, the diversity of images is wide, and reasoning maybe fast, but not necessarily efficient (See section 2.1.2). Dagmara also says: *It was not like an examination, it was like a game.*

On my part I have to say that it was difficult not to share the excitement of the girls as they commented on the proceedings of their interview; they talked fast, one over another and all were enthusiastic about their experiences during the Speaking test. I started the interview in a sympathetic tone with a question 'What it was that worried you about the Speaking test', and finished it off envying their experience during the Speaking test. Unfortunately, the main interview turned out to be just the opposite.

9.2   The main interview in year 2000

9.2.1   Aims

The main reason why I organised the second interview was my experience during the observation in School 2, I felt I had to speak to test-takers who had been treated so inadequately (see Table 15). There were also other aims:

- examine the causes of anxiety in the Speaking test

- examine the interaction between the affective variables

- see how the test takers' comments depict the interaction between affective variables and performance strategies

- compare the results of the two interviews (1999 and 2000).

9.2.2 Procedure

After observation of the administration of the Speaking test at School No.2, I asked if any of the test takers would agree to meet me to discuss their experiences during the

interview. Three girls volunteered (I had not met any of them before). We arranged to meet the next day after the examination was over.

As I was not allowed to take the recording of the test-takers' performance from the Curriculum and Examination centre, I gave the test takers both student's and teacher's papers of the Speaking exam, and asked them to comment on their experiences while answering each of the questions of the examination papers. I asked them questions based on my observations during the examination (see Table 15). The interview was conducted in Latvian, it lasted for one hour and was recorded, later transcribed and translated (see the translation of the transcript in Appendix 3).

9.2.3 Site

The interview took place in a quiet classroom (a different one from the room where the examination had taken place) in test-takers' own school.

9.2.4 The test takers

The three girls were from the same form. They were 17 to 18 years old, their language level was intermediate or upper intermediate (see Table 17). I had never met them before the examination, therefore the tone of the interview was formal at first.

Table 17 Language proficiency of the interviewees

Test takers	Code	Paper	Speaking proficiency					Speak. Result	Mean profic.
			Commun. Strat.	Task ach.	Accuracy	Fl.uency	Pron.		
Antra	045	10	6	6	4	6	6	95%	84%
Helena	023	8	6	5	5	5	5	92%	73%
Elina	035	10	4	5	4	4	5	73%	58%

9.2.5 Results of the study

The analyses of the results of the interview are divided into two parts: causes of anxiety and the interaction between different individual variables during the Speaking test.

9.2.5.1 Causes of anxiety

I started the second interview as I had the first, by asking the test-takers if they were anxious during their Speaking test. When the test takers started commenting on their feelings and their worries during the examination (see Table 18), their greatest anxiety, once again was vocabulary, not only the unknown words but also of the words they had forgotten, were told and immediately forgot again:

Helen: *I liked my picture, when I saw it I thought, I will say that it is a biology lesson, but then I forgot what 'skeleton' is in English. I said a poster with bones on it and I asked the teacher 'how do I say it in English', she said 'skeleton'. I thought it was exactly as in Latvian; and immediately forgot the word again, so funny. Otherwise I did not have any problems.*

Table 18 Test takers' comments on the causes of anxiety in 2000

Year	Name/ Code Nr	P. Nr.	Comment	Cause
2000	Antra (045)	10	*There were some specific words I did not know (seat, toilet, sponge). *"Education polishes good natures and corrects bad ones", I tried to think of something to say.	*Vocabulary problems, *Lack of ideas
	Helen (023)	8	*When I was waiting for the interview I was a bit worried, but when I came in all my anxiety was gone. *When you know that you have only one minute and you are not sure about your language, it just beats me, the time limit. *I forgot how to say 'skeleton' in English, asked the teacher and forgot again	*state anxiety *time limit *vocabulary + test anxiety lack of confidence

	Elina (035)	10	*When preparing for the examination, we recorded it and there were many pauses and when you listened to it, it was awful. *I was afraid of becoming too philosophical; if I do not know one word I forget what I am going to say. *I did not know the word 'spot'. *I depended on the teacher's reaction… she made such a terrible face.	*lack of confidence *language anxiety *vocabulary *interaction

Apart from vocabulary, there were also other causes of anxiety: time limit (Helena), lack of confidence (Helena and Elina) and, surprisingly, lack of ideas, which was in a striking contrast with the exuberance of ideas and willingness to speak that could be felt in the comments of the interview in 1999. None of the test takers mentioned the interlocutor or assessor as a cause of anxiety, which again surprised me, as during the observation I had been struck by the inappropriacy of the conduct of the interlocutor, for example, lack of eye-contact or in fact and the obvious lack of interest in what the test-takers were saying.

9.2.5.2 Interaction between anxiety and other affective variables

Anxiety was not the only affective variable that was activated during the language test. Unprompted, motivation and confidence were also mentioned when the students were asked about anxiety. It seems that the different affective variables interact: confidence and motivation are used to sooth anxiety caused by the approaching examination as the test takers say that they spent time and effort preparing for the test:

Antra: *I cannot say that I was worried, as I knew the papers and we had practised with the papers for the last year and I knew that the pictures are rather odd and can have all kinds of objects in them.*

However, none of the test takers said that they wanted to ask something themselves or, that they enjoyed their interview. Even when Helen said she liked the tasks, she was referring to the fact that she had done a similar paper when preparing for the examination:

I liked the other questions, because I remembered that one day, before the examination I caught the teacher during one break and we tried out the paper from the previous year, and its theme was "Education", the questions were quite similar. I was so relieved that it had all been discussed beforehand.

Thus doing the test well was their main goal and their motivation was extrinsic. When asked whether they were worried by the lack of support from the interlocutor, they immediately made reference to their confidence, another affective variable that is used to compensate for the anxiety. The test-takers said they had confidence in their knowledge of their teacher. Antra said that she was not surprised by the distant manner of the interview as they knew the interviewer well:

Antra: *I don't know, I am used to her, with this teacher we do not have any relationship or any attitude, and in fact it did not disturb me.*

Helen, however, immediately added that her reaction had been totally different when she had seen the teacher for the first time, 3 years ago:

Helen: *Now that I have studied with her I am used to her, but last year, no, in form 9, when I had not studied with her, I was worried. I entered the room, gave her some flowers, said hello. I was looking at her and still smiling, but she looked out of the window. Everything died in me. I thought, that's it, I do not need anything any more, she is not interested. She is going to assess my presentation from how I entered the room…*

Bachman and Palmer's (1996) description of the role of affective schemata says that it not only decides the effectiveness of our reaction to the task, but also whether we choose to do it at all. For Helen, however, it is not the task, but the interlocutor's manner that stopped her from performing as well as she could.

Apart from knowing the teacher, there is also the students' confidence in their knowledge of the language; both of these help the test takers to compensate for the lack of support from the interlocutor:

Antra: *Whatever, we have the feeling that English is like a native tongue, and then we do not have to worry. If I am sure of the things I know, it does not matter to me whether she is looking at me or at the ceiling. Some time ago, when this feeling was not so strong, there were times when such attitudes did undermine my confidence though.*

Elina, who did not feel confident about her English, depended more on the interlocutor's manner and the fact that she knew the teacher was her only support:

In task 3, I asked whether I could get information instead of asking what information I could get. The teacher was puzzled. I depend on the teacher's reaction. She made such a face. It is very important to me to see even the smallest smile on teacher's face, or to see that she looks kindly at me. However, if she looks at me as if I have said something wrong, I feel even less safe… But then, we prepared for the exam with her, so we knew her already…and we knew that she would not suddenly become our friend. (Everybody laughs) *She often covered her face or looked out of the window also during the classes.*

The effect of the necessity to compensate for the negative affect (or as test takers say *lack of relationship*) was lack of ideas, and a limitation of one's performance:

Elina: *if the teacher smiles or you have a closer contact, you feel freer, get more ideas and start speaking more…*

Int.: *If the person does not look at you, she does not see that you want to say something else*

Elina: *Yes, she stops your thoughts…*

Antra: *Yes, there were a couple of times when I wanted to say something, but that is quite personal, the more worried you are the more you are conscious of how people react.*

When asked whether they thought that their presentation would have been different with a different interlocutor

Antra says: *Yes, but I do not know whether it would have been any better though…*

Helen is of the same opinion, but again stresses the importance of knowing the person before the examination:

If, for example I had been with Charmen, or if I had been with somebody totally unknown and they had not smiled, I would not be thinking about what to say, but what I did wrong. If I look at the teacher, and she makes such eyes, that is it, nothing more to say, I'd rather keep quiet (Laughs).

Although all the test takers say that they were not worried by the interlocutor's manner because they knew the language and knew their teacher, their lack of excitement when talking about the examination was in such contrast with the test takers' comments in 1999, that I decided to compare the impact of the affective variables on their thoughts while taking the examination.

9.3 Comparison of the results of the two interviews

The questions I asked the test-takers in both interviews concerned anxiety: I told them that my research topic was test anxiety and I asked them to go through the tasks and comment on everything (e.g. test tasks, environment, interviewer) that caused any anxiety. When transcribing, I noticed, however, that their comments on the causes of anxiety were connected with the cognitive aspects: *I was worried because I could not remember what this word in English* (Dagmara), *I was worried about the picture and forgot to ask the questions* (Ieva 1999), *I forgot what I was told because I was*

worried (Helen 2000), *I do not know one word, I lose my thought* (Elina 2000). Thus, anxiety as well as other affective variables, were all connected with the test-takers' goals, their plans and the reaction to their own assessment of their performance or of the tasks. Therefore I decided to use Bachman and Palmer's (1996) framework of meta-cognitive strategies to see how the comments on anxiety fitted the framework of meta-cognition.

The other aim in using Bachman and Palmer's framework was to compare the test-takers comments in 1999 and 2000 from the point of view of meta-cognition: how the different affective reaction to the interviewers' personality (positive in 1999 and negative in 2000) had affected the use of meta-cognitive strategies (see Table 19).

Table 19 Comparison of the test-takers' comments in 1999 and 2000

Str.	Areas	1999 comments (Dagmara, Karina, Ieva)	2000 comments (Antra, Helena, Elina)
Goal Setting strategy	Identifying the test tasks	*If the person is normal, then you feel as if it is a normal conversation, so that you are not worried, can concentrate. (Dagmara) * It was not like an examination, it was like a game. (Dagmara) He is the sort of person with whom you want to talk to although I do not know him. He creates such a good atmosphere. (Dagmara)	I cannot say I was worried, as I knew the papers and we had practised with the last year's papers and I knew that the pictures were rather odd and can have all kinds of objects in them. (Antra)
	Choosing one or more tasks from a set of possible tasks	It is better now when we do not have to learn topics by heart. God forbid, forget a sentence and everything else will get mixed up. (Ieva)	
	Deciding whether or not to attempt to complete the tasks selected	I think that when you take an examination you concentrate all your energies, go and do it. (Ieva)	*Everything stopped in me, I thought, that is it, I do not need anything any more, she is not interested, she is going to assess my presentation from only how I entered… (Helen) * if she looks at me as if I said something wrong, I feel even less safe. (Elina)
Assess-ment strategy	Assessing the characteristics of the tasks	I actually liked the role-plays - you say something and get something back. I really liked the theme I had for questions about books, and the second task was good, the themes easy to discuss… (Dagmara) *Sometimes I think I have already talked too much, the questions were such that I could talk and talk. (Ieva)	*I took the paper, I looked at it, it was OK, but when I had to read the next task, my hands started trembling, when I start talking, I am fine, when I stop, my hands start trembling again (Helen) *I liked the other questions, because I remember that one day before the examination I was so relieved that it had all been discussed beforehand. (Helen)

	Assessing our own knowledge	*Sometimes I need a word, I remember it in German and forget it in English, and you start making funny constructions. It means we have not had enough practice in English (Ieva)	When I got my paper I was not very happy as there were some specific words that I did not know (toilet, sponge). I tried to think of something to say instead (Antra)
	Assessing the correctness or appropriate ness of the response	*That's right, when you talk, you do not think about anything else, who is going to listen, how they are going to assess you. May be at that moment you don't, but afterwards…Angry with yourself when you say something stupid and know that you could have put it better. *When I came in I thought, why aren't I worried? (Karina) *I was not afraid, but my tongue was all tied up (Karina)	*When I was preparing for the examination, we recorded it and there were many pauses, and when you listen to it, it is awful. (Elina) *In task 3 I asked whether I could get information instead of asking what information I could get. The teacher was puzzled; and I depend on the teacher's reaction, she made such a face… (Elina)
Planning strategy	Selecting elements from the areas of topical knowledge and language knowledge	* I was running all around the school in a very good mood, and all the thoughts were floating in my head, how can you catch them, but when you start speaking… (Ieva)	*I could not say much in task 2, I was afraid of becoming too philosophical, I do not know English very well. (Elina) *Everything depends on the teacher's facial expression and attitude, if the teacher smiles, or you have a closer contact, you feel freer, get more ideas and start speaking more. (Elina)
	Formulating one or more plans for implementing these elements	*I do not think that the exam tests your ability to concentrate; I do not think when I speak, I do not think at all, even when I write I do not think. (Dagmara) *I think exam does test your ability to let yourself go. (Ieva) *I was playing for the time to think of something. (Karina)	When I start thinking and miss a word I lose my thought, therefore I am just telling what I know. (Elina)
	Selecting a plan	Then suddenly I had an idea to ask him something personal, he got confused, but was not shocked. I also got confused; I thought he decided on his own to remind me to take the sandwiches. (Ieva)	I was thinking all the time, do not stop, go on talking in spite of the stress. (Elina)

In Table 19 we can see that the comments on anxiety elicited in both the interviews, contain comments on the test-takers' goals, their plans and their self-assessment. One can also detect the difference between the comments elicited in both the interviews (in 1999 and 2000) in all three areas, therefore I will discuss each separately.

9.3.1.1 Goal setting strategies

The goals of the test takers in the two interviews differed: the goal of the 1999 interviewees was to interact with the interlocutor using the test tasks as a means, while for the year 2000 interviewees it was the tasks that were an aim in themselves. As a

result, the year 2000 test-takers were quite satisfied that there was no warm-up before the Speaking test. All they wanted to do is come in, sit down and do the tasks:

Antra: *I don't know, I am used to her; with this teacher we do not have any relationship or any attitude, and in fact it did not disturb me. Of course it depends on the way the person has been trained, whether he has a personal relationship, but if you just come in and go out, with no smile, no anything.*

This is just the opposite to Ieva's enthusiasm about the interaction during the Speaking test in 1999. She considers an interactive kind of test to be much more appropriate as it does not depend so much on memory which can be affected by test-anxiety:

It is better now when we do not have to learn topics by heart, God forbid, will forget a sentence and everything else will get mixed up.

Deciding whether to do the tasks or not did not occur to the test takers in either case, although it had occurred during the year 2000 comment on Helen's previous experience (see Table 17). In addition, Antra (2000) said that she might not have said a word if this had not been a person whom they knew well. Which again can be opposed by Ieva's (2000) comment about *concentrating all your power* and doing your best without any reservation (see Table 19).

9.3.1.2 Assessment strategies

Assessment strategies could be discerned in the test-takers' comments on their attitudes to the tasks and their own performance:

Antra (2000): *I did not understand when she said 'describe your favourite spot', what did she mean, here, or in general? Otherwise there was nothing to puzzle me. And in fact we had quite a well-connected dialogue.*

The self-assessment of the correctness of one's own response, of the interviewees of year 1999 is more critical:

Ieva (1999): *The main thing, you feel that it is wrong, but do not know the right word... I was also worried about the picture and forgot to ask any questions, the formality at the beginning before the dialogue, everybody listening, you get afraid of your own voice, everybody is listening.*

Karina: *Angry with yourself when you say something stupid and know that you could have put it better.*

When assessing the tasks and their own performance the year 1999 interviewees seemed to be more dissatisfied with themselves. It seems that the test-takers' positive attitude to the interviewer in Year 1999 Speaking test led to their striving to perform as well as possible, but at the same time caused dissatisfaction with their own performance. In addition, the test takers assessed not only their performance but also their own reactions to test, for example, Karina became aware of the lack of anxiety, which seemed wrong:

 Karina: *When I entered, I thought, why aren't I worried*?

9.3.1.3 Planning

The year 1999 test-takers considered that a test tests your ability to let yourself go (this reminds of the 'flow' concept (see section 2.1); they constantly stressed their lack of planning, although they did not deny the re-examination of their performance afterwards:

Karina: *Yes, when you speak, you do not think about anything else.*

Dagmara: *No I do not think about anything else, even when I write, I do not think what I will have in the end.*

> Karina: *who is going to listen, how they are going to assess.*

However, Dagmara, when listening to the recording of her interview during the examination, explained how she used a special strategy, changing the topic, to cover a gap in her knowledge:

Dagmara: *Here I had to react quickly, before he could ask me anything else. If I felt I would not be able to answer a question, I said something quickly, so that he would not ask. Here I do not understand what he said, so I will ask again…*

The test takers of 1999 and 2000 have different views of the Speaking test: the year 1999 interviewees stress their spontaneity, and seem to use planning simply as a last resort, while the year 2000 (Elina) stress the constant need for control and planning of what to say. To illustrate, here is a short extract of the dialogue recorded in 1999:

> Ieva : *I think that an examination is when you concentrate all your power, go and do it, but not try once, and then again and then they send your best effort to the ministry, it is best to concentrate and do it.*
>
> Karina: *That's it.*
>
> Dagmara: *I do not think that exam tests your ability to concentrate, I think…*
>
> Ieva: *I think the exam tests your ability to let yourself go. I was in a constant lack of time …I lost any feeling of time.*
>
> Dagmara: *I wanted to speak more, speak and speak.*

Elina's (2000) comment is in striking contrast to the preceding dialogue:

Elina*: I could not say much in task 2, I was afraid of becoming too philosophical. I do not know English very well. When I started thinking and missed a word I lost track of my thought, therefore I was just said what I knew.*

9.4   *Findings of the interview study*

The main findings of Study 2 are the following:

- The interviewees do not consider the Speaking test to be more anxiety provoking than the other tests (this agrees with Madsen 1982).

- Nevertheless, the test-takers experienced anxiety before the Speaking test, during the Speaking test and after the Speaking test was over.

- The anxiety before the test was utilised as a motivator to prepare for the examination better. Their anxiety in the morning before the test was so strong that it was impossible to study, but as the interviewees entered the examination room and started performing, they were able to concentrate. They did not think of their performance while they were performing, but afterwards, they remembered the problems they had encountered and blamed themselves for not being able to remember what was said to them

- The attitude of the interviewer matters to the test-takers and affects their performance: if the interviewer is supportive and understanding, the test interview can develop into a genuine interaction with the interviewer, in this case the test-takers have lots of ideas to talk about, they initiate and ask questions that are expected of them (depending on their goals).

- During the interview test-takers think of ways of covering the gaps in their knowledge (planning strategy).

- If the interviewer is abrupt, or nervous him or herself, and cannot give his or her support, the test-taker starts thinking of what he or she has done wrong instead of the test itself and is not motivated to carry out the tasks fully (self-assessment strategy).

- If, however the interviewer is well known to the test-takers, his or her attitude is expected to be the same as previously: then the lack of support does not worry the test-takers as they use their confidence in themselves and in their knowledge about the interviewer to compensate for the lack of interviewer' support (knowledge is used to compensate the lack of affect).

I could summarise the comparison of the effect of affect (positive attitude to the interviewer in 1999 and negative in 2000) on the use of meta-cognitive strategies:

1. The goal setting strategies were different (the year 1999 interviewees wanted to interact, while the year 2000 interviewees wanted to do the tasks in the best possible way).

2. The planning strategies were different: year 1999 interviewees thought about how to develop the discourse, while the year 2000 interviewees thought of how not to say anything extra and how to follow the task demands as closely as possible (see Helen). This resulted in lack of initiation and improvisation.

3. The use of assessment strategies was different: after the interview the year 1999 interviewees were critical of themselves and thought they could have done better, while the Year 2000 interviewees, although noticed problems in their performance, congratulated themselves on the fact that they had managed in spite of the interviewer's attitude.

The comparison of the interaction between the comments of the year 1999 and year 2000 interviewees suggest that the test-takers' affective reaction to their interviewers during the examination affect their choice of meta-cognitive strategies (goal-setting, planning and assessment). This supports Bachman and Palmer's (1996) model of interaction between affect and meta-cognitive strategies.

Chapter 10   Study 3 - Questionnaires

This chapter includes the description of the two studies (in years 1999 and 2000) where I used questionnaire as a research method. The questionnaires had one part in common, that of a scale measuring the level of anxiety during the Year 12 examination, but the second part differed:

1. in year 1999 I used Spielberger et al's (1978) Test Anxiety Inventory (TAI) to investigate the relationship between general test anxiety and my measurement of foreign language test anxiety,

2. but in year 2000 I used Purpura's (1999) questionnaire of meta-cognitive strategies to investigate the interaction between my measurement of foreign language test-anxiety and meta-cognitive strategies.

10.1 Questionnaire study in year 1999

This section contains the description, the analyses of the results and the findings of a questionnaire study carried out in 1999.

10.1.1 Aim of the study

The aim of the year 1999 study was to assess the level of anxiety, measure its impact on the Year 12 exam (see examination materials in Appendix 1) and find out the reasons for test anxiety.

The TAI (Spielberger et al 1978) was administered together with the anxiety measurement questionnaire (see the questionnaire in Appendix 5) to find out the impact of the test anxiety variables (Emotionality and Cognitive worry) on different language skills and groups of different language proficiency levels.

The TAI results were compared with the test-takers' performance on all skills, but the self-reported levels of anxiety were registered immediately after the written part of the

examination and therefore did not comprise information on the anxiety level of the Speaking test, which took place later and was explored using interview as a research method (see section 9.1).

10.1.2 The test-takers

All the test-takers were secondary school graduates. Out of more that 4000 questionnaires received, 250 responses were selected using a hierarchical multi-stage stratification method. The sample was made up from 8 different regions of the country, half of them from the capital Riga (half of the population lives in Riga), with a wide range of proficiency levels: there were 40 test-takers for each level of proficiency group (see section 10.1.5).

After being assessed and evaluated at the same time as the other school leavers, the selected test-takers' questionnaires were separated for the analysis of the level of anxiety and its impact on test-performance.

10.1.3 Instruments

The questionnaire (see Appendix 5) consisted of two parts: the first was a scale for measuring the level of anxiety during each task of the written part of the examination. The questionnaire was developed in Latvian to encourage the test-takers of lower levels of English language proficiency to respond. The questionnaire started with a request to circle the number (1 to 4) in order to reveal the test-takers' level of anxiety (from *did not cause anxiety* to *caused strong anxiety*) for each task. For each task the test-takers were also requested to give the reason(s) for their anxiety. They were informed that their responses would in no way influence the evaluation of their work, but would help us to develop a better test next year.

The second part of the questionnaire was Spielberger et al's (1978) TAI, which differentiates between the Worry and Emotionality components of test-anxiety (see

section 4.2.). The Inventory described the physiological manifestations experienced when taking examinations in general and the thoughts that might have occurred to the test-takers during the examination. The test-takers had to state how often (from 1: *nearly never* to 4: *nearly always*) they experienced these thoughts (e.g. *thoughts about my possible mark interfere with my work*) or feelings *(e.g. I feel jittery during the examination*).

Spielberger et al. developed the Test Anxiety Inventory as an *objective, relatively brief self report scale that would correlate highly with other widely used measures of test anxiety* (Spielberger et al 1978, p.180), and which could be used for measuring the Worry and Emotionality components of anxiety. Worry was described as a primarily cognitive concern about the consequences of failure, and Emotionality was defined as autonomic reactions evoked by evaluative stress. Liebert and Moris (1967) provided evidence that worry was associated with performance decrements on cognitive tasks, whereas emotionality was unrelated to task performance (Spielberger, 1978, p. 173).

The reasons for using the Inventory in this study were on the one hand to let the test-taking population in Latvia know what is understood by test-anxiety and to introduce the idea that it is a general phenomenon experienced by most test-takers, not just by students of English in Latvia. On the other hand it was used to find out if test anxiety can be facilitating as well as debilitating and to discover how it influences performance in different language skills.

The Year 12 examination was used to measure the test-takers' language proficiency (see Chapter 7) in spite of the security problems that occurred before the test in Riga and Daugavpils. As the statistical analyses of the test results were satisfactory (see section 7.6.) I decided to use the exam results for this study.

10.1.4 The administration of the questionnaire

The staff of the Curriculum and Examination Centre (CEC) photocopied the forms and packed them together with the examination papers. The test-takers filled them in immediately after the exam; the exam invigilators collected them and sent them back to the CEC together with examination materials.

10.1.5 Results of the study

The mean anxiety level of each task was calculated in order to compare the anxiety level each task evoked. Their means were compared to the difficulty level of each task to see if there was any relationship between them.

The test-takers' levels of anxiety were correlated with their performance levels on the relevant task and with their overall language proficiency in order to assess the impact of anxiety on performance and to see whether it was general language ability that was affected by anxiety or whether it was skill specific ability.

Using the mean and the standard deviation the whole sample was split into four groups according to their level of proficiency. The group with the highest level of proficiency (P4) comprised all the students whose scores were more than one standard deviation above the mean, the next group (P3) comprised the test-takers within plus one standard deviation of the mean, the following group (P2) comprised all the test-takers a standard deviation of minus one of the mean and the last group (P1), those test-takers who scored more than 1 standard deviation below the mean.

 For each group correlations were calculated to assess the interaction between anxiety and performance in different proficiency groups. In the description of the study, however, the results of only two groups have been included: the strongest (P4), whose language proficiency was more than 1 standard deviation above the mean (45 test-takers) and the weakest (P1), whose language proficiency was more than one standard

deviation below the mean (43 test-takers). It was done to highlight the role of proficiency in its interaction with anxiety.

Using the Test Anxiety Inventory results the mean for each question was calculated to assess the level of the Emotionality and Worry factors. The test-takers' performance on each task was correlated with the emotionality and worry factors to assess the facilitating and debilitating features of test-anxiety in each language skill.

The reasons for anxiety were coded and their frequency was calculated. Then similar comments were grouped and their frequency calculated.

The results of the two instruments were compared to assess the interaction between the test-takers' reactions to examinations in general and their reaction to the English language examination in particular.

10.1.5.1 Anxiety level

The level of anxiety of the whole sample ranged from 39% (Writing task 1) to 66 % (Listening task 3) (see Table 20). The anxiety level changes follow the same pattern
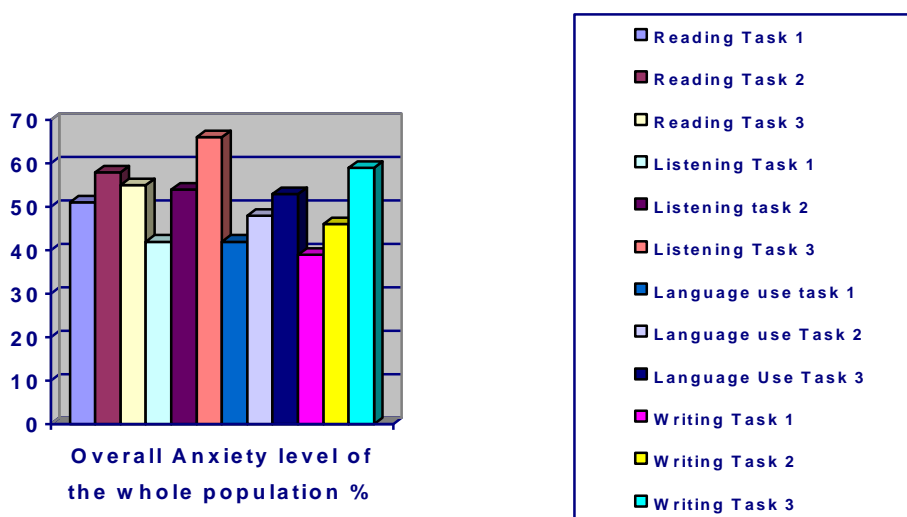
Figure 16 The overall anxiety level of the whole population

as difficulty level (see Figure 16): the first task of each test is the easiest and the least anxiety inducing, the second is more difficult and the third is not only the most difficult but also the most anxiety inducing in that test. The overall level of anxiety is 51%, which is lower than the debilitating threshold (60%) referred to by Madsen (1982) (see section 4.3).

The only exception to the steady progression of difficulty level, as we have already seen in section 4.1.6, is the Reading test where the second task was more difficult than expected. The most anxiety-inducing task, however, is not the most difficult reading task, but the third listening task (66%), which is above the debilitating threshold (Madsen 1982). Here one cannot blame the difficulty level of the task, as it appears to be quite easy (mean 62%).

Table 20 Anxiety level according to proficiency level

Task	Mean % of the task	Overall Anxiety level %	Anxiety level in the highest proficiency group (P4)	Anxiety level in the lowest proficiency group (P1)
Reading	**52**	**55**	**37**	**61**
Matching	69	51	34	58
Gap-filling	36	58	43	63
Multiple-choice	52	55	35	61
Listening	**69**	**54**	**43**	**59**
Gap-filling	80	42	37	50
Multiple-choice	67	54	41	58
True/false	62	66	52	70
Language Use	**66**	**48**	**33**	**55**
Multiple choice	78	42	31	51
Editing	64	48	34	54
Gap-filling	57	53	34	61
Writing	**69**	**41**	**38**	**55**
Postcard	77	39	30	43
Letter	71	46	37	52
Report	60	59	46	70
Total	64	51	38	58
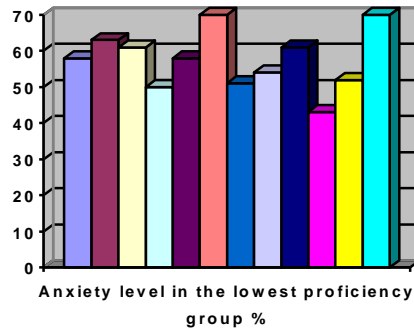
Figure 17 Anxiety level in the lowest proficiency group in Reading, Listening, Language Use and Writing Tests

When the sample was split to assess the level of anxiety in different language proficiency groups the pattern in general was the same: both groups reacted to the difficulty level of the tasks. The two charts (Figure 17 and Figure 18) show the fluctuation of anxiety level during the test in the highest and lowest proficiency groups and the bars represent the tasks in the order given in Table 20 (3 tasks per test). The highest peaks in both groups are for the third Listening task, the most difficult of the three tasks, the second highest is for the Writing task 3. The highest proficiency group (Figure 18), however, does not seem to be affected by the Language Use tasks although these become progressively more difficult, all three bars
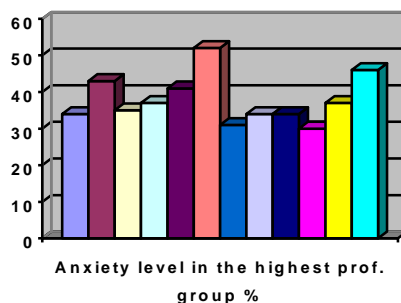
Figure 18 Anxiety level in the highest proficiency group in Reading, Listening, Language Use, Writing tests

representing Language Use appear to be quite similar.

In the highest proficiency group (P4) the level of anxiety in most tasks is around 30% with three peaks distributed regularly throughout the whole examination. This suggests that although the test-takers were anxious, they also managed to recover during other tasks. The lowest proficiency group (P1) was constantly anxious: their anxiety level is close to 60% with only one respite: the postcard task in the Writing test (see Figure 17). Thus the results suggest that the test-takers of the lowest proficiency group were carrying a double load: doing the tasks and constantly worrying. This may have tired the least proficient group more than the most proficient group (see Selye' s GAS theory in section 4.1.4). In addition, for this group, 5 tasks out 12 were above the debilitating threshold (see section 4.3.).

10.1.5.2 Correlations between level of anxiety and level of performance

Mathews et al (2000), following Yerkes and Dodson law say that the *relationship between arousal level and performance may be expressed as an inverted U curve. The cortex functions most efficiently at moderate levels of arousal, when the person is alert and wakeful, but not highly excited or agitated* (Mathews et al 2000, p. 165).

Dyer (1997), on the other hand says that correlation coefficients can be used to reliably measure the relationship between two underlying variables only if we have a reasonable belief that they are related in a linear fashion; that is, *the increases and decreases to the value of one variable are accompanied by corresponding and parallel increases or decreases to the value of the other throughout the whole range of possible values of the variable* (Dyer 1997, p. 293).

If we take into account what Mathews et al (2000) and Dyer (1997) say, it is evident that correlation coefficients are inappropriate as a method of measurement of the relationship between anxiety and language performance. Nevertheless, much of the research discussed in Chapter 4 (on anxiety in psychology as well as linguistics) use

correlation coefficients to measure the relationship. As Dyer (1997) does not suggest a different method for non-linear relationship measurement, I will follow the example of other researchers and also use correlation coefficients.

Another reason why I decided to use correlation coefficients was the fact that many of the correlations between performance and self-reported anxiety level were significant although not very strong. As I had more than 200 test-takers' replies, the critical value of the Pearson product-moment correlation coefficient at the probability 0.01 was .182 (two-tailed, as I was interested in both positive and negative correlations). The correlations below .182 are marked as non significant (n.s.) (see Table 21).

All the correlations between the self-reported level of anxiety and the test-takers' level of performance are negative (see Figure 19). The impact of anxiety is highest for Reading test (-.438), which was also the most worrying **test** according to the test-takers' evaluation (see Table 20).

In the most worrying **task**, the third listening task (mean 66%), on the other hand, the worrying does not seem to have had much impact, although the correlation is negative (-.228) for the whole population, and very small but still positive for the group of the lowest proficiency group. This suggests that we should be calling it 'emotionality', instead of 'worrying'. According to cognitive interference theory (see section 4.2) it is the cognitive variable of test anxiety that interferes with the performance level, while emotionality does not affect it.

Another interesting feature is the fact that the impact is the highest in the first task (see Table 21) in all tests except writing, where the performance on the third task has been most influenced by anxiety. This may suggest that the cognitive variable is more important than emotionality in the third task.

Table 21 The correlation between anxiety level and proficiency level in the two proficiency groups

Task	Overall Correl. with prof. in the relevant skill	Correlation with overall prof. level	Correlation with prof.in same task for the strongest group (P4)	Correlation with prof.level for the weakest group (P1)
Reading				
Matching	-.438	-.370	n.s.	n.s.
Gap-filling	-.374	-.324	n.s.	n.s.
Multiple-choice	-.423	-.353	-.223	-.195
Listening				
Gap-filling	-.293	-.208	n.s.	-.572
Multiple-choice	-.242	-.209	-.271	n.s.
True/false	-.228	n.s.	n.s.	.n.s.
Language Use				
Multiple choice	-.280	-.336	.313	-.277
Editing	-.310	-.278	.207	n.s.
Gap-filling	-.385	-.348	n.s.	n.s.
Writing				
Postcard	-.207	-.243	n.s.	n.s.
Letter	-.186	-.224	n.s.	-.566
Report	-.232	-.322	n.s.	-.326
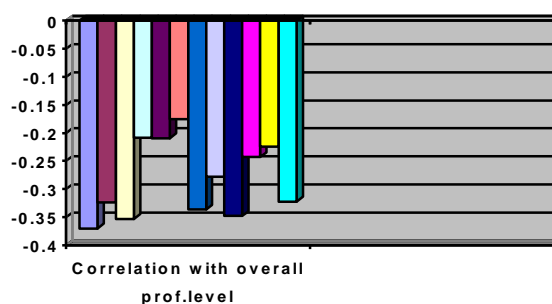
Figure 19 Correlation between the anxiety level and proficiency level for the whole sample

The correlations between the level of anxiety and performance level are negative in the whole sample (see Figure 19) . If we compare the chart in Figure 19 to the charts depicting the difficulty level and anxiety level changes which were lowest for task 1, rather higher for Task 2 and highest for the Task 3, depending on the difficulty level of the task, here in Reading and Listening tests, we can sometimes see the opposite: the anxiety in the first and easiest tasks in all the tests affect the performance more than the other two.

If the sample is split according to proficiency level, the correlations become more varied. In both the groups there are some positive correlations between level of anxiety and proficiency level, the highest (.313). is Language Use Task 1 (see Table

21) in the highest proficiency group. Both the highest negative correlations for the lowest proficiency group are Listening task 1 (-.572) and Writing task 2 (-.566).

10.1.5.3   Causes of anxiety

The test-takers have given different reasons for their anxiety. Their main reasons are: test-situation (lack of time or space, nervousness about the result, unknown task types), language specific (unknown words, nervousness about grammar and vocabulary), administrative problems (could not hear, bad recording, acoustics of the room) and personal problems (a broken finger, someone died).

The reasons provided were classified into 24 groups, coded and data processed. The results for each task can be seen in Table 22. Next to each reason we can see how many times each reason was mentioned in each task (in Reading (R), Listening (L), Language Use (LU) and Writing (W) tests).

Table 22 The reasons for anxiety and the tasks in which they occurred

		R1	R2	R3	L1	L2	L3	LU1	LU2	LU3	W1	W2	W3	Total
1.	Beginning of the test, the first task	34	9	8	0	0	0	0	0	0	1	0	0	51
2.	Could not concentrate	0	0	0	0	0	0	0	0	1	0	1	0	2
3.	Too many unknown words in the texts or other vocabulary problems	11	30	15	2	1	1	2	3	10	11	6	6	97
4.	Did not know what to do, did not understand the task	6	7	2	3	3	3	3	3	2	1	2	4	38
5.	Difficult text	5	8	5	1	3	2	1	1	1	0	0	0	26
6.	Anxious about the results	8	11	10	12	11	12	21	15	15	13	14	13	154
7.	Unclear speech in listening tasks	1	0	0	34	42	46	1	0	0	0	0	0	123
8.	Background noise in the listening test	0	0	0	1	13	10	0	0	0	0	1	0	24
9.	Speech in the listening texts too fast	0	0	0	13	18	46	0	0	0	0	0	1	77
10.	Acoustics of the room or other admin. problems	1	1	1	7	2	5	1	1	1	1	1	1	23
11.	Noises on the street, coughing in the room	1	1	0	0	0	0	0	0	0	0	1	1	4
12.	Unexpected task type	1	0	0	0	0	0	0	3	1	3	0	1	9
13.	A difficult theme	0	0	1	1	0	4	0	3	12	3	18	56	97
14.	Too long a composition	0	1	0	0	0	0	0	0	0	0	1	1	3
15.	Too short composition	0	0	0	0	0	0	0	0	0	2	1	5	8
16.	Lack of space in writing task	0	0	0	0	1	0	0	0	0	0	10	2	12
17.	Lack of time	24	24	34	2	0	1	3	8	7	4	4	8	118
18.	Broken finger, difficult to write	1	1	1	1	1	1	1	1	1	1	1	1	13

		R1	R2	R3	L1	L2	L3	LU1	LU2	LU3	W1	W2	W3	Total
19.	No reason	7	6	7	7	6	4	5	5	5	6	6	5	69
20.	Tired at the end of the test	0	0	0	0	0	0	0	0	0	3	3	3	9
21.	Difficult task	8	19	7	1	1	3	5	7	10	2	4	6	73
22.	Positive agitation	1	1	1	1	1	0	0	0	0	1	1	2	9
23.	Listening text and task in different order	0	0	0	0	1	1	0	0	0	0	0	0	2
24.	Total													1021

The reason most often mentioned was anxiety about the result of the test. This is mentioned 154 times (some test-takers gave the same reason for several or even all tasks). This unspecified type of anxiety created by the test situation can be seen throughout the examination, but many students mention the beginning of the examination as a cause of anxiety. The same happens at the end of the examination: candidates become tired and this becomes another reason for anxiety. Selye discusses this phenomenon in his GAS theory (1956), which was discussed in section 4.1.4. Here, however, one can see in practice how "the sum of all non-specific systemic reactions of the body to long-continued exposure" can be divided into three stages: alarm reaction (beginning of the test), resistance (worry without any reason, cannot concentrate) and exhaustion (tiredness at the end of the test).

Although all the feedback on the break before the Writing test, which divides the test into two approximately equal parts of 90 minutes each, has been positive, one cannot be sure that after the break when the students have to start anew this does not add to the strain. On the other hand, the fact that there are only nine cases of tiredness affecting task performance in 258 questionnaires is encouraging. Another encouraging feature is the fact that there are nine cases where the test-takers have recognised that test anxiety can be a positive reaction to a test situation.

Table 23 General state anxiety

Anxious about the results	154
Beginning of the test, the first task	51
Tired at the end of the test	9
Positive agitation	9
Could not concentrate	4

No reason	69
Total	296

The total number of cases when anxiety causes could be classified as state anxiety caused by the evaluative situation is 296 out of the whole (1021 all in all), that is 29% (see Table 23).

In this study test-takers have also mentioned several causes that can be classified as foreign language anxiety (see section 4.3). Some reasons (such as unclear speech) could be classified both as an administrative problem (if the tape recorder was bad) or language anxiety if the test-taker is not used to natural speech; others, like, *I am never good at writing*, or *a difficult text* or *many unknown words*, are definitely foreign language anxiety. All in all foreign language anxiety could underlie 457 reasons out of 1021, that is 45% (see Table 24).

Table 24 Foreign language anxiety

Unclear speech in listening tasks	123
Too many unknown words in the texts	97
Speech in the listening texts too fast	77
Background noise in the listening test	24
Difficult text	26
A difficult theme in writing	97
I am never good at writing	2
Too short composition	8
Too long a composition	3
Total	457

A different kind of reason for anxiety is test design problems. This was also uncovered by this questionnaire. Lack of time is a popular complaint in language tests and test designers often use shortage of time on purpose as a discriminator. In this case, however, the 82 complaints of lack of time out of total of 118 in foreign language anxiety group altogether were related to Reading. The test team decided that the complaint was valid and that the time for next year's reading test would be increased in spite of the fact that the Year 12 test was supposed to be a power test.

There were two cases where a person considered that the questions in the listening tasks were presented in a different order from the relevant material in the text, but this was considered to be untrue. There were also 38 cases where the test-takers had problems with understanding the tasks. As these did not concern one task, but referred to different tasks it was perhaps unfamiliarity with the specific task type.

All in all test design caused 170 complaints or 17% of all complaints (see Table 25). These are what Bachman (1990) calls 'systematic errors', as they affect all the test-takers equally and can decrease the validity of the measurement.

Test administration problems, however, lead to unsystematic differences, which decrease the reliability. There were 48 such cases according to the test-taker feedback (4%). Unfortunately we cannot know whether these were the only variations in administration; none of the test-takers mentioned any problems with test security, although that was a major concern for all the test administrators.

There were also personal problems, such as a broken finger that caused difficulty in writing and a funeral of a classmate on the previous day. These problems decreased the reliability of the test and may have affected the test scores.

Table 25 Examination design problems

Did not know what to do, did not understand the task	38
Listening text and task in different order	2
Lack of time	118
Lack of space in writing task	12
Total	170

Table 26 Examination administration or training problems

Acoustics of the room	23
Unexpected task type	9
Noises on the street, coughing in the room	4
The invigilator was standing at my desk all the time	12
Total	48

All the causes of anxiety could be depicted in the form of a pie chart (see Figure 20)
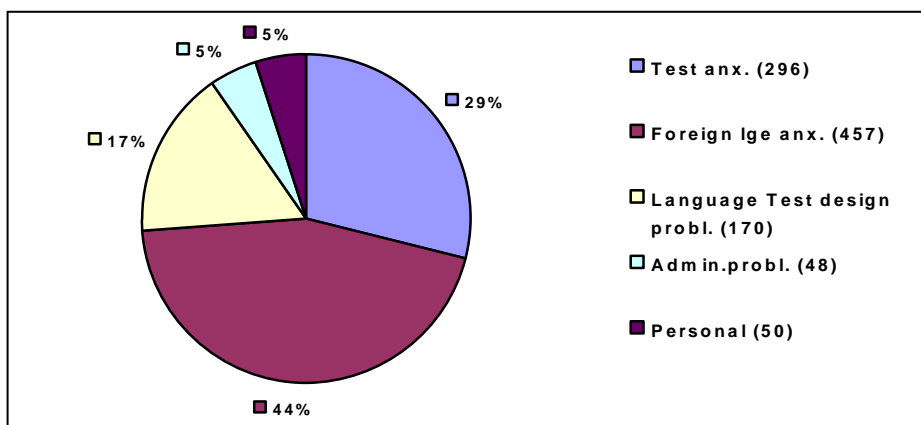
Figure 20 Causes of anxiety during the examination

that clearly demonstrates the overwhelming influence of foreign language anxiety, which is even higher than test anxiety. The two smallest segments are test administration and personal problems during the test.

10.1.5.4   Test Anxiety Inventory analyses

The aim of administering the Test Anxiety Inventory (Spielberger et al 1978) (see the questionnaire in Appendix 4) was to investigate the experiences the test-takers had gone through during the test and to see how these agreed with the responses given in the first part of the questionnaire. The third aim was to see if it is possible to single out the Worry and Emotionality components in test anxiety and compare their impact on language performance.

The test-takers were asked how often they felt as described, from *never* to *very often*. Their answers were converted into percentages: *never* was 0% and *always* was 100%. The results as seen in Table 27 suggest that the frequency of the Emotionality factor is 53%, and the Worry factor is 43%. Spielberger et al's research suggested that it was the Worry component that affected performance: thinking about the grade, getting

through the test, doing so poorly, so that this interferes with concentration. In the case of the Year 12 examination the answer is not so straightforward.

The difference between the strongest (P4) and weakest groups (P1) can be seen in the first question: that of confidence about the examination, where there is the only positive correlation with the test performance (see Table 27). However, if we compare frequencies, we can see that many of the same students who chose 'confidence' (Question 1) must have also chosen the most frequent answers like *Worry before important examinations* (Question 15)*, Uneasy before getting the test paper back* (Question 10)*.

Thus confidence can be combined with what one student called a *hunter's excitement* The difference between the level of confidence in the highest and lowest proficiency groups is 23%, and the differences between the two groups *feeling upset* (Question 2) and *feeling jittery during the examinations* (Question 8) is also 22% (see Table 27). *Anxious during the test even when well prepared* (Question 9)*, Uneasy when getting the paper back* (Question 10) and *Wish exams did not bother so much* (Question 11) and *thinking about failing* (Question 16) differentiate the groups by 20% or more. But such features as *so tense stomach gets upset during exams* (Question 12), *defeat self on tests* (Question 13), *feel panicky during tests* (Question 14), worry before *important tests* (Question 15)*, heart beating fast during tests* (Question 17) and *nervous, forget facts during exams* (Question 19) do not differentiate between the groups by more than 10%.

In both the groups the Worry (W) component forms a smaller part of anxiety than Emotionality (E). The difference between the means of W and E in the highest proficiency group (P4) is 4%, but in the lowest proficiency group (P1) it is 6%. If we compare the two groups the difference in the W component is 11%, and in Emotionality it is 13% (see Table 27).

Table 27 The frequency of experiencing test anxiety in different groups of proficiency

Test anxiety inventory	Worry or emotion factor	Frequency of the exp. in the whole sample (%)	P4 gr. (%)	P1 gr. (%)
1. Confident and relaxed during the examination		65	81	58
2. Uneasy, upset feeling during finals	E	57	45	67
3. Thinking about the grade interferes with work	W	50	47	58
4. Freeze up on finals	E	47	42	52
5. Thinking about whether I will get through the school	W	45	38	47
6. Confused when working on tests	E	49	45	55
7. Thoughts of doing poorly interfere with concentration	W	52	45	61
8. Feel jittery during tests	E	55	42	64
9. Anxious during tests even when well prepared	E	70	60	80
10. Uneasy before getting test paper back	E	53	44	61
11. Wish exams did not bother me so much	E	68	55	76
12. So tense stomach gets upset during tests	E	32	30	33
13. Defeat self on tests	W	68	36	38
14. Fell panicky during tests	E	33	32	36
15. Worry before important tests	E	69	61	71
16. Thinking of failing during tests	W	47	38	60
17. Heart beating fast during tests	E	51	44	57
18. Worrying after exam is over	E	43	35	48
19. Nervous, forget facts during exams	W	46	42	48
Mean frequency level		**52**	**44**	**56**
Number of test-takers		**250**	**45**	**43**
Worry			**41**	**52**
Emotionality			**45**	**58**

10.1.5.4.1 Correlations between Test Anxiety Inventory answers and language performance

All the correlations between the Spielberger et al's Test Anxiety Inventory measurement of the frequency with which test-takers experience test anxiety (either Worry or Emotionality) and the level of language performance for the whole sample are negative, although some are stronger than others (see Table 28). The positive correlations are marked bold.

Question 16, *thinking about failing during the examination*, is mostly negative (, -.344), and *uneasy, upset feeling during finals* (Question 2, -.295), but again, as is the case with correlation coefficients, it is difficult to be certain that it is the *cause* of failure, rather than the *result*.

Another damaging feeling is being anxious in spite of being prepared (Question 9). This had a frequency of 70%. This means that a large part of the sample chose it and marked it as a frequently experienced feeling. The wish to be less worried by exams also scored 66% (Question 11) and correlates negatively with performance (-.229).

Feeling tense (Question 12), *panicky* (Question 14), *defeating self* (Question 13), and *forgetting facts* (Question 19), do not significantly correlate with performance in the Year 12 examination. *Forgetting facts* is the only Worry component that does not produce a significant negative correlation with performance, but then in a language examination one would not be expected to demonstrate a knowledge of facts.

Table 28 Correlations between level of test anxiety and linguistic performance in different proficiency groups

Test anxiety inventory	W/ E.	Correlation with overall performance	highest correlations (whole sample)	highest correlations for Group P4	highest correlations for Group P1
1.Confident and relaxed during the examination		.318	**.348 R**	**.200R**	n.s.
2.Uneasy, upset feeling during finals	E	-.295	-.317 R	**.379 LU1**	-.351 L1
3.Thinking about the grade interferes with work	W	-.223	-.278 L1	**.213 L1**	-.467 L1
4.Freeze up on finals	E	-.237	-.264 L1	-.439 L1 **.365 W2**	-.488 L
5.Thinking about whether I will get through the school	W	-.143	-.205 R3	n.s.	-.279 L
6.Confused when working on tests	E	-.211	-.264 R	-.359 R1	-.305 R
7.Thoughts of doing poorly interfere with concentration	W	-.241	-.266 L	-.337 R2	-.313 W2
8.Feel jittery during tests	E	-.237	-.308 R1	-.243 R2	-.284 L1
9.Anxious during tests even when well prepared	E	-.246	-.222 R	**.246 L**	-.245 L3

Test anxiety inventory	W/E.	Correlation with overall performance	highest correlations (whole sample)	highest correlations for Group P4	highest correlations for Group P1
10.Uneasy before getting test paper back	E	-.219	-.226 Sp	-.208 Sp **.208 R1**	-.276 W2
11.Wish exams did not bother me so much	E	-.229	-.250 Sp	**.230 LU 1**	**.468 W** -.344 Sp
12.So tense stomach gets upset during tests	E	n.s.	n.s.	-.398 LU1	-.414 L
13.Defeat self on tests	W	n.s.	n.s.	-.320 R1	-.393 L
14.Feel panicky during tests	E	n.s.	n.s.	**.288 L**	-.331 W2
15.Worry before important tests	E	n.s.	-.205 L1	**.310R2**	-.397 W
16.Thinking of failing during tests	W	-.344	-.325 R	-.237 W2 **.232 R2**	-.383 W
17.Heart beating fast during tests	E	-.215	-.231 R	-.269W1 **.259 R2**	-.407 L
18.Worrying after exam is over	E	-.200	-.256 L1	-.443L1	-.526 L1
19.Nervous, forget facts during exams	W	n.s.	n.s.	-.271 R1 **.216R2**	-.365 R1

The correlations between the frequency of test anxiety experiences and language performance in the whole sample is negative, but again as the sample is split according to proficiency level there are several cases of positive correlations (just like with the first part of the questionnaire that was measuring the foreign language anxiety), especially in the highest proficiency group. These are marked bold (see Table 28). The differentiation between positive and negative correlations in different proficiency groups suggests the non-linear relationship between tets-anxiety and performance.

There are 11 cases for the objective tasks (receptive skills) when test anxiety actually helped the test-takers; 6 out of these reports relate to the reading tasks. For the weakest group, listening is the most negatively affected (11 cases). In this case it is the state of anxiety in general that affects the language skill, as the students are not talking about any particular language task. This shows how deeply language is connected with study skills in general: Spielberger et al (1978) found that Test Anxiety Inventory had a negative correlation with a study skills test of -.56. The high

and overwhelmingly negative correlations with general test anxiety in the weaker group look alarming and lead us to wonder whether what is being measured here is language skill at all.

10.1.5.4.2 The impact of Worry and Emotionality factors on language skills in the whole sample

If we compare the impact of anxiety levels marked by test-takers for each skill, there is little difference between Worry and Emotionality factors. However, if we look at the highest correlations, it is mostly the receptive skills, reading and listening that have stronger correlations. The only exceptions are the two Emotionality features where the correlations are highest for speaking *Uneasy before getting test paper back*, (-.23) and *Wish exams did not bother me so much* (-.25).

Although all the correlations are negative (except Question 1, which measures the level of confidence), the Emotionality factor for Question 9 *Anxious during tests, even when well-prepared*, productive skills' correlations are lower than receptive. The Worry Question 16 *Thoughts about failure during tests* impact is equally negative for all skills (apart from one Listening task and one Language Use task) (see Table 29).

Table 29 Correlations between Worry and Emotionality factors and test performance for the whole sample

Task	Correl.with *Confident and relaxed during exams*	Correl.with *Anxious in spite of preparedness*	Correl.with *Thoughts about failure during exam interfere*
Reading	**.348**	**-.259**	**-.325**
Matching	.268	-.308	-.267
Gap-filling	.327	-.202	-.310
Multiple-choice	.294	-.187	-.238
Listening	**.270**	**-.248**	**-.287**
Gap-filling	n.s.	-.227	-.239
Multiple-choice	.234	-.197	-.175
True/false	.272	-.186	-.303
Language Use	n.s.	n.s.	**-.247**
Multiple choice	.199	n.s.	-.252
Editing	n.s.	n.s.	-.186
Gap-filling	n.s.	n.s.	-.244
Writing	n.s.	n.s.	**-.240**
Postcard	n.s.	n.s.	-.247
Letter	n.s.	n.s.	-.242

Task	Correl.with *Confident and relaxed during exams*	Correl.with *Anxious in spite of preparedness*	Correl.with *Thoughts about failure during exam interfere*
Report	n.s.	n.s.	-.240

Table 29 suggests that it is Reading and Listening skills that benefit from test-takers' confidence (Question 1) and suffer from being anxious in spite of preparedness (Question 10), while thinking about failure (Question 16), harms the test-takers throughout the examination. On the other hand, it could as well be the self-assessment strategy that would be telling the test-taker that he or she is not doing well, that agrees with the test-takers' performance on the test.

10.1.5.5 Correlations between levels of anxiety and the Test Anxiety Inventory

The correlations between the measurements of anxiety in each language task during the examination and feelings about examinations in general are positive and in most cases significant (apart from Question 12: *so tense the stomach gets upset* and *forget facts during the examination*) (see Table 30). This allows us to conclude that the two parts of the questionnaire were at least in part measuring test anxiety. The only negative correlation is between the level of anxiety during the language exam and the absence of anxiety (Question 1) in the Test Anxiety Inventory. The highest correlation can be found in the second reading task (Gap-filling), which was too difficult for most of the students, so that the test anxiety component must have been stronger than language anxiety. The next statement: *uneasy and upset during the finals* (Question 10) also shows how much of the anxiety was caused by the general state anxiety in each skill. Again the strongest correlation is in the Reading test, which seems to be the most related to study skills in general.

Table 30 Correlations between Language test anxiety and the Test Anxiety Inventory

Questions	R1	R2	R3	L1	L2	L3	LU1	LU2	LU3	W1	W2	W3

Questions	R1	R2	R3	L1	L2	L3	LU1	LU2	LU3	W1	W2	W3
1.Confident and relaxed during examinations	-.390	-.451	-.360	-.263	-.368	-.378	-.389	-.362	-.350	-.319	-.369	-.378
2.Uneasy, upset feeling during finals	.414	.413	.433	.309	.400	.378	.301	.309	.257	.268	.387	.274
3.Thinking about the grade interferes with work	.271	.248	n.s.	.235	.229	n.s.	.189	n.s.	n.s.	.222	n.s.	.168
4.Freeze up on finals	.287	.323	.334	.275	.243	.173	.217	.262	.221	.187	.250	.271
5.Thinking about whether I will get through the school	.245	.324	.305	.138	.239	.224	.206	.293	.263	.254	.284	.285
6.Confused when working on tests	.164	.292	.254	n.s.	.202	.258	.183	.212	.183	n.s.	n.s.	n.s.
7.Thoughts of doing poorly interfere with concentration	.281	.265	.286	.269	.242	.242	.217	.219	n.s.	.288	.249	.213
8.Feel jittery during tests	.260	.318	.299	.238	.283	.334	.255	.276	.194	.216	.296	.269
9.Anxious during tests even when well prepared	.183	.195	.183	.176	.198	n.s.	n.s.	.205	.234	n.s.	n.s.	.186
10.Uneasy before getting test paper back	.254	.251	.272	.259	.239	n.s.	.197	.235	n.s.	.286	.307	.285
11.Wish exams did not bother me so much	.219	.342	.268	.199	.299	.309	.228	.271	.268	.252	.311	.335
15.Worry before important tests	.312	.351	.249	.185	.221	.194	.196	.235	n.s.	.222	.281	.241
16.Thinking of failing during tests	.238	.252	.272	n.s.	n.s.	.208	n.s.	.230	n.s.	.193	.225	.202
17.Heart beating fast during tests	.238	.253	.307	.194	.222	.170	.232	.185	n.s.	.191	n.s.	n.s.

Although there are 5 questions that do not have significant correlation with foreign language anxiety: (Questions 12, 13, 14, 18 and 19), nevertheless there is enough evidence that there is a relationship between the two anxieties. On the other hand, it is clear also that the relationship between the two anxieties does not depend whether Spielberger et al consider the question to be measuring Cognitive or Emotionality factors.

10.1.6 Findings of year 1999 study

The main findings of the year 1999 study were the following:

- Foreign language anxiety during the test is measurable and its level changes during the tests. It differs between different test tasks and different language skills. It differs also in groups of different language proficiency. The group with the highest level of proficiency has the lowest level of anxiety, and the group with the lowest level of anxiety has the highest proficiency level.

- The correlations between foreign language test anxiety and language performance are significant, although low and negative for both the highest as well as the lowest proficiency groups of test-takers.

- Test takers are aware of the causes of anxiety. They mention causes connected with their general level of anxiety (anxious about the results is the most often mentioned, 154 cases), anxiety about language use (unclear speech in the listening test 123 cases), examination design problems (lack of time 118 cases) and examination administration problems (acoustics of the room, 23 cases).

- Test anxiety is experienced by the highest proficiency group (44% frequency) as well as the lowest proficiency group (56% frequency). Test anxiety has a significant low, positive correlation with foreign language test anxiety and slightly higher, but negative correlations with language proficiency level in all tasks.

10.2 Questionnaire study in year 2000

I used the same measurement instrument for foreign language test anxiety in Year 2000 study, but substituted test-anxiety questionnaire part (Spielberger et al 1978) with Purpura's (1999) questionnaire on meta-cognition, thus the focus in year 2000 is on the interaction between anxiety, language proficiency and meta-cognition.

10.2.1 The aims

The aim of the Year 2000 study was to assess the level of test and classroom anxiety with the help of a questionnaire, measure the relationship between these and performance on the Year 12 exam, find out the reasons for test anxiety, examine the meta-cognitive strategies used by the test takers and study the interaction between meta-cognition, test anxiety and language proficiency.

10.2.2 The test-takers

All the test-takers were graduating from secondary school. Out of approximately 2000 questionnaires received, 248 responses were selected using a hierarchical multi-stage stratification method (Dyer 1997). As in the previous study, the sample was made up from different regions of the country, and from different types of schools (specialised language schools, city schools, small country schools and boarding schools) to represent the whole range of proficiency levels and to preserve the proportions of the type of students of the whole country (see Table 31).

Table 31 Representativeness of the sample

Type of school	Number of schools	Number of questionnaires
Big secondary schools in Riga (the capital)	4	21
Riga region schools	2	31
Small town schools	13	60
Big city schools	10	97
Specialised English language schools	2	28
Boarding schools and part time schools	3	11
Total	**34**	**248**

After being assessed and evaluated at the same time as the other school leavers, the selected test-takers' papers and questionnaires were separated for the analysis of the level of anxiety, use of strategies and the impact of these on test-performance.

10.2.3 Instruments

There were 2 instruments for this study: Year 12 English language examination results in year 2000 and a questionnaire.

10.2.3.1   Questionnaire

The year 2000 questionnaire (see Appendix 6) contained 31 questions that examined the notion of anxiety and the meta-cognitive strategy use:

- 12 questions referred to anxiety caused by the test

- 2 questions referred to the students' experiences during foreign language classes

- 1 question addressed the issue of whether the respondents had discussed test anxiety during the test preparation period

- 16 questions referred to the use of meta-cognitive strategies.

The first part of the questionnaire was identical with that of the year 1999. It started by asking the test-takers to circle the number (1 to 4) that would reveal the test-takers' level of anxiety (from *did not cause anxiety* to *caused strong anxiety*) for each task of the exam. For each task the test-takers were also requested to state the reason for their anxiety. They were informed that their responses would in no way influence the evaluation of their work, but would help us to develop a better test next year.

The second part of the questionnaire was targeted at the examination of the interaction between meta-cognitive strategies and affective schemata (see Bachman and Palmer 1996 discussion in section 3.4). The questionnaire was based on Purpura's (1999) Meta-cognitive Strategy questionnaire, separating goal-setting (*I set goals for myself in language learning*), planning (*Before I write a composition in English I plan my work)* and assessment (*When I speak English I know when I make grammar mistakes)* processes; it also included questions about language and classroom anxiety (for example, *Foreign language classes cause more anxiety even when I am well prepared)*. The test-takers had to state how often (from 1: *never* to 5: *always*) they applied these strategies or experienced these feelings.

Before examining the results of the questionnaire of the year 2000 I will address the issue of reliability of the questionnaires. Dyer (1997) defines reliability of the research instruments as *the extent to which the procedure is capable of returning as accurate result in spite the presence of factors which might influence the outcome in one direction or anther* (Dyer 1997, p. 129). Dyer considers that the reliability of a questionnaire consists of its reliablity to resist the influence of the passage of time,

fatigue or changes to motivation. Dyer suggests using either test-retest or split half reliability. Purpura (2000), however, used Cronbah's alpha for reliability study of his questionnaires (see section 5.2.1.2).

Davies et al (1999) define Cronbach's alpha as a measure of internal consistency and reliability. The authors propose that Alpha indicates how well a group of items measure the trait of interest by estimating the proportion of test variance due to common factors among items. Davies et al say that if all items measure the same underlying dimension, then the items will be highly correlated with all other items.

Cronbach's alpha for my for the whole questionnaire is .914 (in 1999, Appendix 5) and .832 (in 2000, Appendix 6) which suggests high level of agreement between the items in 1999 and lower, although still sufficiently high in 1999. The lower correlation of the year 2000 questionnaire is caused by Purpura's questionnaire (separately Cronbach alpha for the questions included in my questionnaire from his questionnaire is .780), which is similar to what Purpra reports in his book (see section 5.2.1.2).

10.2.3.2   The language test

The Year 12 examination was used to measure the test-takers' language proficiency. The test was prepared, administered and marked as in previous years (see Chapter 7). The reliability and the validity of the test were checked by ITEMAN (see Appendix 9) and inter test correlations were carried out (see Table 32). The items in each skill were weighted to achieve equal weighting between the separate tests (20% for each test).

The objective part of the test was analysed by ITEMAN (the Alpha coefficient for the Reading Test was .89, for Listening .88 and in Language Use .87). The correlations between the first and the second marking in Speaking were .59, and in Writing .76. The lowest correlation was between the first and second marking of Task 1 (.40), which was caused by some problems that will be discussed below.

After a third marking and averaging, the correlation between the total result and Speaking test results was .84 and for Writing test results .85. It is interesting to note that the live marking in Speaking again had a higher correlation with the total (.80) and had a higher mean (20.58 points) than the second marking (.64) and (18.42) just like in year 1999 study.

The standard deviation of the whole exam was 15%, which was adequate for the needs of a proficiency test. The minimum score was 8%, the maximum 98%; the mode was 53% and the mean 56%. The standard errors of measurement for the objective tests ranged from 2.3 (Reading) to 2.5 (Listening).

The impact of each language skill on the total test score can be seen in the correlations of the different tasks which ranged from .48 (Writing task 1) to .85 (Reading Task 2) (see Table 32 and Appendix 13 for the whole matrix of intercorrelations). The correlation for the first writing task is very low and probably was caused by the extremely low word limit, for which many students were penalised because they exceeded it. As a result this task was measuring not only language skills, but also learning strategies (see the interaction between the meta-cognitive strategies and performance in this task in section 10.2.5).

Table 32 Inter correlations of the Year 12 exam

Task	Nr.of items or points	Standard deviation	Correlation with the total
Reading	**40**	**7.46**	**.912**
Matching	10	1.80	.543
Gap-filling	15	3.56	.853
Matching	15	3.57	.783
Listening	**40**	**6.70**	**.900**
Gap-filling	20	3.62	.846
Multiple-choice	13	2.08	.734
Paraphrasing	7	1.94	.739
Language Use	**40**	**7.14**	**.915**
Multiple choice	13	2.54	.764
Editing	14	2.97	.780
Gap-filling	13	2.72	.834
Writing	**50**	**7.66**	**.851**
Email	8	1.28	.484
Recipe	17	3.75	.696
Diagram descr.	25	4.47	.727

	Nr.of items or points	Standard deviation	Correlation with the total
Task **Speaking**	**30**	**5.22**	**.842**
Total	200	15%	N.A.

10.2.4 The administration of the questionnaire

The staff of the Curriculum and Examination Centre photocopied the forms and packed them together with the examination papers. The test-takers filled them in during the exam; the invigilators collected them and sent them back to the Examinations Centre together with examination materials.

10.2.5 Results of the study

Using the mean and the standard deviation the whole sample was split into four groups according to their level of **proficiency** and into four groups according to their level of **anxiety** (for the method of dividing the test-takers into the groups see section 10.1). The aim was to separate the language proficiency and anxiety variables (see tables 33 and 34).

Table 33 Distribution into proficiency groups

Groups	Name of the group	Number of students	Mean proficiency level	Mean anxiety level
Proficiency S.D >1	P4	41	80%	47%
Proficiency S.D 1…0	P3	88	65%	55%
Proficiency S.D 0…-1	P2	79	51%	56%
Proficiency S.D <-1	P1	40	35%	58%
Whole sample		248	58%	55%

Table 34 Distribution into groups according to level of anxiety

Groups	Name of the group	Number of students	Mean anxiety level	Mean proficiency level
Anxiety S.D >1	A4	34	77%	53%
Anxiety S.D 1…0	A3	89	62%	56%
Anxiety S.D 0…-1	A2	79	48%	60%
Anxiety S.D <-1	A1	46	35%	63%
Whole sample		248	55%	58%

Just as in year 1999 study the mean of level of anxiety for each task (Questions 1-12) was correlated with the mean of language proficiency, but this time it was compared not only in the four groups of different levels of proficiency, but also in groups of different levels of anxiety. The comments on the causes of anxiety were also grouped like in year 1999 study, but again they were examined not only in groups of proficiency, but also in groups of different levels of anxiety.

The mean for the frequency of each meta-cognitive strategy (Questions 13-32) was calculated to compare the use of each strategy in different groups of proficiency and anxiety. The correlation between language proficiency and strategy use was calculated to examine the interaction between the use of strategy and test performance. Afterwards, the correlations between strategy use and anxiety level were calculated for each proficiency and anxiety group in order to examine the interaction between goal setting, planning and assessment strategies with foreign language anxiety.

10.2.5.1   Foreign language test anxiety level

According to Madsen (1982) one of the main reasons for anxiety is the perception of task difficulty. Therefore I will start with the difficulty level of the items used to measure the level of anxiety. Then I will compare the level of anxiety in the groups of different levels of **anxiety** and different levels of **proficiency** and finally I will look at the interaction between test anxiety and performance in each task.

As the Year 12 examination in 2000 was prepared according to the same test specification as the previous years', the mean as in all the secondary school graduation examinations in Latvia was 58%, that is within 50-60%.

If we look at the fluctuation of the level of difficulty of the separate tasks (see Table 35), we can see that the most difficult task was Language Use task 3, where the mean

was only 26%, and the next most difficult were the third tasks in Listening and Reading (37% and 34%). Although the level of difficulty of these tasks is very high, this complies with the test specifications requirement that the level of difficulty of separate tasks should increase, as the third task is targeted at learners of English as a first foreign language, who have been studying English for 12 years. The easiest task in the whole test was the first task of the Reading test (86%), which was meant to encourage and reassure the test takers. The first tasks in all the tests were the easiest, as planned by the test specification and the third the most difficult (with the exception of the Writing, where the second task turned out to be more demanding).

The anxiety level of the whole sample (248 test takers) ranged from 41% (Reading task 1 and Writing task 1) to 70% (Listening task 3) (see Table 33). The tasks that had an anxiety level above Madsen's 'debilitating threshold' of 60% (see section 4.3), were Reading task 3, Listening task 3, Language use task 3 and Writing task 2 and according to Madsen should be removed from the examination as they could be biasing the test-results in favour of the students who are less anxious.

Table 35 Level of difficulty and test anxiety level in different groups of sample

Task	Whole sample		Groups of different **proficiency** level		Groups of different **anxiety** level	
	Mean (%) of the task	Mean (%) anxiety of the task	P4 (%)	P1 (%)	A4 (%)	A1 (%)
Reading	**48**	**54**	**45**	**54**	**76**	**35**
Matching	86	41	34	45	62	32
Gap-filling	37	52	42	51	77	32
Multiple-choice	34	69	59	64	90	41
Listening	**61**	**56**	**49**	**56**	**81**	**33**
Gap-filling	65	47	40	49	72	29
Multiple-choice	68	52	46	51	79	31
True/false	34	70	60	69	93	40
Language Use	**50**	**56**	**46**	**61**	**80**	**33**
Multiple choice	62	47	35	54	71	28
Editing	60	55	46	59	78	32
Gap-filling	26	65	56	71	90	38
Writing	**54**	**53**	**48**	**62**	**72**	**39**
Email	69	41	38	47	61	30
Recipe	54	66	63	74	78	52

Task	Whole sample		Groups of different **proficiency** level		Groups of different **anxiety** level	
	Mean (%) of the task	Mean (%) anxiety of the task	P4 (%)	P1 (%)	A4 (%)	A1 (%)
Diagram description	58	53	44	64	76	36

If we look at the reactions across different levels of **proficiency,** for the highest proficiency group (P4), the most anxiety inducing task was Writing task 2 (63%) and the least worrying were the Reading and Language Use tasks 1 (means 34% and 35%) and only one of the Writing tasks (Recipe) exceeded the 60% debilitating threshold (Madsen 1982), while in the lowest proficiency group (P1) there were 5 tasks that were higher than 60%: Tasks 3 in Reading, Listening and Language Use and Task 2 and 3 in Writing.

Figure 21 shows the fluctuation of the mean of the level of anxiety for the whole sample. Just like in year 1999 study it follows the changes of the real difficulty level and not the expected one (Writing task 2 was more anxiety inducing than task 3 and also more difficult)
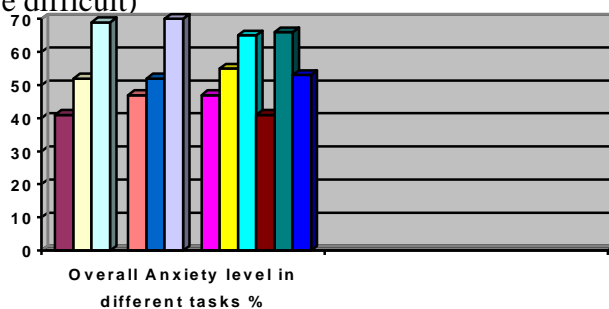
Figure 21 Anxiety level fluctuations in Reading, Listening, Language use and Writing tasks

If we look at the reactions of the test takers across different levels of **anxiety**, the group with the highest level of anxiety had only 2 tasks that were lower than the debilitating threshold. If we remember that the mean proficiency level of this group was 53% and it formed nearly 14% of the whole sample (see Table 34) then we might extrapolate from this to say that approximately 1 in 10 students of different levels of proficiency of English was working above the debilitating threshold level in the Year

12 examination. That was approximately 500 students of the 5184 population of the year 2000.

Madsen's (1982) suggestion that tasks that are too difficult should be excluded from the test would not help this group, as the group's mean proficiency level was 53% and the tasks were not too difficult for them.

The causes of the anxiety will be discussed below. Here I will just mention that the most surprising result for the test development team was Writing task 2, which asked the students to write a recipe for a typical Latvian dish (under the rubric "instructions" in the test specifications). Judging by the students' responses in the questionnaire the topic of food was considered to belong to the beginners' level and had not been repeated in the secondary school. As a result it came as a surprise to the test takers (see the 'specific theme' as a cause of anxiety in section 10.2.5.3.).

10.2.5.2   Relationship between foreign language test anxiety and test performance

In Table 36   we can see that most of the correlations between anxiety and test performance for the whole sample are negative although not all of them are significant. The correlation between level of anxiety and performance in a particular task is significant (at p<0.01), as opposed to the correlation with the performance in the overall proficiency (apart from Language Use), which suggests that anxiety is the reaction to the problem in the particular task that the test-taker has become aware of.

Table 36 The correlation between anxiety level and proficiency level

| Task | For the whole sample | | Groups of different levels of proficiency | | Groups of different levels of anxiety | |
	With the overall proficiency	With the result in the particular task	P4	P1	A4	A1
Reading	n.s.	n.s.	**-.195**	**.417**	n.s.	n.s.
Matching	n.s.	-.438	n.s.	.219	n.s.	n.s.
Gap-filling	n.s.	-.374	n.s.	.292	n.s.	-.281

Task	For the whole sample		Groups of different levels of proficiency		Groups of different levels of anxiety	
	With the overall proficiency	With the result in the particular task	P4	P1	A4	A1
Multiple-choice	n.s.	-.423	n.s.	.504	n.s.	.206
Listening	n.s.	**-.207**	**-.262**	n.s.	n.s.	**-.273**
Gap-filling	n.s.	-.293	-.290	n.s.	-.341	-.310
Multiple-choice	n.s.	-.242	-.236	n.s.	n.s.	n.s.
True/false	n.s.	-.228	-.396	.230	n.s.	n.s.
Language Use	**-.246**	**-.258**	**-.387**	**.198**	n.s.	n.s.
Multiple choice	-.266	-.280	n.s.	.233	n.s.	n.s.
Editing	-.184	-.310	n.s.	n.s.	n.s.	n.s.
Gap-filling	-.193	-.385	-.351	.238	n.s.	n.s.
Writing	**-.198**	**-.257**	**-.456**	n.s.	**-.279**	**-.245**
Email	n.s.	-.207	-.346	n.s.	-.299	n.s.
Recipe	n.s.	-.186	-.283	n.s.	-.191	n.s.
Diagram description	-.248	-.232	-.377	n.s.	-.505	n.s.
Total	-.219	N.A.	-.355	.194	-.117	-.148

The second pair of coloumns shows the interaction between test anxiety and performance in groups accroding to different levels of **proficiency**.

We can see that most of the correlations between proficiency and anxiety level in the highest proficiency group are negative, while in the lowest proficiency group they are positive, although very low and mostly insignificant. This supports the view that the relationship between performance and anxiety is a curvilinear (Mathew et al's 2000), therefore if the population is split into groups using the distribution curve, then one group will have positive and another negative correlation.

The highest positive correlation for group P1 is in Reading test (.42). It seems that the weaker test takers have also been the least anxious and the students who scored higher also understood what to worry about.

The group that has the highest negative correlation is the highest proficiency (P4) group (-.36) and evidently the highest proficiency level test-takers were concerned about inability to perform as well as they wished. This agains suggests the connection between assessment strategies and anxiety: the higher proficiency group's assessment strategies are evidently more focussed.

If we look at the correlations between levels of anxiety and proficiency across different levels of **anxiety**, we notice a change of pattern. The only significant correlations (p<0.01) are in the Listening and Writing tests. If we look at the correlations of the highest anxiety level group (A4) which has a mean level of anxiety in Reading Task 3 of (90%) (see Table 36), it is surprising that the correlation between the performance level and the anxiety level is non-existent: (.029). The level of anxiety for the same group (A4) in Listening task 3 is 93.4% but the correlation is (-.34). The correlation for Writing task 3 is (-.51), although the level of anxiety in this task is "only" 75.7%, which is not very high for this group. The interaction between the level of anxiety and language performance is so different in different language skills, that it appears that Cheng et al (1999) may be right in saying that what we should be talking about is language skill specific anxieties and not foreign language anxiety in general (see section 4.3).

10.2.5.3   Causes of anxiety

The test takers have provided 566 causes of anxiety for the 12 tasks, that is, 2.29 causes per person. The two tasks, which were most often commented on are Writing Task 2 (93 times) including, *unexpected theme* (29 times) and *unexpected task* (27 times) and Reading Task 3 (75 times), including *unknown words* (17 times), *difficult text* (22 times), *lack of time* (16 times).

I have divided all the causes of anxiety mentioned by the test takers into three groups:

1. feelings experienced during the examination

2. foreign language use

3. examination design and administration

10.2.5.3.1 State anxiety

While I was examining the causes of anxiety provided by the test takers I often encountered comments that were not connected with language use. Test takers often assessed their ability to satisfy the test demands:

- A large group of causes is directly connected with concern about the difficulty of the test and the tasks, for example: *Difficult to understand the complicated words. It was all rather difficult.* These statements could be classified as task assessment (Bachman and Palmer 1996).

- Test takers often say that they were worried by their own reactions, for example: *My own agitation* or *Lack of confidence in what I am writing and in myself* or *My first impression about the task. I am not sure that I am doing the right thing. I was not sure about my knowledge. Difficult to think, if anxious, cannot think at all!* This according to Bachman and Palmer (1996) could be classified as self-assessment.

- There are also phrases that suggest that some test-takers knew themselves and were prepared for their reaction: *It's an exam!* Or *You should know yourself!* Also *It is natural!* And even *There is no concentration without anxiety,* which is an illustration of Csikszentmihalyi's (1998) theory that anxiety helps to focus and achieve flow (see section 2.1).

- Test takers use the word 'afraid' to describe their feelings before, during and after the test, for example: *Afraid that I will not know anything, afraid that I will not manage or will not hear the necessary information or say the wrong thing. I am afraid I have filled in the wrong word, and I am sure I'm right in my feeling.* This supports Onwuegbuzie et al's (2000) proposal

that foreign language anxiety consists of input, processing and output anxiety.

All the causes that were not directly connected with language use or features of the test itself, are grouped together under the title 'State anxiety', as it is the test takers' reaction to their being in the test situation (see Table 37).

If we compare the frequency of mentioning their experiences between different proficiency groups, we can see that all the causes are distributed equally, therefore I will analyse the frequency of comments across all the levels of **anxiety** (see Table 37)**.**

The most often mentioned cause of anxiety is general anxiety about the test (56), out of which the highest anxiety group alone mentioned this 33 times, which is more than a half. In this group I have placed all the remarks quoted above and also those cases when the word "EXAM!" in capital letters was written in the place where the reason for anxiety in each task was sought.

The next most often mentioned cause by the highest anxiety group was the complaint about the difficulty of the task (14 times), which interestingly, was seldom mentioned by the lowest language proficiency group (7 times), but most often by the medium proficiency groups (20 times). Another complaint of the high anxiety groups is the difficulty in choosing the right version in multiple-choice tasks (10 times) and being restricted by the word limit in writing (4 times).

The complaints in the lower anxiety level groups are: *lack of confidence* (24) *forgetting words* (4), *not knowing what to write* (3), *inability to concentrate* (2) and *freezing hands* (1). If we compare the most often mentioned comments in these groups we notice that the high anxiety group's (A4) chief concerns are with writing

too much and the difficulty in circling the right choice, while the low anxiety group (A1) test-takers do not know what to write, cannot concentrate on the task and are even too cold. This is an example of May's (1979) theory that healthy anxiety activates and provides ideas about how to reach one's aim. At the same time this also agrees with Stevick's (1999) proposal that affect may place so many ideas on the worktable that one is at a loss to know how to cope with everything.

Table 37 The frequency of self-reported causes of anxiety connected with test state

	All	A4	A3	A2	A1	P4	P3	P2	P1
General anxiety	56	33	11	8	4	6	30	18	2
Lack of confidence	24	4	4	4	12	0	9	10	5
Task too difficult	53	14	17	13	9	6	20	20	7
Forget words	7	0	2	4	1	3	1	0	3
Difficult to choose the right version	13	0	10	2	1	9	0	4	0
Did not know what to write	5	1	0	3	1	0	1	2	2
Too short composition	**5**	**1**	**4**	**0**	**0**	**2**	**2**	**1**	**0**
Could not concentrate	2	0	0	2	0	0	0	1	1
Boys do not have to know recipes	2	0	0	1	1	1	0	0	1
Hands were freezing, could not write	1	0	0	0	1	0	0	1	0
Total	168	53	48	37	30	27	63	57	21
Number of test takers	248	34	89	79	46	41	88	79	40
Number of responses per person	0.69	1.23	0.56	0.47	0.65	0.71	0.72	0.72	0.53

10.2.5.3.2 Foreign language use anxiety

The causes connected with language use (267) as compared to state anxiety (168) are mentioned by the test takers much more often. (see Table 38). I have not separated the different language skills, because most of the reasons mentioned by the test takers are repeated in all the language skills. Only listening stands out with 100 causes out of 267, which is more than 1/3 and supports Vogley's (1998) statement that *Anxiety that accompanies the listening comprehension task is difficult to detect, but is potentially one of the most debilitating, because in order to interact verbally the listener must first understand what is being said* (Vogley 1998, p.67).

It is possible, that some of the reasons mentioned were caused by technical problems with tape recorders, but there were no complaints from the invigilators, and we can see that students in the lowest anxiety level (A1) complain of the bad recording very seldom (2) as opposed to the group A3 (22). On the other hand, there are also the 11 test-takers in the highest proficiency level (out of 41) who complain about the bad recordings. In spite of the high correlations between the listening test and the total proficiency level, perhaps the invigilators' observation questionnaire should contain a question about the quality of the recording.

Table 38 Foreign language anxiety in different groups of the sample

	All	A4	A3	A2	A1	P4	P3	P2	P1
Difficult text	46	10	18	13	5	5	20	17	4
Specific topic	40	1	16	17	6	6	15	18	1
Unknown words	43	5	15	22	1	4	15	14	10
Vocabulary problems	19	4	6	5	4	6	7	6	0
Bad recording	39	3	21	14	2	11	27	8	3
Speech in the listening texts too fast	32	2	13	15	2	6	13	13	7
Unclear speech	11	0	7	1	3	0	4	3	4
Listening too fast	18	1	8	5	4	2	4	9	3
Do not know grammar	14	0	2	9	3	6	3	3	2
Do not know spelling	5	0	3	2	0	0	3	2	0
Total	267	26	109	103	30	40	91	93	34
Number of test takers	248	34	89	79	46	41	88	79	40
Number of responses per person	1.08	0.76	1.22	1.30	0.65	1.0	1.03	1.18	0.85

Apart from listening problems one can notice the preoccupation with problems of what Bachman (1990) calls organizational competence, both grammatical and textual. Vocabulary problems are referred to in different forms: *unknown words* in comprehension (43 times) as *vocabulary problems* in production (19 times), the lack of knowledge of vocabulary in the *specific topic* (40 times). All in all there are 102 complaints and if we also add 7 cases of *forgetting words in agitation* from the state anxiety reason group, vocabulary causes even more worry than the Listening test does. This supports Stevick's (1999) proposal that affect controls access to memory: the words cannot be accessed because one is conscious of stress. As a result, the task demands cannot be satisfied, anxiety level grows and it becomes even more difficult to access words.

Textual competence has also caused anxiety and is mainly referred to as *difficult text* in both listening and reading (46 times). This group of causes may also contain socio-linguistic competence problems as they could be one of the reasons for a text being difficult, and were sometimes mentioned in the comments on the Listening test (for example, *How can a person speak so fast?* or *the use of English was unusual)*. Most of the students complain of speed, not hearing and not understanding the text without providing reasons.

Language use problems as opposed to the comments on general state anxiety are mentioned more often in the medium proficiency and medium anxiety groups. The lowest proficiency (0.85 answers per person) and the lowest anxiety (0.65 answers per person) group commented on the language use problems very seldom.

10.2.5.3.3 Examination design, training and administration

Although the examination was prepared according to the test specifications, and all the administrators were trained in test administration and all the tasks were pre-tested, the test takers' questionnaires still reveal problems that caused anxiety (see Table 39). The most common complaint is about the unexpected tasks and themes, which occur in all skills, but the most surprising concern was the Writing Test (12 cases concerning Writing Task 1 and 27 for Writing Task 2) (see the discussion about the reasons this chapter above). This supports Shohamy's (1982) view that the unexpected in a test is a frequent cause of anxiety (see also 'novelty check' during the appraisal in Scherer 2000, section 2.3.7).

Table 39 Examination design, training or administration problems (random factors)

	All	A4	A3	A2	A1	P4	P3	P2	P1
Unexpected task or theme	49	4	23	14	8	7	29	8	5
Lack of time	47	2	23	16	6	7	16	17	7
Did not understand the task	15	2	4	8	1	3	2	7	3
Layout of the task caused problems	5	0	2	2	1	1	3	1	0
Unclear pictures in reading	8	0	5	2	1	2	4	2	0
Noise behind the door	3	0	3	0	0	1	1	1	0
Did not like the task	2	0	2	0	0	2	0	0	0
What do I write in the Draft?	2	0	0	1	1	1	1	0	0
Total	131	8	62	43	18	24	56	36	15
Number of test takers	248	34	89	79	46	41	88	79	40
Number of responses per person	0.52	0.24	0.67	0.54	0.39	0.54	0.64	0.46	0.38

The second most frequently mentioned cause of anxiety in the examination problems is lack of time (47). This is mentioned most often in Reading (21 times) in spite of the fact that the time was increased this year from 40 minutes to 50 minutes after the last year's complaints. The complaint about the lack of time is made by all groups, both

across proficiency levels and anxiety levels, so we can assume that it does not discriminate between the groups.

If we compare the frequency of comments on the examination problems, we can see that it is the higher proficiency level test-takers (0.54 and 0.64 comments per person in Groups P3 and P4) who are more conscious of these, as compared to the lowest proficiency level test-takers (.38 comments per person).

10.2.5.3.4 Distribution of the causes of anxiety

The most often mentioned causes of anxiety in the whole sample (see Figure 22) are connected with language use (1.08 responses per person), test state anxiety is mentioned much less often (0.69 responses per person), and test problems are commented upon the least often (0.52 responses per person).
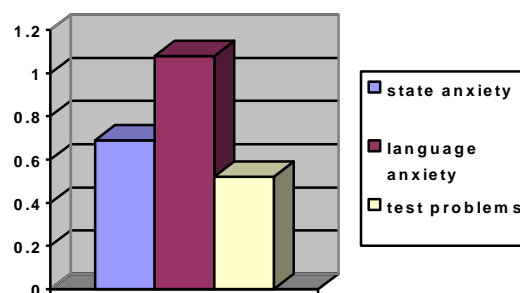
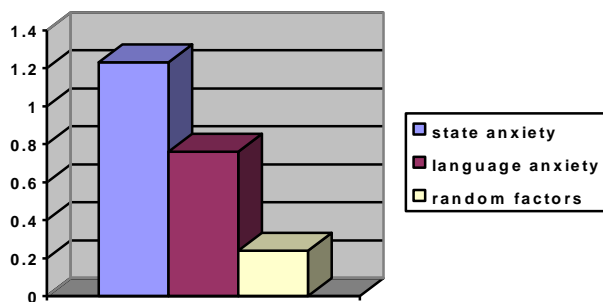**Figure 22 Distribution of the causes of anxiety in the whole sample**

If we compare the distribution of the causes of anxiety according to the different groups of the sample (see Table 40), we can see that the most comments were made by the group A3 (2.45 causes per person) where anxiety level is just above the mean. The least comments (1.69 causes per person) were made by the group A1 which had the lowest level of anxiety and the group P1 which had the lowest level of proficiency (1.76 causes per person).

If we compare the distribution of causes of anxiety of the whole sample (see Figure 22), we can see that the test takers are more concerned with the use of language (1.08 responses per person), less with the test situation (.69) and still less with the random factors such as problems within the test (.52).

Table 40 Distribution of the causes of anxiety

	All	A4	A3	A2	A1		P4	P3	P2	P1
State anxiety Number of responses per person	0.69	1.23	0.56	0.47	0.65		0.71	0.72	0.72	0.53
Language anxiety Number of responses per person	1.08	0.76	1.22	1.30	0.65		1.0	1.03	1.18	0.85
Random factors Number of responses per person	0.52	0.24	0.67	0.54	0.39		0.54	0.64	0.46	0.38
Total	2.29	2.23	2.45	2.31	1.69		2.25	2.39	2.36	1.76

If we compare the distribution of causes of anxiety in the different **proficiency** level groups, we can see the same tendency: the more proficient the students, the more they are disturbed by test problems: groups P4 and P3 mention the random factors .54 and .63 times per person, while Group P1 make them only 0.38 times per person.

Figure 23 Distribution of causes of anxiety in the group of highest anxiety

If we look at the distribution of the causes of anxiety of the different groups of **anxiety** (see Figure 23), we can see that the group with the highest anxiety level has a totally different pattern from the other groups: most of the comments relate to test experience (1.23 responses per person), the language anxiety is mentioned less often (0.76 per person) and the test problems are mentioned least often (0.24). This supports

Wine's (1971) theory that the test-anxious test-takers are least concerned with the test and its problems, as they are focused on their own experiences. In the case of test design problems it would mean that the anxious test-taker blames his or her incompetence for his or her inability to perform in a satisfactory manner and not the test.

10.2.5.4   Relationship between foreign language classroom anxiety and foreign language test anxiety

Horwitz (1986) found strong correlations (.56) between classroom anxiety and test anxiety. The test takers in 1999 questionnaire did not mention their previous experience in language learning as being directly connected with their test anxiety. Therefore in the year 2000 questionnaire I decided to include three questions focusing on classroom anxiety and one on attempts to overcome test anxiety during the preparation period. The results of the analyses can be seen in Table 41. The test-takers had to respond to statements by choosing numbers from 1 (*never*) to 5 (*always*). These were converted to the mean per cent so that they could be interpreted as frequency (from 100% (always) to 0% (never).

Table 41 Classroom anxiety means in different groups

Whole sample			Proficiency level groups				Anxiety level groups			
	Question	All	P4 %	P3 %	P2 %	P1 %	A4 %	A3 %	A2 %	A1 %
13	While preparing for the examination we discussed with the teacher the effects of test anxiety on the examination score	43.4	41.5	45.3	42.6	43.0	49.1	43.6	43.8	38.3
14	Foreign language classes worry me more than other classes even if I am well prepared	35.7	28.3	33.1	39.5	41.5	50.3	35.9	33.7	28.3
15	I am afraid that my classmates will laugh at me if I speak English	28.6	22.4	25.3	32.3	35.0	40.0	29.8	25.1	24.3

The whole sample reports that the frequency of discussions of test anxiety with their teacher was 43%, the highest frequency is reported by the same group (A4) that reported the highest level of test state anxiety (49%). This group has also reported

foreign language classes as being more worrying than other classes (frequency 50%), and 40% frequency of being afraid that the other students will laugh at them if they use English during the class.

However, only 35% of the weakest proficiency group (P1), where one would expect students to make mistakes when speaking, reports that they are afraid of being laughed at. Thus it is not the level of proficiency that causes foreign language classroom anxiety.

**10.2.5.4.1 Correlation between level of classroom anxiety and level of performance**

Scherer (2000) suggested that the use of linear correlations for investigation of the relationship between cognitive and affective variables is impossible as in the production of emotions different systems operate parallely (physiological, cognitive elements and action tendencies in the form of motivation) each adding their specific characteristics to the emotional experience (see section 2.3.7).

If we look at the scatterograms depicting interaction between the mean anxiety during the examination as measured by anxiety questionnaire in 2000 (see Appendix 6) and the test performance as measured in year 12 examination in 2000 Language Use and Speaking test we see that they do not suggest significant relationship (see Figures 1 (Language use) and 2 (Speaking test)). On the other hand, the scatterograms do not suggest that the relationship is curvelinear as proposed by Yerke-Doddson law either. Thus the distribution of the points in both scatterograms suggest that the linear correlation can be used to investigate the relation between the language performance and anxiety level, because, although there can be different systems involved as suggested by Scherer (2000), the relationship may still be linear.
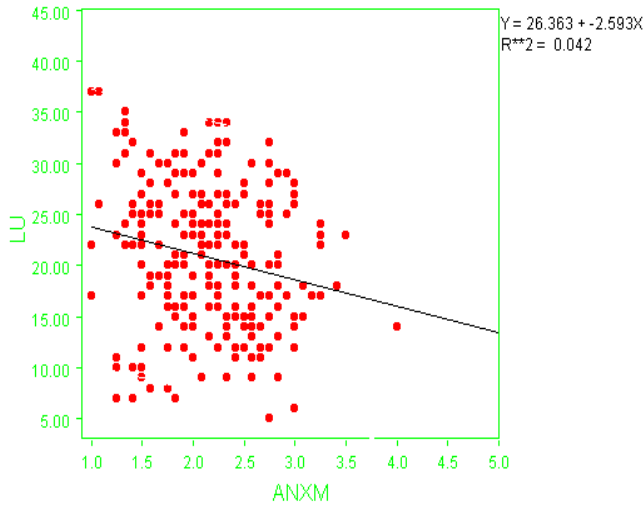
Figure 24 Scatterogram depicting relation between Language Use test performance and mean anxiety level
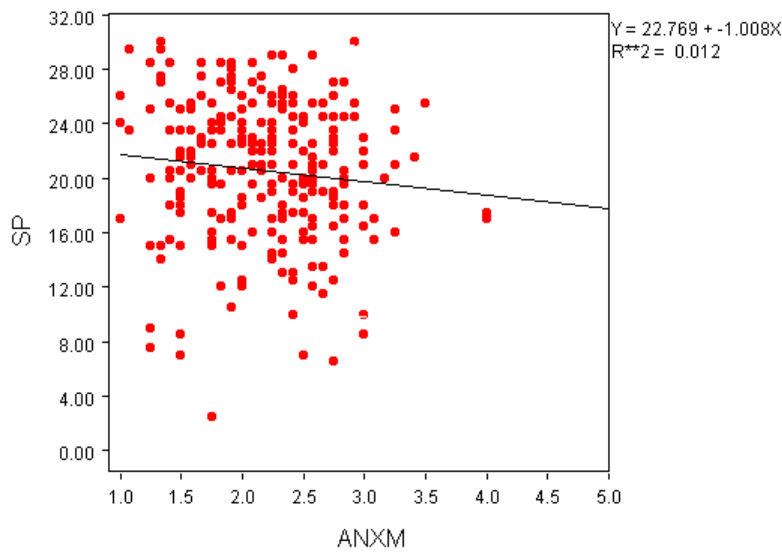
Figure 25 Scatterogram depicting relation between Speaking test performance and mean anxiety level

As predicted by the studies of Horwitz and Gardner (see section 4.3) the analyses of the responses of the whole sample show that there is a negative correlation between level of classroom anxiety and level of foreign language proficiency. The question about the discussion of the effects of anxiety (Question 13) showed that there was no

significant correlation between the reported frequency of discussions and the performance of the whole sample.

If we look at the correlations between the same responses and the level of performance in the different **proficiency** level groups (see Table 42) we can see that there are some fairly strong correlations with some individual language skills (I have included only the significant ones at p<0.01).

Table 42 Correlations between classroom anxiety and language proficiency according to groups of proficiency

		All	P4	P3	P2	P1
13	While preparing for the examination we discussed with the teacher the effects of test anxiety on the examination score	n.s.	.436 LU .341 W3 -347R1 .346R2	n.s	.265 LU3 .272 R1 -.226R2	.219LU2
14	Foreign language classes worry me more than other classes even if I am well prepared	-.280	-.521LU2 -.207L3	.249W -.322S2 -.218L1	-.387TOT -.364L1 -.305R	.347LU3 -.310L1
15	I am afraid that my classmates will laugh at me if I speak English	-.290	-.357L -.240 S -.299LU2	.225LU .291W3 -.259S2 -.235L1	-.218LU3 -.228W1	.234LU1 .210 S2

It is interesting to note that the discussions about anxiety (Question 13) correlates positively with Language Use performance for all proficiency levels (except group P3) and, logically, being afraid of classmates' laughter if one speaks in English (Question 15) also correlates negatively with Speaking test performance in all proficiency groups, (again except group P3).

Apart from the two above-mentioned features, the correlations between the level of anxiety and language performance are very inconsistent and it is difficult to explain why they are positive for one task and negative for another. I will, therefore, now turn to the relationship between the level of classroom anxiety and test-performance in different groups of test anxiety.

If we examine the correlations between the level of classroom anxiety and test performance in the **anxiety** groups (Table 43) we find regular correlations. The

strongest correlations in the A4 group is Question 14: (*Foreign language classes worry me more than other classes even if I am well prepared*); it correlates negatively with the performance in all skills, with the highest correlation relating to Speaking (-.335).

In medium anxiety level groups (A3) and (A2) the responses to both classroom anxiety questions correlate negatively in all skills and all the tasks (see Table 43), although Question 15 (*I am afraid that my classmates will laugh at me if I speak English)* actually refers to speaking anxiety. In the group (A2) the speaking performance in this question has the highest negative correlation (-.401), but for group (A3), the Language Use test performance correlates slightly higher than Speaking (-.297).

Table 43 Correlations between classroom anxiety and performance in different anxiety groups

	A4	A3		A2		A1	
Task/question	14	14	15	14	15	13	15
Reading	**-.297**	**-.288**	**-.271**	**-.306**	**-.282**	**.305**	**-.231**
Matching	-.259	-.243	-.237	-.219	n.s.	n.s.	-.265
Gap-filling	-.255	-.276	-.198	-.187	-.227	.264	n.s.
Matching	-.200	-.198	-.249	-.337	-.309	.299	-.265
Listening	**-.212**	**-.272**	**-.291**	**-.319**	**-.288**	**.189**	**-.321**
Gap-filling	-.256	-.244	-.304	-.317	-.293	.193	-.411
Multiple-choice	n.s.	-.311	n.s.	-.301	-.217	.220	-.270
Gap-filling	n.s.	n.s.	-.297	-.175	-.210	n.s.	n.s.
Language Use	**n.s.**	**-.267**	**-.297**	**-.233**	**-.204**	**.207**	**-.200**
Multiple choice	n.s.	n.s.	n.s.	n.s.	-.232	n.s.	n.s.
Editing	n.s.	-.252	-.222	-.252	n.s.	n.s.	-.244
Gap- filling	n.s.	-.243	-.333	n.s.	-.293	.309	-.212
Writing	**n.s.**	**n.s.**	**-.225**	**-.188**	**-.293**	**.230**	**-.234**
Email	n.s.	n.s.	-.190	n.s.	n.s.	.382	n.s.
Recipe	-.206	n.s.	-.191	-.246	-.248	.209	-.230
Diagram descry.	n.s.	n.s.	-.225	n.s.	-.217	.385	n.s.
Speaking	**-.335**	**-.305**	**-.257**	**-.309**	**-.401**	**.283**	**-.185**
Total	**-.259**	**-.296**	**-.305**	**-.310**	**-.334**	**.266**	**-.253**

The group that has the lowest level of anxiety has an overall positive correlation between the discussion of anxiety level in the class and test performance. The only other significant correlations that Question 13 produced were with the highest anxiety

group (A4): (.42 in Language Use and -.24 in Speaking and -.24 in Reading). This suggests that the discussion of test anxiety in the class has different effects on the performance of the different anxiety groups.

If we compare the impact of foreign language test anxiety and foreign language classroom anxiety on test performance as depicted by the correlations in Tables 34 and Table 28, we cannot see much difference as the correlations are weak although some of them are significant and one might come to the conclusion that the impact of foreign language test state anxiety and classroom anxiety are similar. However, the structural equation (see section 10.3) suggests that the impact of classroom anxiety is at least twice as big as the impact of foreign language test anxiety. Evidently Scherer (2000) was right to say that linear correlations are not the most appropriate method for evaluating of the interaction between different nervous system components (see section 5.2.2). Nevertheless, the correlations do reveal the existence of an interaction between anxiety and performance. I will now examine the interaction between the two types of anxiety (classroom and test).

10.2.5.4.2 Correlation between foreign language classroom anxiety and foreign language test anxiety

Correlations between the classroom anxiety question responses and foreign language anxiety are significant and positive for the whole sample in both questions (Question 14: .336; and Question 15: .248). This was already predictable and suggests that classroom anxiety is also a cause of test anxiety (see Table 44).

Table 44 Correlations between Classroom and Test anxiety levels in each task

Gr.	Qt.	Test anxiety level in each task												Test anx. Mean
		R1	R2	R3	L1	L2	L3	LU1	LU2	LU3	W1	W2	W3	
P4	13	n.s.	-.244	n.s.	-.257	-.290	-.258	-.247	-.221	-.348	n.s.	-.215	n.s.	-.297
P3	13	n.s.	.186	.222	.292	.432	.263	.234	.193	.247	n.s.	n.s.	n.s.	.331
	14	n.s.	.234	n.s.	n.s.	n.s.	n.s.	n.s.	.317	n.s.	n.s.	n.s.	n.s.	.192
	15	n.s.	.278	.234	.144	.140	.337	.303	.273	.313	n.s.	n.s.	.186	.358
P2	14	.413	.271	.240	.323	.287	.104	.281	.335	.219	.446	.179	.284	.439

P1	13	.285	.254	n.s.	n.s.	.187	.132	.371	.282	.187	.298	.348	.183	.353
	14	n.s.	n.s.	n.s.	.303	.251	.264	.223	.285	.317	.036	.425	.293	.368
	15	n.s.	n.s.	n.s.	.343	.205	.284	.190	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
All	14	.271	.216	.164	.196	.183	n.s.	.270	.318	.244	.244	n.s.	.225	.336
	15	n.s.	n.s.	n.s.	n.s.	n.s.	.210	.201	n.s.	.204	.140	n.s.	.203	.248

For the highest proficiency group (P4) there are no significant correlations between Questions 14 and 15 and foreign language test anxiety in any of the tasks. Evidently they separated the problems that they encountered in the test from their previous anxieties. The only question that has significant correlations is Question 13: *While preparing for the examination we discussed with the teacher the effects of test anxiety on the examination score*. This correlates negatively only with this group, that is the more often the discussions were held, the less anxiety was felt by the highest performance group (P4) (-.297).

Question 13 correlates positively with anxiety level in Groups P3 (.331) and P1 (.353) creating an impression that in the medium proficiency level, the more teachers discussed anxiety, the more anxious test-takers felt during the test, although the level of anxiety is low for these groups and there is no significant correlation with anxiety level in the whole sample. So it seems that it is perhaps better for teachers to leave the question of test anxiety alone and not discuss it as only 40 people out of 240 benefited and nearly a hundred felt that the more they discussed it the more worried they felt in the test.

The highest positive correlation between classroom and foreign language test anxiety is for proficiency group P2 (.44) and although the correlations are lower than they were in Horwitz's (1986) study (.56), they still support the interaction between these types of anxiety.

10.2.5.5 Relationship between meta-cognitive strategies and foreign language test anxiety

Bachman and Palmer's (1996) model of language use shows interaction between 'affective schemata' and 'strategic competence' (see section 3.4). In their model strategic competence consists of cognitive and meta-cognitive strategies. It is the meta-cognitive strategies that are of most interest to me because of the similarity between their functions and those of affect (see section 3.3.1).

Purpura (1999) defines meta-cognitive strategy use as *a set of conscious or unconscious mental or behavioural activities which are directly or indirectly related to some specific stage of the overall process of language acquisition, use or testing* (Purpura 1999, p.6).

His interpretation of strategies as 'behavioural activities' suggest that the questions he uses to research meta-cognition can be used to research the impact of anxiety as well because I do not believe that any person would change his or her behaviour unless this was motivated by both cognition and affect. As Purpura's questionnaire is also based on Bachman and Palmer's (1996) language use model, his meta-cognition questionnaire contains questions, which explore all three areas of meta-cognition:

1. Goal setting process items in his questionnaire are designed to measure the extent to which the test-takers feel they use higher order executive function of identifying and choosing specific goals and objectives before or during the activity

2. Assessment process items are designed to assess the extent to which the test-takers feel they use their executive skills for assessing the situation (taking stock of conditions surrounding a language task by assessing one's own knowledge, available internal and external resources and constraints of the situation before engaging into it), monitoring (determining the effectiveness of the performance

while engaging in the activity), evaluating (determining the effectiveness after the performance)

3. Planning process items assess the extent to which the test-takers feel they use the executive function of generating an overall plan of action before they engaged in an activity; learning to learn (which Purpura also included in planning processes) items are designed to measure the extent to which the test-takers feel they arrange for the presence of conditions that helped them successfully accomplish the task and learn about how language learning, use and testing work (Purpura 1999, p.54).

I chose questions from his questionnaire to represent all three areas of meta-cognition (goal-setting, assessment and planning). Here I will present my findings on the use of meta-cognitive strategies and their interaction with language proficiency and anxiety using simple means and Pearson Product Moment correlation methods. In section 10.3 the same issues are discussed using the Structural Equation Modelling method.

This section consists of three parts

- analyses of the frequency of use of different strategies
- the interaction between the use of strategies and level of performance in different groups according to their level of performance or level of anxiety
- the interaction between strategy use and level of anxiety.

10.2.5.5.1 The frequency of use of meta-cognitive strategies

The mean of the frequency of use of meta-cognitive strategies in the different groups varies from 43.9% to 92.7% with an overall mean use of 68.7% (see Table 45), so all the strategies included in the questionnaire are well represented. The strategies that were used most frequently are the planning or monitoring strategies *Trying to understand when somebody speaks English* (Mean 89.6, Question 21) and *Trying to concentrate when writing the test* (Mean 86.1, Question 27). The least frequently used strategy is a goal setting strategy *When I begin to study, I plan what I am going to do*

(Mean 47.2, Question 16). If we compare the areas of meta-cognitive strategies, the use of planning is more frequently reported than are goal setting and assessment strategies.

I have compared the frequency of use of strategies across different levels of anxiety and proficiency groups. I will start with the use of strategies in the **anxiety** groups.
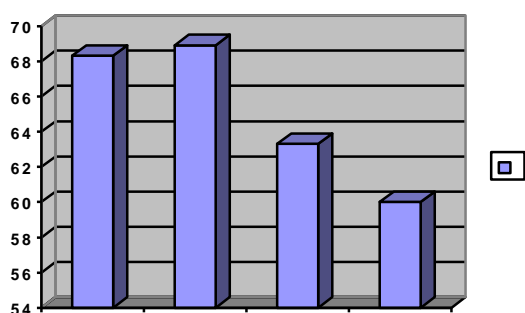
Figure 26 Frequency of use of assessment strategies in groups A4 (1st bar), A3 (2nd bar), A2 (3rd bar) and A1(4th bar)

If the sample is divided into groups according to level of **anxiety**, we can observe that in all groups the goal setting strategies again are the least frequently used, but the use of the self-assessing strategy in the high anxiety groups is higher (68.3 and 68.9) than it is in the lowest anxiety group (60%) (see Figure 24). If we remember Sarason's (1960) and Wine's (1971) claims that highly test-anxious persons are self-dissatisfied and instead of dealing with the tasks, are spending time on self-ruminations (see section 4.2), we can presume that this is caused by the too frequent use of assessment strategies.

Table 45 Frequency of the use of strategies in groups of different levels of anxiety

	Strategy	Process	All %	A4 %	A3 %	A2 %	A1 %
16	When I begin studying English, I plan what I am going to do	Goal setting	47.2	46.7	50.0	44.6	46.5
17	I set goals for myself in language learning	Goal setting	74.3	70.9	78.2	70.1	76.5

18	I do everything to improve my knowledge of English	Goal setting	68.0	67.3	71.8	65.6	65.2
19	I think of the ways in which to continue studies after school	Goal setting	68.9	68.5	68.2	69.9	69.1
20	I try to find a way that is best for learning new words	Goal setting	60.9	63.6	62.7	58.2	60.0
Goal setting total			**63.9**	**63.4**	**66.2**	**61.7**	**63.5**
21	I try to understand when someone is speaking English	Planning	89.6	89.1	92.7	89.1	84.8
27	Before writing a composition I plan my work	Planning	58.3	63.0	59.1	57.2	55.2
29	When I am taking a test I try to concentrate	Planning	86.1	81.8	90.2	85.8	81.7
Planning			**78.0**	**78.0**	**75.4**	**80.7**	**73.9**
28	Before doing an English assignment I think whether I have enough English to do it	Assessment of situation	53.7	55.2	59.5	52.4	43.9
26	Before writing a test I try to find out how it is going to be scored	Assessment of situation	58.1	66.1	60.0	54.4	55.2
30	When I am taking an English test I know how much time has gone by	Assessment (self monit.)	80.1	84.2	85.9	78.2	69.1
31	Before handing in a test I check my work	Assessment (self monit.)	71.0	72.7	75.9	67.8	65.7
23	When I speak English I notice when I make mistakes	Assessment (self-mon.)	63.0	59.4	61.8	62.8	68.3
24	When I listen to English I recognise when other people make mistakes	Assessment (monitoring)	61.5	57.6	60.2	63.3	63.5
32	When I have handed in my English test I think how I could have written it better	Assessment (evaluation)	67.7	76.3	75.7	62.5	55.1
25	When I have finished speaking I think how I could have said it better	Assessment (evaluation)	67.8	74.5	72.0	65.1	59.6
Assessment			**63.4**	**68.3**	**68.9**	**63.3**	**60.0**
Mean language proficiency			58	53	56	60	63
Mean frequency of use of the strategies			**68.4**	**70.2**	**70.2**	**68.6**	**65.6**

If we compare strategy use in all the anxiety groups we also notice that the higher the group's level of anxiety, the more frequent the use of strategies. This is not surprising if one remembers Lewis' (1996) interpretation of anxiety as a temporary state whose aim is to develop strategies to reduce or eliminate itself .

If we look at each group separately, each of them has a different pattern of strategy use:

- A4 scores highest in Assessment strategy questions (questions 20, 25, 26, 27, 32)

- A3 scores highest in Goal setting questions (16, 17, 18) and assessment strategy questions (28, 30 and 31)

- A2 scores highest only on one strategy, that of Planning (Question 19)

- A1 scores highest on Assessment strategy, both questions concern noticing mistakes in language use (Questions 23 and 24), which is logical, as their language proficiency is the highest.

This allows me to conclude that anxiety level not only induces the more frequent use of meta-cognitive strategies, but also decides (or depends on) the type of strategy we use.

The division of the sample into groups according to their level of **proficiency** produces more variety (see Table 46): for example, the lowest frequency of strategy use is for an assessment strategy in the highest proficiency group (P4), (Mean 39%, Question 28) and the highest is for a planning strategy in the second highest proficiency group (P3), (Mean 93.8%, Question 21).

Table 46 The frequency of use of meta-cognitive strategies according to the level of proficiency

	Strategy	Process	All %	P4	P3	P2	P1
16	When I begin studying English, I plan what I am going to do	Goal setting	47.2	46.3	47.8	47.7	45.5
17	I set goals for myself in language learning	Goal setting	74.3	71.2	76.6	73.6	74.0
18	I do everything to improve my knowledge of English	Goal setting	68.0	68.8	69.7	66.4	66.5
19	I think of the way how to continue studies after school	Goal setting	68.9	69.8	73.8	69.2	57.0
20	I try to find a way that is best for learning new words	Goal setting	60.9	55.6	64.6	57.9	64.0
Goal setting total			**63.9**	**62.3**	**66.5**	**67.4**	**62.96**
21	I try to understand when someone is speaking English	Planning	89.6	85.9	93.8	89.5	84.5
27	Before writing a composition I plan my work	Planning	58.3	62	57.0	57.4	59.0
29	When I am writing a test I try to concentrate	Planning	86.1	83.4	87.8	87.4	82.5
Planning total			**78**	**77.1**	**79.5**	**78.1**	**75.3**
23	When I speak English I notice when I make mistakes	Assessment	63.0	67.3	64.8	61.5	57.5
24	When I listen to English I recognise when other people make mistakes	Assessment	61.5	68.8	63.7	57.2	57.5
25	When I have finished speaking I think how I could have said it better	Assessment	67.8	60.5	69.2	70.3	67.5
26	Before writing a test I try to find out how it is going to be scored	Assessment	58.1	56.1	59.3	58.7	56.5
28	Before doing an English assignment I think whether I have enough English to do it	Assessment	53.7	39	54.0	57.4	61.0

30	When I am taking an English test I know how much time has gone by	Assessment	80.1	78.5	80.2	81.3	79.0
31	Before handing in a test I check my work	Assessment	71.0	69.3	74.3	70.5	66.5
32	When I have handed in my English test I think how I could have written it better	Assessment	67.7	52.7	67.4	76.7	66.2
Assessment total			**65.4**	**64.0**	**66.6**	**66.7**	**64.0**
Mean frequency of use of strategies			**64**	**64.7**	**69**	**67.7**	**62.0**
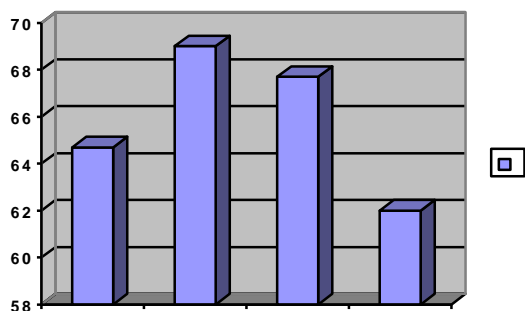
Figure 27 Self-reported frequency of strategy use in proficiency groups [P4 (1st bar), P3 (2nd bar), P2 (3rd bar) and P1 (4th bar)]

Both the weakest and the strongest language users say that they use fewer strategies than the two groups in between. The lowest proficiency group (P1) seems to have been the least active in using any strategies (Mean 62%), whereas the highest proficiency group used the fewest strategies in Assessment (64%) and used only slightly more Goal setting strategies (66.3%) than the weakest proficiency group. More frequent use of strategies can be seen in the medium proficiency groups in all the three meta-cognitive strategy areas: Goal setting, Planning and Assessment. This could be explained by the fact that the mean proficiency level of the highest anxiety group is 53% (see Table 34).

10.2.5.5.2 Correlations between the use of strategy and language performance

Strange as it may sound the interaction between the meta-cognitive strategies and the performance of the test takers is not always positive (see Table 47). As it has already been pointed out by Purpura, not all the learning and test taking strategies that we are trying to teach our students have a positive impact on their performance (Purpura 1999) although one has to hold in mind that these are the results of self reported

questionnaire and that what I am measuring is in fact the test-takers' awareness of their use of strategies. It is possible that if the use of strategies has not reached the level of automaticity and needs conscious attention, it can interfere with the performance level.

There are only two strategies that have similar correlations for the whole sample: one of these is negative *Before doing an English assignment I think whether I have enough English to do it* (-.319, Question 28), and one positive *When I listen to English I recognise when other people make mistakes* (.254, Question 24). None of the other strategies have significant correlations with the overall level of performance for the whole sample.

If however, we look at the correlations between level of anxiety and proficiency level when the sample is split according to their level of **anxiety**, we see that correlations have a distinct pattern.

There is one assessment strategy (Question 28) that has a negative correlation in all the anxiety groups (except the highest anxiety group where it does not correlate (.058), and four that are positive for all the groups: two goal setting strategies (Questions 18 and 19) and two assessment strategies (Questions 23 and 24); all the other strategies have different correlations in different groups.

- The group with the highest anxiety level (A4) has four significant ($p<0.01$) positive correlations with level of proficiency:

19. *I think of ways in which to continue English studies after school*	(goal setting)	.398
21. *I try to understand when someone is speaking English*	(planning)	.497
29. *When I am taking a test I try to concentrate*	(planning).	.478
31. *Before handing in a test I check my work*	(assessment)	.302

Use of all the other strategies produces less strong, but still positive correlations. Thus this group not only uses most of the meta-cognitive strategies most frequently, but also more effectively than other groups.

- The next group (A3) has 10 negative correlations, which although very small and insignificant, still are persuasive because of their similarity. The only strategy that has a significant correlation with performance is examining the amount of English knowledge necessary before doing a task (-.316, Question 28). There are also four significant positive correlations in this group between performance and assessment strategies:

18. *I do everything to improve my knowledge of English*	.212
26.*Before taking a test I try to find out how it is going to be scored*	.253
23. *When I speak English I notice when I make mistakes*	.220
24. *When I listen to English I recognise when other people make mistakes*	.249

Thus, although this group said used strategies more frequently than the next two groups, the relationship between strategy use and performance level is ambiguous.

Table 47 Correlations between the use of strategy and performance level in different groups of anxiety

	Strategy	Process	All	A4	A3	A2	A1
16	When I begin studying English, I plan what I am going to do	Goal setting	n.s.	n.s.	n.s.	n.s.	.259
18	I do everything to improve my knowledge of English	Goal setting	n.s.	n.s.	.212	n.s.	n.s.
19	I think of the ways in which to continue studies after school	Goal setting	.182	.398	n.s.	n.s.	.372
21	I try to understand when someone is speaking English	Planning	n.s.	.497	n.s.	n.s.	n.s.
27	Before writing a composition I plan my work	Planning	n.s.	.193	n.s.	n.s.	.299
29	When I am taking a test I try to concentrate	Planning	n.s.	.478	n.s.	n.s.	.184
28	Before doing an English assignment I think whether I have enough English to do it	Assessment	-.319	n.s.	-.316	-.380	-.320
26	Before taking a test I try to find out how it is going to be scored	Assessment	n.s.	n.s.	.253	n.s.	n.s.
30	When I am taking an English test I know how much time has gone by	Assessment	n.s.	n.s.	n.s.	n.s.	.297
31	Before handing in a test I check my work	Assessment	n.s.	.302	n.s.	n.s.	.267
23	When I speak English I notice when I make mistakes	Assessment	n.s.	n.s.	.220	n.s.	.303
24	When I listen to English I recognise when other people make mistakes	Assessment	.254	n.s.	.249	n.s.	.457
	Number of Strategies that have positive correlation		2	5	4	0	8
	Number of Strategies that have negative correlation		1	0	1	1	1

- The third group (A2), has practically no correlation between meta-cognitive strategy use and language performance. The correlations are insignificant and mostly negative.

- The use of strategies was most efficient in the group A1: the positive correlations are stable throughout the different language skills in eight strategies, while in the group of high anxiety level there were only five.

To bring out the similarity of interaction between the use of strategy and the performance in different language skills I have included all the significant correlations (above .182, p <0.01) here for all the tasks for the low anxiety group A1 (see Table

48) and the high anxiety group A4 (see Table 49). The use of strategy effect is similar across all the skills, while the interaction between anxiety level and the performance level differs in different language skills.

Table 48 The significant correlations between the strategy use and proficiency level in low anxiety level group (A1)

Task/Question	Strategies in A1 group							
	16	19	22	23	24	30	31	28
Reading	**.286**	**.303**	**.303**	**.293**	**.384**	**.283**	**.272**	**-.321**
Matching	.294	.291	.441	.284	.298	.230	.257	-.268
Gap-filling	.351	.299	.237	.231	.353	.323	.218	-.316
Multiple-choice	n.s.	.217	.221	.260	.324	.180	.246	-.246
Listening	n.s.	**.296**	**.255**	**.198**	**.455**	**.224**	**.217**	**-.336**
Gap-filling	n.s.	.325	.330	n.s.	.454	.283	.240	-.243
Multiple-choice	n.s.	.213	n.s.	n.s.	.349	n.s.	.192	-.307
True/false	n.s.	.218	n.s.	n.s.	.389	.207	n.s.	-.395
Language Use	**.184**	**.372**	**.270**	**.262**	**.427**	**.278**	**.189**	**-.251**
Multiple choice	.251	.317	.293	.220	.293	.232	n.s.	n.s.
Editing	n.s.	.387	.290	.272	.433	.279	.187	-.278
Gap-filling	.201	.317	.161	.228	.443	.252	.244	-.246
Writing	**.301**	**.354**	**.231**	**.293**	**.402**	**.286**	n.s.	**-.291**
Email	.448	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	-.271
Recipe	.265	.438	.313	.322	.494	.381	.231	-.185
Diagram descr.	.381	.278	.215	.302	.330	.192	.204	-.207
Speaking	**.296**	**.379**	**.304**	**.339**	**.431**	**.285**	**.363**	**-.277**
Total	.259	.259	.299	.303	.457	.297	.267	-.320

The only exception in the positive correlations between strategy use and performance in group A1 is Writing Task 1, which does not correlate with any meta-cognitive strategy apart from Question 16 (*when I study English, I set goals for myself*). This can be explained by the fact that the task had a very low word limit and many students were penalised for surpassing it and as a result performance on this task had a low correlation (.484, see Table 30) with the performance on the whole examination. The correlation between the performance in this task and language performance on the whole test is only a fraction higher than that of the strategy of ability to plan (.448).

Table 49 The positive correlations between strategy use and proficiency level in high anxiety level group (A4)

	Strategies in group A4			
Task/Question	19	21	29	31
Reading	**.432**	**.497**	**.559**	**.310**
Matching	.188	.559	.323	.268
Gap-filling	.374	.439	.524	n.s.
Multiple-choice	.415	.234	.434	.417
Listening	**.327**	**.332**	**.442**	**.233**
Gap-filling	.337	.292	.430	.277
Multiple-choice	n.s.	.234	.387	.182
True/false	.314	.345	.266	n.s.
Language Use	n.s.	**.370**	**.219**	**.198**
Multiple choice	n.s.	.412	.214	n.s.
Editing	.216	n.s.	.241	.303
Gap-filling	n.s.	.352	n.s.	.204
Writing	**.334**	**.301**	**.265**	**.313**
Email	.199	.336	.339	.385
Recipe	n.s.	n.s.	.288	.209
Diagram descr.	.447	.302	.236	.297
Speaking	**.367**	**.492**	**.418**	**.174**
Total	.398	.497	.478	.302

If, however, the students are divided into groups according to their **proficiency** level, one can observe that the regular pattern disappears (see Table 50), and we can make only general observations:

- most of the strategies have either positive and negative correlations with language performance in the highest and lowest proficiency groups,

- the medium proficiency groups do not have any significant correlations [apart from Question 28 (*thinking whether one's English is sufficient for the task)* which correlates negatively with most of the skills and Question 24 (*When I listen to English I recognise when other people make mistakes)* which correlates positively].

Table 50 The correlation between the use of strategy and performance level in different groups of proficiency

	Strategy	Process	All	P4	P1
16	When I begin studying English, I plan what I am going to do	Goal setting	n.s.	.344 LU1 .325 R2	.344W1
17	I set goals for myself in language learning	Goal setting	n.s.	.353 R2 .350 LU3	-.461LU -.344L2 -.341R2
18	I do everything to improve my knowledge of English	Goal setting	n.s.	.314LU3 .292R3	.332W1 .249L3
19	I think of the ways in which to continue studies after school	Goal setting	.182	.316 W -.256LU1 -.276R	.392 TOT .242R1
20	I try to find a way that is best for learning new words	Goal setting	n.s.	.267 R2 -.254 SP	.210R2
21	I try to understand when someone is speaking English	Planning	n.s.	.394 R1 .244 L	-.200W2
27	Before writing a composition I plan my work	Planning	n.s.	-.293S2 .225W1	.451LU .396W1 .233SP .363R
29	When I am taking a test I try to concentrate	Planning	n.s.	.226 W1 -.213 S2	-.205L3
28	Before doing an English assignment I think whether I have enough English to do it	Assessment	-.319	-.487 W3 -.251S1	-.408R3 -.262L -.291L1 .327W
26	Before taking a test I try to find out how it is going to be scored	Assessment	n.s.	.276R1 -.219 W1 .270LU2	-.338L1 .206LU
30	When I am taking an English test I know how much time has gone by	Assessment	n.s.	.248 R1 .263 L1 .235LU1	-.395L2
31	Before handing in a test I check my work	Assessment	n.s.	.329R1 .200LU	.437W1 .305S
32	When I have handed in my English test I think how I could have written it better	Assessment	-.182	-.242 W -219R1	.373L3 .338 S .231W3 .306LU3
23	When I speak English I notice when I make mistakes	Assessment	n.s.	.340L .249 W2	-.273L2 .223W1
24	When I listen to English I recognise when other people make mistakes	Assessment	.254	.490L3 .312 R1 .296 LU	.342R .324L3 .336W1 .273LU
25	When I have finished speaking I think how I could have said it better	Assessment	n.s.	-.280 S -.210LU2	-.375L2 .230S
	Nr.of positive correlations			24	23
	Nr.of negative correlations			12	12

Purpura (1999) compared the use of meta-cognitive strategies for his high and low ability groups and found that:

- low proficiency group loadings for assessing the situation, self-evaluating and monitoring were higher than those of the high proficiency group's,

- planning was equivalent across the two groups (Pupura 1999, p.175).

It is difficult to compare his results with mine as I cannot compare the two tests that were used to stream the groups. In addition I have 4 groups while he has only 2. Nevertheless, the observation that the assessment loadings in the factor analyses were higher in the low proficiency group suggests that it is possible that the level of anxiety of the low proficiency group was higher during the test used for Purpura's study.

If we compare the two methods of analysing the interaction between meta-cognitive strategy and performance interaction, that is streaming the groups according to their level of anxiety versus proficiency, the results of this study suggest that the division according to the level of anxiety produces more interpretable results than when analysed according to the proficiency level groups. This once again suggests that there exists a close relationship between anxiety and meta-cognition and meta-cognition cannot be adequately explained without taking into account affect.

10.2.5.5.3 Correlations between meta-cognitive strategies and test anxiety

Test anxiety is usually expected to have a negative impact on language production, but use of strategies is expected to have positive correlations, therefore one would expect the correlations between the two to be negative or, if they represent two totally different phenomena, non-existent.

However, if we examine the interaction between strategy use and anxiety level for the whole sample, we find only two correlations that are not significant (-.126 (Question 23) and .128 (Question 24)), all the others are positive, with the highest correlations for the assessment strategies (see Table 51).

If the sample is split according to their level of **anxiety** we get a mixture of positive and negative correlations. For example, in the lowest anxiety group, the correlation between trying to concentrate during a test correlate negatively with anxiety in Reading Task 1 (-.411), positive with Listening task 3 (.253) and negative with the mean anxiety level (-.211). As a result, it is difficult to see any pattern and it is

impossible to say anything other than that there is an interaction between anxiety and strategy use. Therefore here I have not presented the table of correlations.

When we look at the correlations between anxiety level and the frequency of the use of strategies in the groups with different **proficiency** levels (see Table 51), the correlations become more regular: throughout a skill the same strategy constantly has either positive or negative correlations. I have included here strategies with the highest correlations in each proficiency group (the full table of intercorrelations see in Appendix 14).

Table 51 Some correlation between the use of strategies and anxiety level in each task in different proficiency groups

Gr.	Strat.	Anx. Mean
P4	16 GS	-.291
	21 PL	.369
	28 AS	.262
P3	18 GS	.270
	20 GS	.218
	25 AS	.253
	26 AS	.273
	27 PL	.277
P2	18 GS	.217
	23 AS	n.s.
	24 AS	-.199
	25 AS	.213
	28 AS	.225
	30 AS	.404
	32 AS	.305
P1	16 GS	.242
	18 GS	n.s.
	24 AS	.249
	30 AS	.303
	31 AS	.334
All	25 AS	.236
	28 AS	.216
	30 AS	.260
	32 AS	.312

Most of the significant correlations (two tailed, non directional, higher than .182, $p<0.01$) between the level of anxiety and strategy use are positive. There are two goal setting strategies (Questions 16 and 19) that have a negative overall correlation with mean anxiety. If we compare the definition of the goal setting strategy (see Purpura 1999) with the goal setting theory of Locke and Latham (1994) (see section 2.3.2.), who suggest that the level of performance is directly connected with the goals we

choose, their content, intensity and difficulty we can see the reason for the negative correlation. Question 16 *When I begin studying English, I plan what I am going to do* has negative correlations with anxiety level across all language skills in the high proficiency group. This perhaps helped the test-takers not only achieve their goal but also to overcome their anxiety.

However, the most proficient group (P4) also has a significant positive correlation between the planning strategy *I try to understand when I hear somebody speaking English* (.369, Question 21) and anxiety level and one can only wonder why students who would try to listen carefully when somebody speaks English, should also have higher listening anxiety (.346,). In addition, why does frequency of examining if one's knowledge of English is adequate for accomplishing the task (Question 28) correlate positively with anxiety level in Reading task 1 (.331)? There is of course another way of explaining the positive correlations, and this is suggested by the frequency of use of strategies (the higher the level of anxiety, the more strategies are activated and the more frequently they are used).

The groups with the medium proficiency level (P3 and P2) have the steadiest correlations between each strategy and level of anxiety. The correlation is higher for groupcP2 in the assessment strategies, the highest of those being the strategy relating to following the timing during the test (.404, Question 30).

For the group with the lowest proficiency level (P1) the assessment strategies have the highest positive correlation with anxiety (Language Use task 1 correlation between anxiety level and the contemplation of the task after it has been handed in even reaches .448, Question 32), which has already been suggested by the comparison of the functions of affect and assessment-strategies (see section 3.3.1.).

The positive correlations between the self-reported frequency of use of strategies and anxiety level might suggest that the high anxiety level is activating the use of a variety

of strategies and in some cases the use of strategies has managed to eliminate the level of anxiety, but in most cases the anxiety is stronger and keeps the strategies active.

Purpura (1999) says that *the use of meta-cognitive strategies alone does not appear to improve performance in testing contexts. Rather, these results show that "thinking" needs to work in concert with "actions" in order for learners to do well on language tests* (Purpura 1999, p.173). I suggest that thinking needs to work first with "emotions", or "affective schemata" (Bachman and Palmer 1996) as behaviour is governed by both thinking and feeling. This home truth has been supported by the results of this study in the following ways:

- the higher the person's level of test anxiety, the more often he or she uses meta-cognitive strategies

- meta-cognitive strategies have steady correlations with proficiency level if test-takers are divided into groups according to their level of anxiety instead of proficiency

- there is a positive correlation between test anxiety and the use of many meta-cognitive strategies by different proficiency levels.

10.2.6 Findings of the year 2000 questionnaire study

The year 2000 questionnaire study

1. measured **level of anxiety** of the test-takers and found the existence of relationship between anxiety and proficiency: the higher the level of proficiency, the lower the level of anxiety

2. examined the **causes of anxiety** and found that the highest anxiety group mentions general state anxiety more often than did other groups,

3. compared **classroom anxiety** with test anxiety and found that a) the group with highest test anxiety also reported the highest classroom anxiety, b) attempts to control test anxiety by discussions in the class only helped the most proficient

group and c) correlations between classroom anxiety and test anxiety differed in the different proficiency groups.

4. measured the frequency of use of **meta-cognitive strategies** and found that the higher the level of anxiety, the more strategies were used.

This study also suggested a way to discover regularities in correlation tables

- if the sample was split in groups according to their level of **anxiety**, it was easier to examine the relationship a) between meta-cognitive strategies and level of proficiency and b) between levels of test and classroom anxiety

- if the groups were split according to their level of **proficiency**, it was easier to examine the relationship a) between meta-cognition and level of anxiety and b) between test anxiety and language proficiency.

Although the correlations suggested interaction between language performance, test anxiety and meta-cognitive strategies, they could not answer the question of causality. Therefore I will now examine the same data using the structural equation method to produce a model of interaction between cognition and anxiety in language use.

10.3 Models of interaction between meta-cognition, anxiety and language use

Language performance as a construct by itself is complicated enough to comprehend, but to understand how it interacts with different variables one needs some kind of visual model. The Structural Equation Modelling (SEM) approach offered by Bentler's (1998) EQS version for Windows is ideal for this because

1. it allows one to create a diagram depicting relationships between different variables using path diagrams,

2. it evaluates how well the theoretical model fits the practical data and also provides measures of probability

3. it assesses the interaction between the different variables that influence language performance (using covariance coefficients)

4. it detects the direction of any influence (using regression path analyses).

Structural equation modelling is used to test theoretical models using empirical data to test relationships proposed by the existing theories. However, there are several preconditions that have to be satisfied before we start using the approach and several steps to be taken to establish the reliability of the models created.

10.3.1 The preconditions

Raykov and Marcoulides (2000) insist that when using SEM we have to take the following steps:

1. Identification of parametres: models have to be identified, that is there must be enough empirical information to allow its unique estimation. If a model contains even one unidentified parameter, the model cannot be relied upon, even though other parts of the model represent a useful approximation of the studied phenomenon. To check this condition Raykov and Marcoulides (2000) suggest that we count the number of parametres and subtract the value from the number of nonredundant elements in the sample covariance matrix ($p(p+1)/2$, where p is the number of observed variables. The difference is called degrees of freedom (df). If the difference is positive, the necessary conditions for model identification are fulfilled, if it is 0, we call a model saturated or just identified. If the difference is negative, the model is unidentified and its data are not reliable as there is not enough empirical information to allow model's unique estimation.

2. Researchers should consider for study only those models that are attached to some 'substantive considerations and offer credible means of data description and explanation. Raykov and Marcoulides (2000) propose that we have several theories explaining relationship between the different parametres that have been empirically

measured then SEM (using the model fit index) will help us establish which of the theories depict the relationship most reliably.

3. The most substantial precondition of reliability of models produced by SEM (as I guess using any other method) is that the variables have been reliably and validly assessed.

10.3.2 The process of fitting the model

Raykov and Marcoulides (2000) describe SEM as a method for quantification and testing theories. It is based on comparing the proposed SEM model to the observed data. This is established by comparing the reproduced covariance matrix to the observed sample covariance matrix until the optimal. The process of fitting a structural equation model is like solving a system of equations: on the one hand there are the observed subsequent numerical entries of the observed sample covariance matrix, but on the other there is the corresponding expression of model parameters defined in the in its own matrix. The process starts with initial estimates of the parametres and ends (or converges) when the fit function does not change by more that a very small amount (.000001) and there is no further improvement in the distance between the two matrices (model and observed sample). The smaller the difference, the better the fit of the model. If the difference is large, then:

1. the proposed model is defficient

2. the data may not be good.

The numerical values obtained at the final iteration are presented as the required estimates of the model parameters. If, however, the iterative processes do not converge, the model is not appropriate for representing the observed data relationships.

10.3.3 Methods of establishing reliability of EQS models

Raykov and Marcoulides pay special attention to warn the researchers against putting too much trust in the models produced and suggest several ways of establishing reliability of the models (sampling the population, checking the data and standard error measurements).

10.3.3.1   Sampling

Raykov and Marcoulide consider the aspects of choice of population that can affect the credibility of the models. Although they admit that there is a common agreement that the larger the population the better (Bentler 1995), they also suggest that there is no common agreement of what is meant by large. Their view is that the sample size should 10 times larger than the free parametres size. If it is smaller it is advisable to use Robust method of parametre estimation.

Raykov and Marcoulides (2000) also suggest that if the same model can be replicated in new samples from the same population; this would greatly enhance the value of the proposed model, although the EQS program also provides a standard error measurement which shows how stable the parametres solution is if a repeated sampling were carried out of the same population.

10.3.3.2   Goodness of fit

The test statistic of the goodness of fit of the model tests the null hypothesis that the covariance matrix of the model fits the covariance matrix of the data observed perfectly. Raykov and Marcoulides (2000) define test statistic chi-square value as [T=(N-1)Fmin], where N is the sample size, and (Fmin) *is the computed minimal value of the fit function for the parameter estimation method used* (Raykov and Marcoulides 2000, p.36).

When the model fit is obtained, the SEM program will judge it in relation to models degrees of freedom and produce its associated p value. Usually the p value is considered as significant at 0.5 and when it is achieved the it is considered that the model is capable of reproducing the analysed matrix of variable relationship indices. If, however, the p value is smaller than 0.5 we should consider rejecting the model(Raykov and Marcoulides 2000, p.36).

Raykov and Marcoulides (2000) present a further word of warning against trusting SEM models: even if we have found a model with a perfect fit, we cannot say that we have found the true model that has generated the analysed data. This can be explained by the fact that SEM differs from classical modelling approaches where we are interested in rejecting the null hypothesis. In SEM we are interested in retaining the proposed model whose validity is the essence of null hypothesis, therefore here we are not interested in rejecting null hypothesis. This however implies that we have to accept a model that proposes a theory that does not have to be the only possible explanation of the parametres measured. Thus even if the model fits the data set well, there can be a plethora of other models that fit the data even better.

10.3.3.3   Substantive considerations

Thus according to the authors the basis for trusting the model is sound body of knowledge about the studied phenomenon. Raykov and Malcoulides suggest that all the models should be conceptualised according to the latest knowledge about the phenomenon under consideration which can be found when carrying out extensive study of the pertinent literature. Then, by producing the models representing the older theories and the latest ones, we can compare which produce the best fit. This agrees with Kunnan (1995) who also suggests that a reasonable explanation is more important than a perfect fit (see Chapter 10).

I will first present Purpura's (1999) model investigating the role of meta-cognition in language use and then move on to my own models investigating the interaction between meta-cognition and anxiety in language use.

10.3.4 Previous studies on impact of meta-cognitive strategies

Purpura uses Structural Equation Modelling because it is useful for estimating, specifying and testing hypothesized interrelationships among a set of meaningful variables, both observed (measured data) and latent variables (factors). All in all the program uses four types of variables:

1. the measured variables (V) allow us to develop a standard path analyses or simultaneous equation model with one way or two way interaction between them. They are independent if a two way arrow points to them or dependent if a one way arrow points to them.

2. the hypothetical factor variables (F) used for the common factors proposed by the exploratory factor analyses allow us to develop further factor analyses or a measurement model

3. the program also supplies the model with the residual variables that are not catered for by the measured variables V (shown in the model as E variables)

4. the residuals of the factor variables F are showed in the model as D variables.

Purpura explains that the method uses the following steps: the interrelationships between the observed variables are measured to establish latent variables and develop a measurement model. Then, on the basis of previous research, interrelationships between latent variables are hypothesized and tested until a good model fit is achieved. Thus the structural model is developed. Any discrepancy between the hypothesized relationship and the observed data is shown by the residual. The final step is the analyses of both the models (measured and structural) to develop a full latent variable model.

The main parameters on which Structural Equation Modelling is based are:

1. the path coefficients of factor loadings

2. the variances of the independent variables

3. the covariances of the independent variables.

Purpura used the maximum likelihood estimation method because most of his data met the distributional assumption of multivariate normality and the maximum likelihood robust method is for data that do not have a normal distribution. The robust maximum likelihood estimation method provides a robust chi-square statistic and robust standard errors that correct the normality in large samples (Purpura 1999, p.64).

To assess the goodness of the fit of the model Purpura used several indices:

1. the chi-square statistic was used to measure the overall goodness of fit of the specified model against the unconstrained or null model; this shows the distance between the sample covariance matrix and the fitted covariance matrix

2. because of the scaled non-normality of his data Purpura also used the Satorra-Bantler scaled statistic to provide a scaling correction for the chi-square statistic

3. the comparative fit index was used as a primary index to compare the hypothesized model with the null model.

At first Purpura hypothesized the meta-cognitive strategies as a two-factor model consisting of two latent variables: online assessment and post assessment processes. The analyses showed that this was an *excellent representation of data* (Purpura 1999, p.102) as the comparative fit of the model (CFI) was 0.999 and the chi-square value was 1.750 at 1 degree of freedom (p>0.05). However, the inter-factor correlation of this model was greater than 1, suggesting that meta-cognitive strategy use, as operationalised in Purpura's study was in fact a uni-dimensional construct. As a result the model was respecified as a single latent variable represented by 4 observed

variables; this improved the CFI to 1.0 and chi-square to 1.960 with 2 degrees of freedom.

10.3.5 Interaction between meta-cognition and anxiety

I used the SEM program to investigate the interaction between anxiety and meta-cognition and their combined effect on language performance (see Appendix 15 for the program control printout).

To begin with I treated foreign language test anxiety as a state anxiety and compared the effect of foreign language test anxiety with foreign language classroom anxiety and I found that the effect of classroom anxiety was much stronger than that of foreign language test anxiety. I explained this by the fact that classroom anxiety was a form of trait anxiety and decided to compare classroom anxiety with test anxiety as it was measured in 1999 using Spielberger et al's (1978) Test Anxiety Inventory (TAI). Test anxiety according to SEM turned out to have a stronger impact on language performance than did foreign language test anxiety as measured by my questionnaires in 1999 and 2000. As the questions concerning classroom anxiety (Questions 14 and 15) in year 2000 and test anxiety (Questions 13-31) in 1999 addressed anxiety as a habitual reaction, as opposed to Questions 1-12, which addressed the test-takers' reactions to the current test tasks, I introduced the distinction between **state** (foreign language test) and **trait** (classroom and test) anxiety.

As I had not measured Test anxiety in the year 2000, I could not have a model that would incorporate both test and classroom anxiety, therefore I compared them using my foreign language test state anxiety scale (Questions 1-12) as "anchor" items as they were the same in the questionnaires in 1999 and 2000.

 The next step was to examine the interaction between the state and trait anxieties and language proficiency, and finally I developed models that contained language

proficiency, anxieties and meta-cognitive variables. These examined the interaction between different factors during the language test.

All the models are general linear structural equation models as they all contain all four types of variables (measured, hypothetical and residuals of measured and hypothetical variables). All the models were obtained by using the maximum likelihood method as the data were normally distributed. I used two-way arrows to develop oblique factor models and to examine interaction between different factors and orthogonal factor models with one-way arrows to examine the causality of the interaction (Bentler 1998).

I allowed the comparative fit index of the model to fluctuate between 0.9 and .97, as Bentler suggests that values of less than .8 are inadequate and values above .9 are acceptable. I did not focus on achieving the model fit index of 1.0, as this would lead to the gradual reduction of more and more variables as it happened in Purpura's (1999) research. Instead I followed Kunnan's (1995) proposition that reasonable explanation is more important than ideal fit.

10.3.5.1   Exploratory factor analyses

Before modelling the interaction between the different variables I carried out exploratory factor analyses for the data of years 1999 and 2000 (Appendices 11 and 12). I will present here only the results of the factor analysis of the year 2000 data as the results from the two studies were similar.

The direct oblimin solution of year 2000 Questionnaire (see section 10.2) data converged after 44 iterations and yielded 11 factors with an eigen-value 1.0 and above. In Table 52 I list the factors that I chose for the equation modelling and all the data sets that scored on that factor. The number next to the name of the data shows how well they scored on that particular factor. The last line of Table 52 shows where

the data came from. For the whole printout of the factor analyses see Appendices 11 and 12.

Performance on the test tasks all scored on the same Factor 1 except for Writing task1 (for the reasons for this see the discussion in section 10.1). Questions 1-12 in the questionnaire asking test-takers to state the level of anxiety in each task, however, scored on separate factors (Language use anxiety, Writing anxiety, Listening anxiety and Reading anxiety). It is difficult to accept the idea that the 4 language skills are more similar than the anxieties experienced during the use of these skills, but the results of factor analyses support Cheng et al's (1999) suggestion that there are four different language skill anxieties (in my research there is also the fifth, Language Use test anxiety).

The areas of meta-cognitive strategies: goal setting (Factor 2), assessment strategies (Factor 6) and planning strategies (Factor 7) also scored on separate factors and did not support Purpura's (1999) research finding that all the three areas in fact represent the same assessment strategy. Factors 9 and 10 consisted of different variables that were difficult to identify as a separate factor (see Appendix 12).

Factor 11 (classroom anxiety) was formed by Questions 14 (being more worried by foreign language classes than others) and Question 15 (being afraid that classmates will laugh). Although one of the questions asked specifically about speaking, the fact that it was about the classroom situation as opposed to the test, kept the classroom speaking anxiety in the classroom anxiety factor and did not add it to speaking performance (which formed part of Factor 1).

Table 52 Results of the exploratory factor analyses for year 2000 data

Factor 1 Language performance (Langprof)*	Factor 2 Language Use anxiety (Luanx)*	Factor 3 Goal setting (Goalset)*	Factor 4 Writing anxiety (Writanx)*	Factor 5 Listening anxiety (Listanx)*	Factor 6 Assessment strategies (Assess-ment)*	Factor 7 Planning strategies (Plan-ning)*	Factor 8 Reading anxiety (Readanx)*	Factor 11 Classroom anxiety (Classanx)*
L1 .83	Luanx1 .79						Ranx1 .67	
L2 .79	Luanx2 .77						Ranx2 .71	
L3 .77	Luanx3 .71						Ranx3 .60	
L .90		Q16 .47						Q14 ..61
LU1 .76		Q17 .72						Q15 ..56
LU2 .74		Q18 .55						
LU3 .76			Wanx1 .67					
LU .86			Wanx2 .70					
Sp1 .71			Wanx3 .75					
Sp2 .70				Lanx1 .81				
Sp. .78				Lanx2 .77				
W1 -.0				Lanx3 .57				
W2 .54					Q22 .32			
W3 .41					Q23 .68			
W .58					Q24 .67			
R1 .46					Q25 .39			
R2 .78						Q21 .46		
R3 .70						Q29 .76		
R .82						Q30 .59		
Language test	Questions 7, 8 and 9	Questions 16, 17, 18	Questions 10, 11, 12	Question s 4, 5 , and 6	Questions 22,23, 24, 25	Questions 21, 29, 30	Questions 1, 2 and 3	Questions 14 and 15

*the factor's name in the models

10.3.5.2 Models of foreign language test anxiety

10.3.5.2.1 Foreign language test state anxiety

Exploratory factor analyses of my data (see Table 52) supported the hypothesis proposed by Saito, Garza and Horwitz (1999) that we should be talking about foreign language reading, writing and listening and speaking anxiety instead of foreign language anxiety as a whole. Language use anxiety also appeared in my data as a separate factor. Therefore the data of the year 2000 questionnaire (Questions 1-12) are represented in Model 1 (Figure 26) in the following way: foreign language test state anxiety consists of four factors (writing anxiety, listening anxiety, language use anxiety and reading anxiety). Each language skill (or language element) (except speaking anxiety, which I did not measure with the questionnaire) anxiety is represented by the measurement of the level of anxiety in each test task. Each forms an independent factor, which is connected by two-way arrows to all the other anxieties.
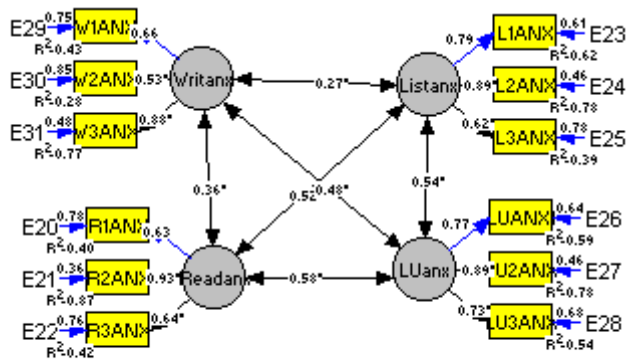
Chi sq.=175.64 P=0.00 CFI=0.90 RMSEA=0.10

Figure 28 Model 1 Foreign language test state anxiety

The Chi square of the model is 175.64, which is high enough to produce a P of <0.001 and to suggest that the probability is satisfactory. The fit of the model is 0.90, which also suggests that it is good enough to explain the interrelationships between the data variables.

The standardised solution (the confirmatory factor analysis) of the model (see Table 53) supports the exploratory factor analyses: each factor is well represented by the measured variables (see Table 52). The loading of the variables vary between .53 (anxiety level during Writing task 2) and .93 (anxiety level during Reading task 2).

Table 53 Standardised solution of Model 1: Language test anxiety

Name of the Variable	Anxiety experienced during	Standardised solution of the covariance between the variable and the factor representing it
W1ANX	Writing task 1	.66
W2ANX	Writing task 2	.53
W3ANX	Writing task 3	.88
R1ANX	Reading task 1	.63
R2ANX	Reading task 2	.93
R3ANX	Reading task 3	.64
L1ANX	Listening task 1	.79
L2ANX	Listening task 2	.89
L3ANX	Listening task 3	.62
LU1ANX	Language use task 1	.77
LU2ANX	Language use task 2	.89
LU3ANX	Language use task 3	.73

The covariances between the four factors representing the four language skill/element anxieties (Reading, Listening, Language Use and Writing) suggest that the interaction

between the four language skill anxieties is significant. It is highest between reading, listening and language use anxiety (.54). It is smallest between writing and listening anxiety (.27) All the correlations are positive.

Table 54 Largest residuals of Model 1

Variable 1	Variable 2	Covariance
L3ANX	R3ANX	.29
LU3ANX	L3ANX	.24
W1ANX	R1ANX	.18
LU3ANX	R3ANX	.16
W3 ANX	L3ANX	.15
R3ANX	R1ANX	-.12
LU1ANX	R3ANX	.12
W1NAX	LUANX1	.12
LU2ANX	L3ANX	.11

The Comparative fit index of Model 1 is .90 as the residuals between the anxiety experienced in different language tasks suggest connections that are not accounted for in the model: Table 54 depicts the largest residuals of Model 1 and the numbers there explain why the CFI is only .90: the level of anxiety characterising the tasks of the same level of difficulty have an interconnection that the model does not account for. This could be explained either by the fact that test-takers knew that each task 3 was going to be the most difficult and reacted in a similar manner in all three tasks. Thus, if the variables representing task 3 anxieties in listening and reading anxiety were connected, the residuals would decrease by .29 and the CFI would accordingly increase. I decided, however, to leave the model as it was (with CFI=.90) as the residuals were low and the interconnections would make the model difficult to understand.

10.3.5.2.2 Interaction between state and trait anxiety

In this section I will examine two trait anxieties (foreign language classroom anxiety and general test anxiety) and one state anxiety (foreign language test anxiety) as well as interaction between state and trait anxiety.

10.3.5.2.2.1 Classroom anxiety

Model 2 (Figure 27) was also developed using the data of year 2000. It represents the interaction between foreign language test anxiety (Questions 1-12) and classroom anxiety, which is a trait anxiety as Questions 14 (*Foreign language classes me worry more than others even when I am well prepared*) and Question 15 *(I am afraid that my classmates will laugh if I speak in a foreign language)* refer to attitudes of the test-takers that have become well established and represent traits.

The exploratory factor analysis suggested that classroom anxiety was a separate factor, and therefore I included Questions 14 and 15 as a separate factor. In the factor analysis factor loadings were not very high: .61 for Question 14 and 0.56 for Question 15. The standardised solution (confirmatory factor analysis) also suggests that it is Question 14 that is loading more strongly: .67. Question 15 has a loading of .43 and as it addresses the question of speaking explicitly, it could be in fact considered as interaction between classroom and speaking anxiety.

Figure 29 Model 2, Interaction between test and classroom foreign language anxiety
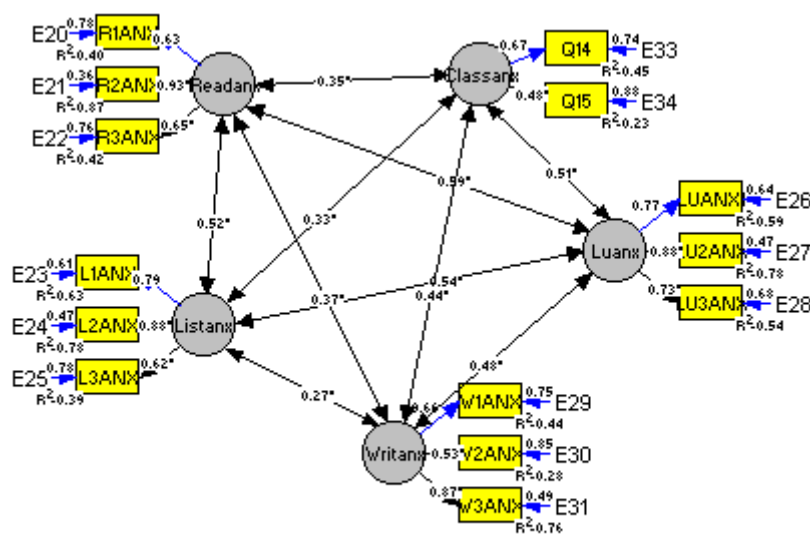
Chi sq.=199.60 P=0.00 CFI=0.90 RMSEA=0.09

Table 55 Interaction between classroom and foreign language test anxiety

Abbreviation	Full name of the factor	Interaction with classroom anxiety (Classanx)
Writanx	Writing anxiety	.44
Readanx	Reading anxiety	.35
Listanx	Listening anxiety	.33
Luanx	Language Use anxiety	.51

Table 55 suggests that all the language skill/element anxieties have positive and significant correlations with classroom anxiety, but interactions between language use anxiety and classroom anxiety is the strongest (.51). The comparative fit index of this model is acceptable: .90 and the probability is also satisfactory (P<0.001). My data suggest that classroom anxiety as a trait is activated by the language test situation and interacts during the Year 12 examination with the separate language skill anxieties.

The detailed model, although explicit, does not allow us to generalise about either the strength of the interaction between foreign language test state and classroom anxiety or its causality. Therefore I produced an orthogonal factor model that contained only two factors: the means of all language skill/element anxieties as the factor of foreign language test state anxiety (Flanx) and classroom anxiety (Classanx) factor (see Figure 29). The arrow that connects them is one-way and the figure that shows the interaction is regression path coefficient that is used to represent a one-way causal flow in a system (Hatch and Lazaraton 1991).

This model fits the data well (although p=0.06 which can be explained by the underspecification of the model) and suggests that it is classroom anxiety that is responsible for causing test anxiety and not the other way round as was suggested by Horwitz (1982). The regression path coefficient (.58) is high enough to suggest the considerable influence of classroom anxiety on causing foreign language test anxiety.

Figure 30 Model 3 Interaction between classroom and foreign language test state anxiety
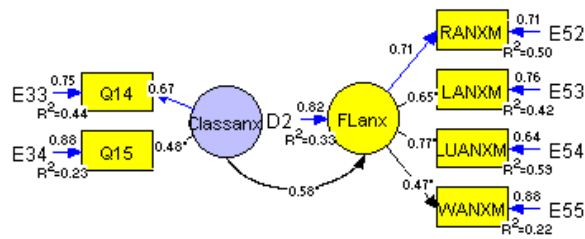
Chi sq.=14.98 P=0.06
CFI=0.97 RMSEA=0.06

10.3.5.2.2.2 Test anxiety

The factor analyses and the models suggest that during the test, state anxieties of the foreign language skills are affected by classroom anxiety, which has itself become a trait. This, however, is not the only trait anxiety that is activated by the test. Spielberger's (1982) and Sarason's (1980) research suggest that apart from subject specific anxieties there is also the test anxiety that is evoked by every test situation and which can also become a trait. The Questionnaire in 1999 (questions 13-31) addressed test anxiety as a trait. Therefore I will use the 1999 data to see how test anxiety as a trait interacts with foreign language test state anxieties (for the description of the questionnaire see section 10.1).

Test anxiety ("Testanx" in the model) was represented by Question 14 (*In the final exams I feel unsafe and am easily upset*), Question 20 (*I am worried even if I am well prepared*) and Question 21 (*I am worried before receiving the results of the test*) which according to Spielberger were subsumed by the emotionality factor. As the cognition factor did not cluster together in the exploratory factor analyses (see Appendix 11) of my data (see the definition of emotionality and cognition in section 4.1.6) I did not include it in the model.

Model 4 (Figure 29) below shows the interaction between foreign language anxiety during the test and test anxiety as a trait. The model fits the data well (CFI is 0.93 and P<0.001) and reveals once more that trait anxiety interacts actively with foreign language state anxiety.

Figure 31 Model 4 Interaction between test trait anxiety and foreign language test state anxieties
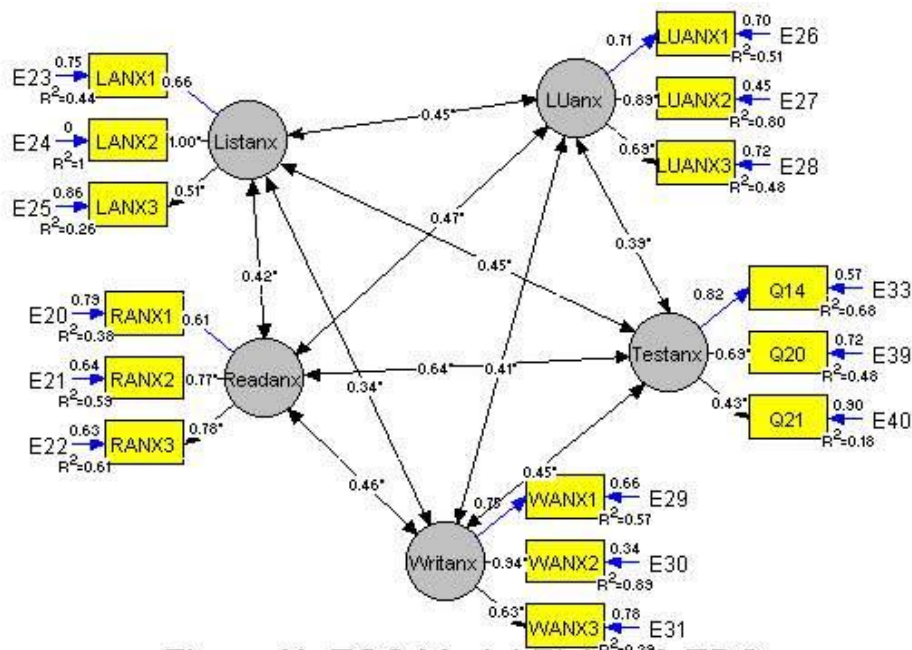
Figure X: EQS Model FLANX9.EDS
Chi sq.=136.94 P=0.00 CFI=0.93 RMSEA=0.07

Table 56 Comparison of the interaction of foreign language state anxiety with classroom and test trait anxiety

Abbreviation	Name of the factor	Interaction with classroom anxiety (year 2000 data)	Interaction with Test anxiety (year 1999data)
WANX	Writing anxiety	.44	.45
RANX	Reading anxiety	.35	.64
LANX	Listening anxiety	.33	.45
LUANX	Language use anxiety	.51	.39

In Table 56 we can see that test anxiety as a trait interacts with foreign language test state anxiety more actively than classroom anxiety in all skills apart from language use anxiety which seems to have a special relationship with classroom anxiety. Test anxiety as a trait interaction with reading is the highest (.64) and with language use

the lowest (.39) which to my mind depicts the difference between classroom anxiety and test anxiety (Gardner and Horwitz (see section 4.3.1) define foreign language anxiety as 'situational' and evidently the first 12 questions of my questionnaire addressed the test situation more than classroom situation).

This however does not answer the question about which of the anxieties is the cause and which the effect. Model 5 (Figure 30) was developed on the basis of two factors: Test anxiety as an independent factor and foreign language test anxiety (represented by all 4 language skill anxieties) as the dependent factor. These are connected by a one-way arrow pointing at foreign language test state anxiety. The fit of the model is .93 which is good and although Chi square is only 58.24, P is still 0.01, which allows us to give credibility to the model.

Figure 32 Model 5 Causal interaction between test trait anxiety and foreign language test state anxiety
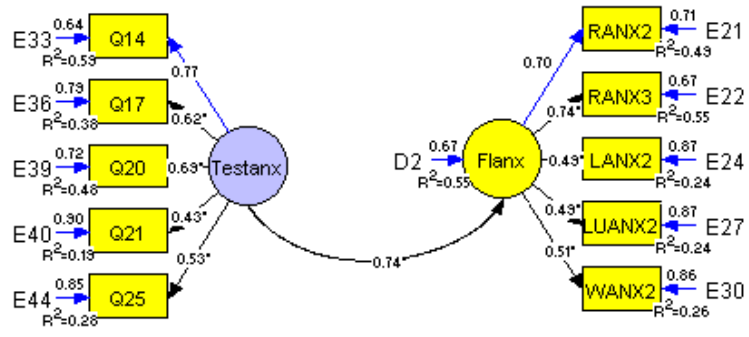
Chi sq.=58.24 P=0.01 CFI=0.93 RMSEA=0.07

The regression path coefficient showing the influence of test anxiety as a trait on foreign language test anxiety as a state is .74, which is higher than the regression path coefficient that showed the influence of classroom anxiety on foreign language test state anxiety (.58). Although the data were taken from two different populations, the comparability of the tests and the questionnaires of the foreign language anxiety allow me to conclude that

1. it is the combined trait anxiety influence (in the form of classroom and test anxiety) that is causing the state anxiety during the foreign language test

2. the influence of the test trait anxiety (.74) is stronger than classroom anxiety (.58) in creating foreign language test state anxiety.

10.3.5.3   Interaction between foreign language performance and anxiety

Having found that my research supports Eysenck's (1992) proposition that anxiety is a multi-dimensional concept, which in this study consists of test, classroom and language skill/element anxieties, the next question to be addressed is how it interacts with language performance during a language test.

10.3.5.3.1 Interaction between foreign language performance, foreign language state anxiety and classroom anxiety

Following Gardner and MacIntyre's (1991) suggestion that interaction between anxiety and language performance is bi-directional, I first produced the so-called oblique factor model where all the factors were connected with two-way arrows allowing bi-directional interaction between the factors (see Figure 31 Model 6). The test anxiety factor consisted of the four language skill anxiety variables that according to the model's standardised solution were well represented, the loadings ranging from .48 (writing anxiety) to .78 (language use anxiety). Language use anxiety was represented by all five skill measured variables (the means of all the language skill test anxiety levels). The language use anxiety factor also represents the variables well (both the language use and reading anxiety variables in the standardised solution load as high as .90) and the writing anxiety factor although loading the lowest (.78), is still high.

Figure 33 Model 6 Interaction between foreign language anxiety, classroom anxiety and language performance
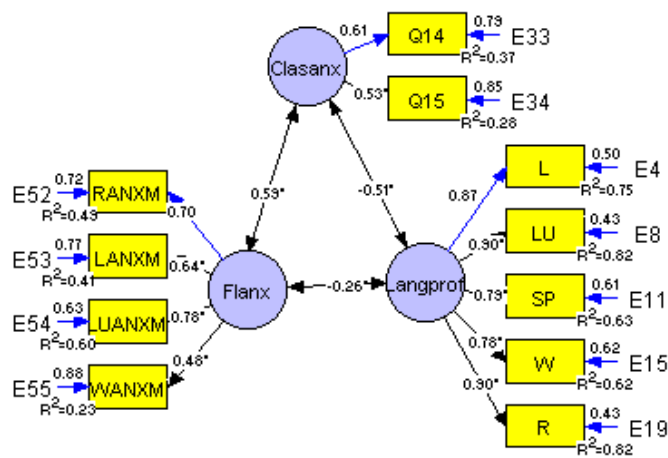
Chi sq.=93.02 P=0.00
CFI=0.96 RMSEA=0.07

The comparative fit index of the model is high: (.96) and the chi-square is high enough (93.02) to ensure P<0.001. This allows us to say that the theoretical model that language performance result is dependent on classroom (trait) and test (state) anxiety fits the data well. The model suggests that the interaction between language performance and anxiety is negative: the correlation between language performance and foreign language anxiety is -.26, but the correlation between language performance and classroom anxiety is nearly twice as big: -.51. The correlation between test anxiety and classroom anxiety is positive: (.59) and this suggests that the two are combined in their interaction with language performance.

This, however, does not answer the question of what causes the interaction. Is it language performance that causes the test state anxiety or test anxiety that affects the language performance? To answer the question of causality, I developed an orthogonal factor model, where the two anxiety factors, which affect language

performance, are interconnected by a two-way arrow, but one-way arrows lead to the language proficiency factor (Figure 32, Model 7).

Although this model fits the data as well as the previous one (CFI is also .96), the measurements of the interaction are different. Classroom anxiety effect has increased to (-.55), but the influence of foreign language test state anxiety (Flanx) on language performance is only (.06) and has become insignificant. If we compare it to the covariance -.26 between language performance and foreign language test anxiety in Model 6, it seems that the influence of language proficiency on foreign language anxiety is greater than that of foreign language anxiety on language proficiency.

Figure 34 Model 7 Influence of classroom anxiety and foreign language anxiety on language performance
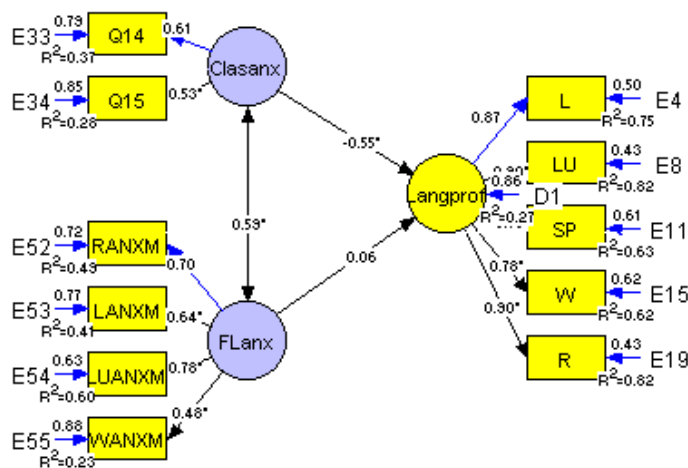
Chi sq.=93.02 P=0.00
CFI=0.96 RMSEA=0.07

This suggests that

1. it is classroom anxiety that causes significant performance deterioration, while foreign language test state anxiety does not cause significant test performance deterioration

2. language proficiency level acts as a major cause of foreign language test state anxiety.

Figure X: EQS Model C:\EQS\CLASANX.EDS
Chi sa.=307.80 P=0.00 CFI=0.93 RMSEA=0.08
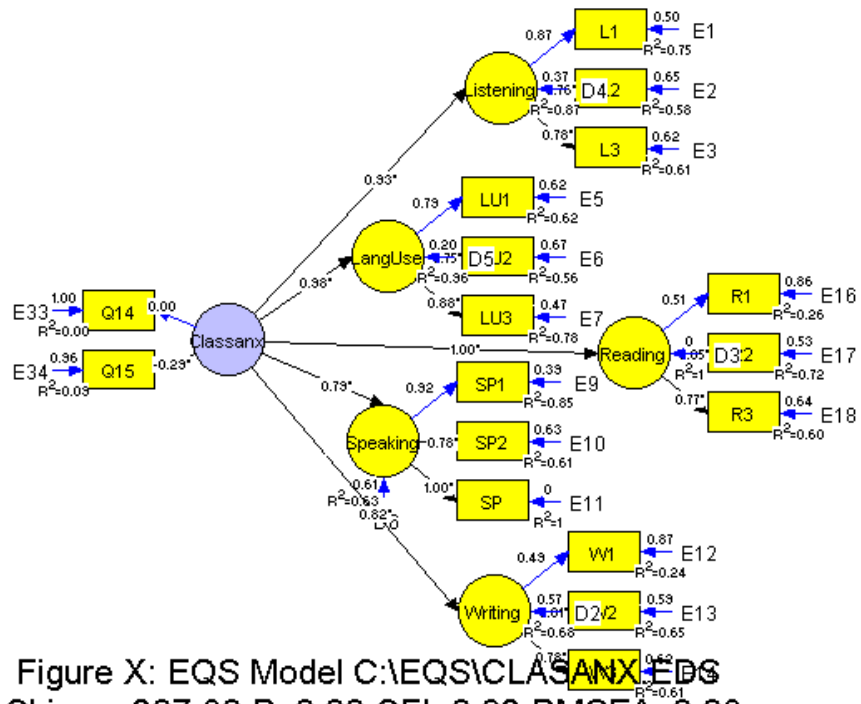
Figure 35 Model 7A, Effect of Classroom anxiety on language skill performance

Figure 33 (Model 7A) shows the effect of classroom anxiety on test performance on different language skills. The effect is significant on all skills: it is greatest on Reading (1.00) and Language Use (.98).

10.3.5.3.2 Interaction between test anxiety and language performance

Model 8 (Figure 34) uses the data of 1999 and depicts the interaction between test anxiety and language performance: we can see that the interaction between both the anxieties (foreign language test state anxiety and test anxiety) is positive, and strong: (.74), and that the interaction between anxieties and language performance is negative [(-.27) and (-.24)].
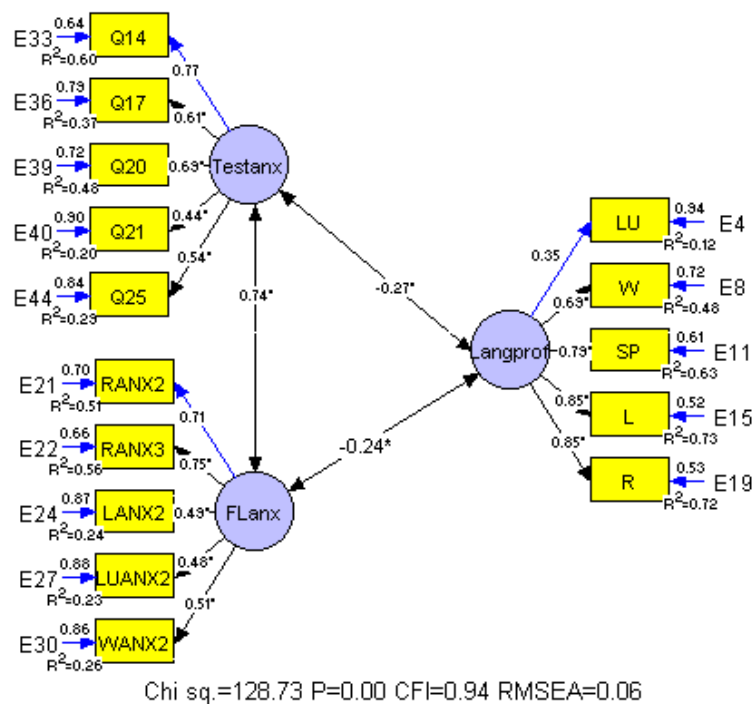
Figure 36 Model 8, Interaction between test anxiety and language performance

If we look at the cause of the influence (see Model 9 Figure 35), we can see that the test *trait* anxiety (Testanx) effect on language performance is twice as strong as that of the foreign language test *state* anxiety effect (-.10); it is evidently language proficiency that partly causes test state anxiety. If compared to classroom anxiety effect (.58), they are both (test trait and foreign language test state anxiety) insignificant (see Model 7).

Chi sq.=128.73 P=0.00 CFI=0.94 RMSEA=0.06
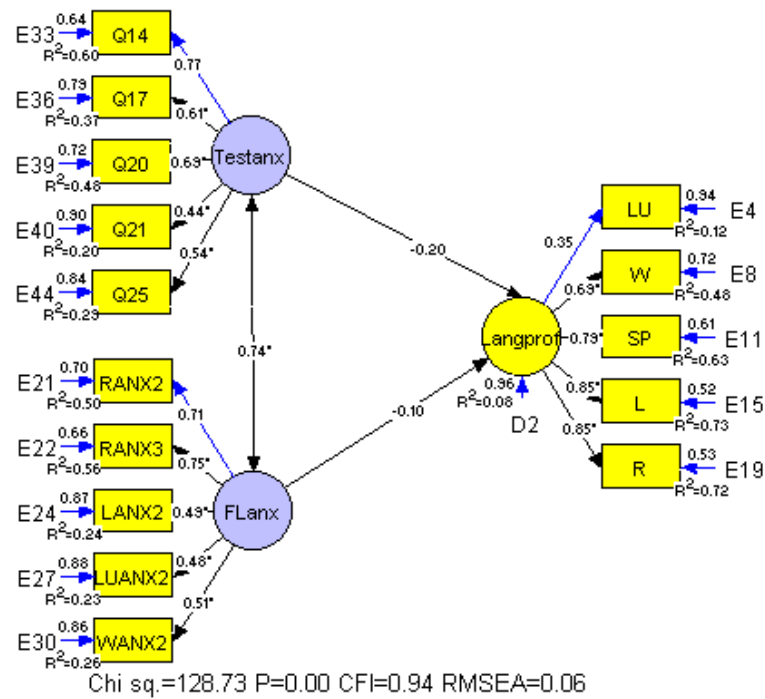
Figure 37 Model 9 Effect of test trait anxiety and foreign language test state anxiety.

10.3.5.4 Meta-cognitive competence interaction with anxiety and language performance

The meta-cognitive strategy questionnaire contained questions referring to all three areas of meta-cognitive strategies: goal setting, planning and assessment. Purpura (1999) suggested that meta-cognitive competence was, according to his research a uni-dimensional construct composed of different assessment skills (ability to assess the needs of the task, one's own competence and the resources needed to fulfil the needs of the task). I used the same questions in my questionnaire and exploratory factor analyses yielded three separate factors. I hypothesized that the meta-cognitive competence consisted of three independent factors each representing one area and interacting with the two other areas. I developed Model 10 (Figure 36) consisting of three independent factors: the three meta-cognitive strategy areas, each interacting with the other two. The theoretical model fits the data well (CFI =0.96), however, the

probability (P=0.08) does not allow us to give credibility to this model. I think this can be explained by the fact that the model is underspecified, as both the language anxiety and language performance are missing. As they are added (see Model 10A) the CFI and chi square increase and p decreases. Therefore I will use Model 10 (Figure 36) just as a confirmatory factor analysis to see that all the three factors are well represented by the measured variables, with the loading ranging from .42 for Question 28 (assessment strategy) to .69 for Question 18 (goal setting strategy).
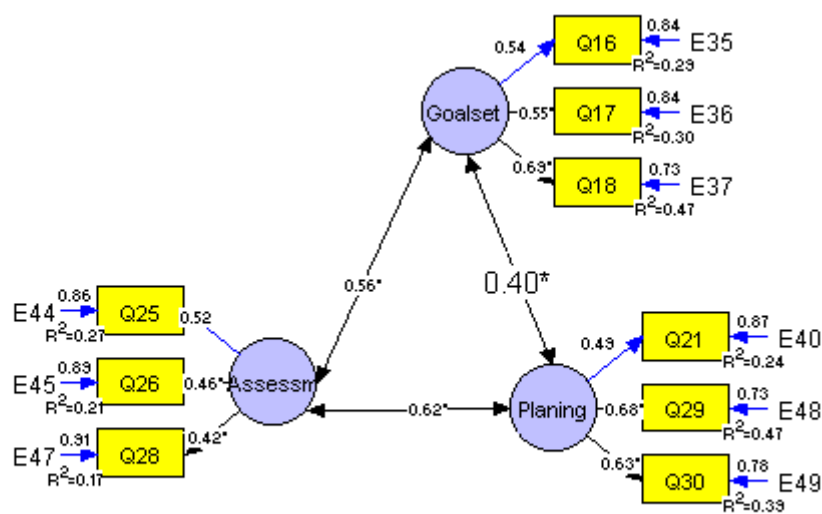
Figure X: EQS Model META.EDS
Chi sq.=34.51 P=0.08 CFI=0.96 RMSEA=0.04

Figure 38 Model 10 Interaction between meta-cognitive strategies

The interaction between the different meta-cognitive strategy areas is the highest between planning and assessment (.62) and the lowest between goal-setting and planning strategies (.40), which suggests strong and significant interaction between the three areas contrary to Bachman and Palmer's (1996) flow-chart predictions (see section 3.3.1.).

10.3.5.4.1 Interaction between meta-cognition and language performance
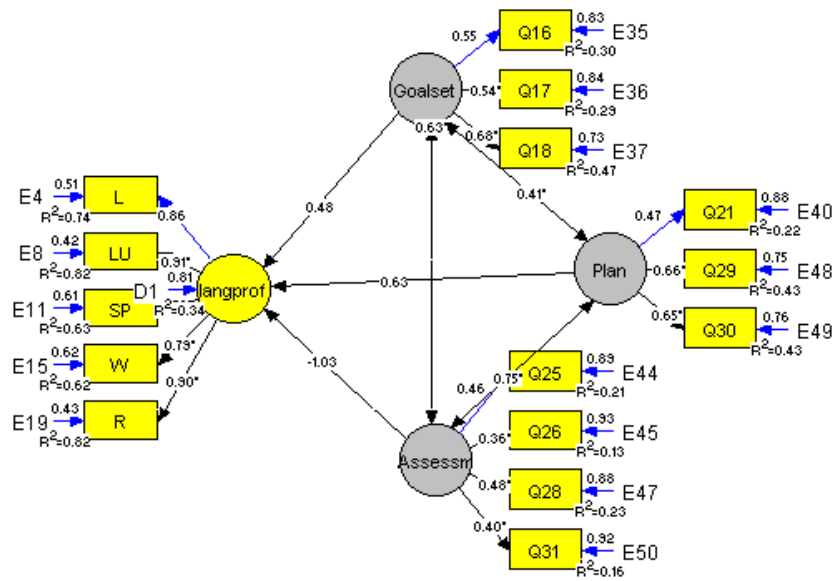
Figure X: EQS Model LANGMET.EDS
Chi sq.=120.50 P=0.01 CFI=0.97 RMSEA=0.04

Figure 39 Model 10 A Interaction between meta-cognitive strategies and language proficiency

Model 10 A explores the interaction between the meta-cognitive strategy areas, the bi-directional interaction between the three areas and the simultaneous impact of each of them on language performance, which in this model is a dependent factor. We can see that all the characteristics of Model 10 A as compared to Model 10 have improved (CFI is 0.97, and the probability is down to 0.01), which suggests that the model fits the data well and we have additional evidence that all the three areas of meta-cognition affect language performance: the reported goal setting and planning strategies have a positive impact (regression coefficients .48 and .60), but the reported assessment strategies have a negative impact (-1.03). I will now proceed to explore the interaction between each language skill and the three meta-cognitive strategy areas. All the language skills are defined as independent variables that interact with each other and simultaneously with meta-cognitive strategy areas.
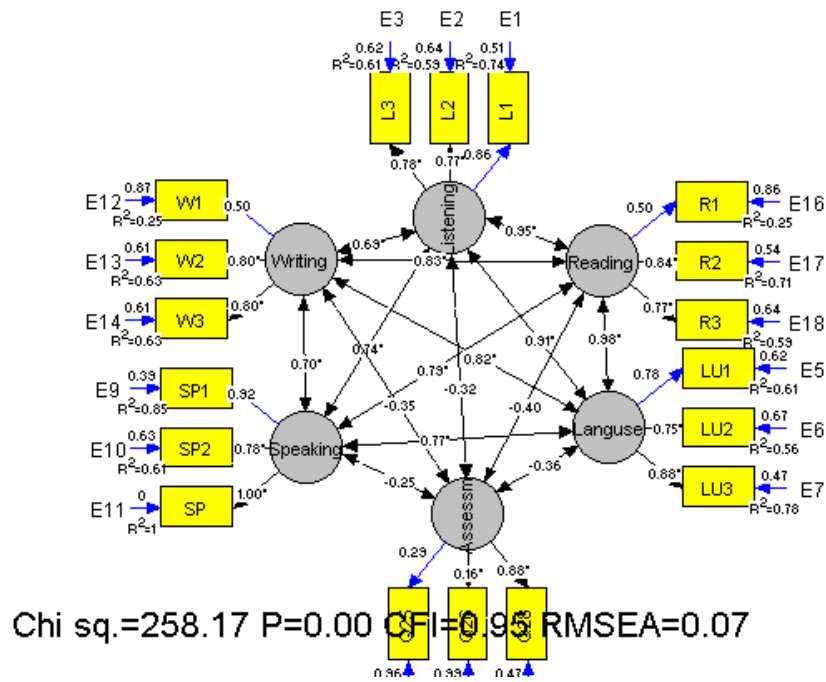
Figure 40 Model 10 B Interaction between meta-cognitive strategies and language proficiency

Model 10 B explores the interaction between the assessment strategies and all the

language skills as represented by the performance on Year 12 examination. The model

fit is good (CFI is .95) and probability is smaller than .01 and this provides credibility

to the model. The estimates of the links between the different factors, however, do not

provide significant information: we can see that assessment strategies have an

insignificant negative correlation with all the language skills or elements. As

expected, there is significant interaction between all the separate language skills. I

will not include the models presenting the interaction between the other areas of meta-

cognitive strategies and language skills, as all the correlations are insignificant.

10.3.5.4.2 Interaction between meta-cognitive strategies and anxiety
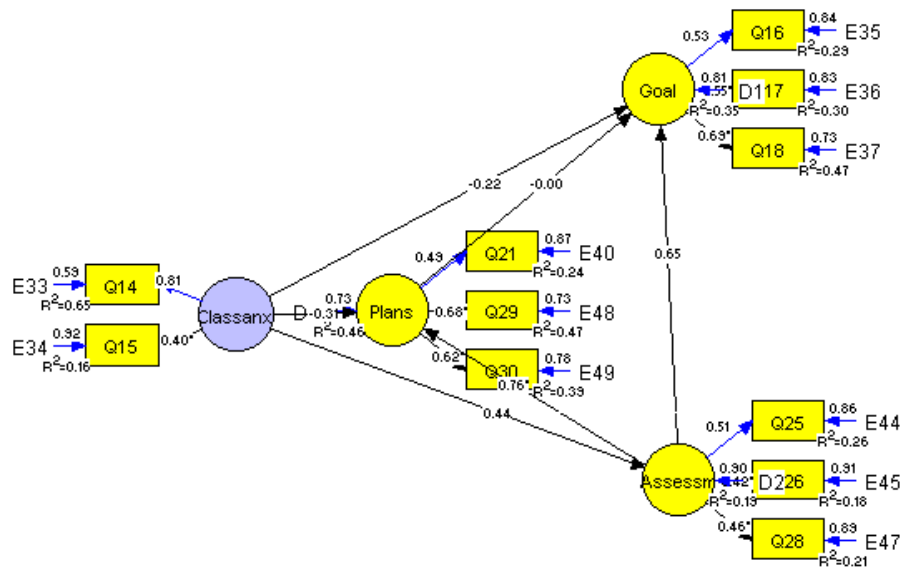
Figure X: EQS Model SAKARS4.EDS
Chi sq.=56.11 P=0.03 CFI=0.94 RMSEA=0.04

Figure 41 Model 11 The effect of classroom anxiety on meta-cognitive strategy use

Model 11 contains all three meta-cognitive strategies and classroom anxiety. We can see that although the CFI has decreased, the probability of the model is higher, (p=0.03). The relationships between the meta-cognitive strategy area have also changed: assessment of the task demands and one's own possibilities affect the goals (.65) and the plans (.76). There is no interaction between plans and goals. This explains why the CFI is lower in this model. The new variable in this model, classroom anxiety, has a positive relationship with assessment (.44) and planning (.31), but a negative relationship with goals (-.22). Evidently, the interaction between plans and goals has been taken over by classroom anxiety.

I will now proceed to examine the interaction between anxiety and each of the meta-cognitive strategy areas and language performance. The interaction between the meta-cognitive strategy areas and anxiety and language performance is usually predicted to be bi-directional and it turned out that with my data too, the models with two-way arrows fitted the data better than one-way arrow models. Therefore I will present only the oblique factor models here.

10.3.5.4.3 Interaction between assessment strategy, language anxiety and language performance

The Model 12 depicting the interaction between assessment strategies, foreign language test state anxiety and language performance as independent factors fits the data well (CFI is 0.96 and P<0.001). The interaction between assessment strategies and both the anxieties is active and positive (from .43 to .57). The correlation between language performance and anxiety is again stronger with the trait anxiety (classroom anxiety -.49) than with state anxiety (foreign language anxiety -.25). The interaction between language performance and assessment strategies is also negative (-.29).

Chi sq.=141.61 P=0.00 CFI=0.96 RMSEA=0.05
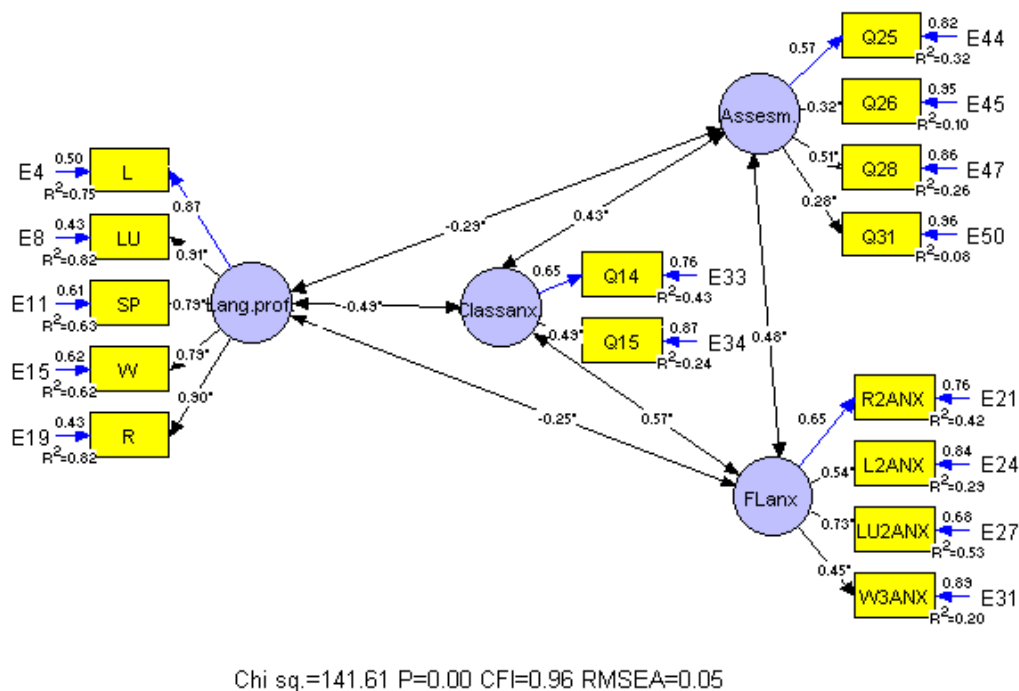
Figure 42 Model 12 Interaction between assessment strategies, anxieties and language proficiency

The interactions between foreign language test state anxiety, classroom anxiety and assessment strategies are positive, significant and fairly strong (.48 and .43). The interaction between the assessment strategies and language proficiency is negative and significant but is not very strong (-.29).

The interaction model (Model 12, Figure 41) does not tell me which is the cause and which is the effect, so I developed an orthogonal model (Figure 42) that makes assessment strategies the cause of both trait and state anxiety. Although this model does not fit the data as well (CFI is 0.93), it still fits well and we can see the extent to which assessment strategies are responsible for causing significant anxiety: (classroom anxiety (.41) and foreign language test state anxiety (.32). Thus Model 13 supports Eysenck's suggestion that anxiety has a cognitive basis (see section 4.1.1.).
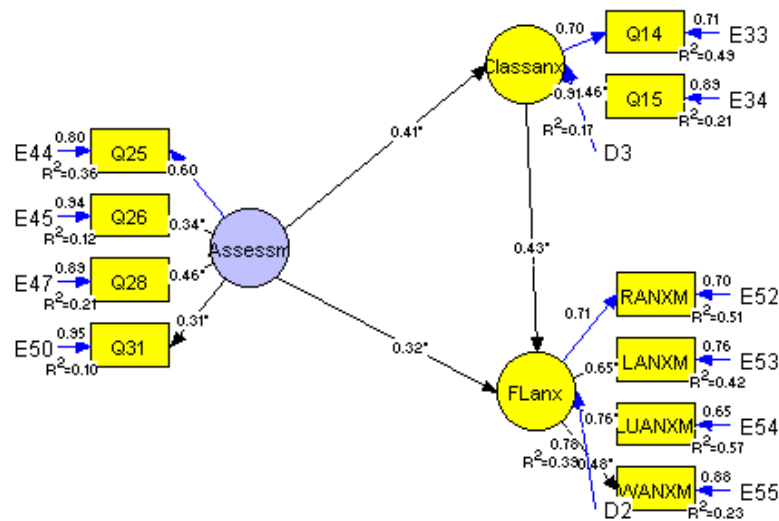
Chi sq.=56.58 P=0.00 CFI=0.93 RMSEA=0.06

Figure 43 Model 13 The causal effect of the assessment strategies on anxiety

10.3.5.4.4 Interaction between planning strategy, language anxiety and language performance

If we compare the interaction between assessment strategies and anxiety and planning strategies and anxiety, we can see that although there is a considerable interaction between planning strategies and foreign language test state anxiety (.29), there is no interaction between planning and classroom anxiety.

Figure 44 Model 14 Interaction between Planning strategies, anxieties and language proficiency

Chi sq.=118.05 P=0.00 CFI=0.96 RMSEA=0.05
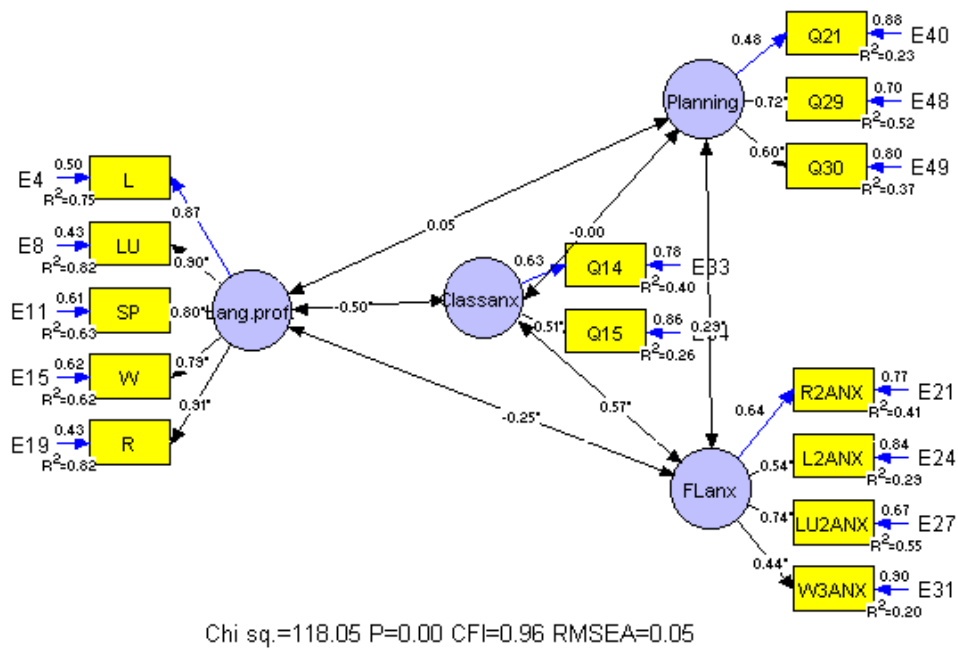
The fact that there is so little interaction between planning and language performance did not seem to be very logical, so I developed a model to examine the effects of planning strategies on anxiety and language performance (see Model 15 Figure 43) without the direct connection between classroom anxiety and language proficiency. Here we can see that although the model's CFI is slightly lower (.94) than in Model 11, it shows the effect of planning strategies on language performance (.18). This suggests that classroom anxiety and planning strategies interact during the test and have a combined impact on language performance.

Figure 45 Model 15 Effect of Planning strategies and anxieties on language performance
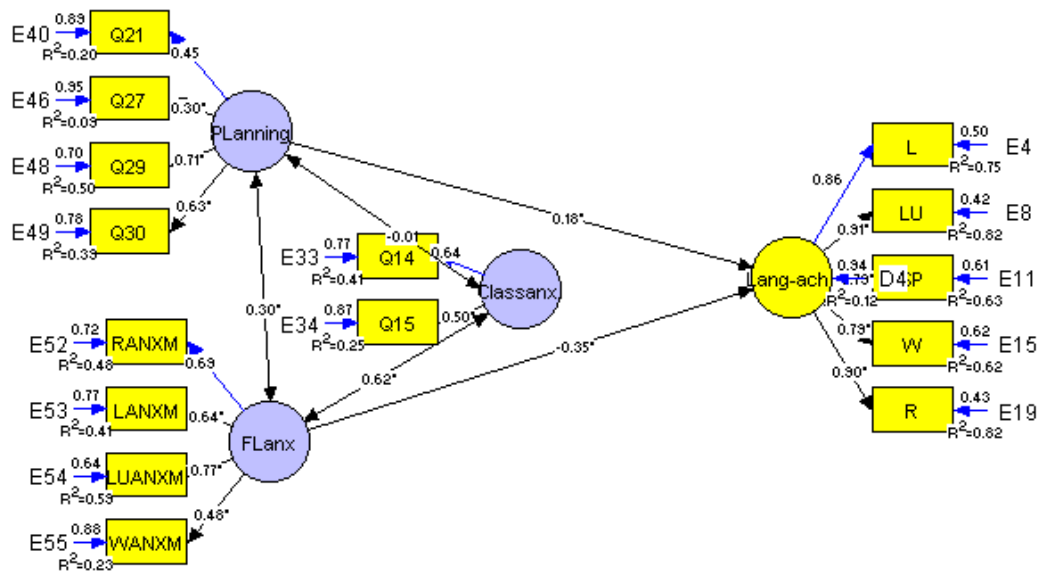
Figure X: EQS Model P6ROCK.EDS
Chi sq.=164.59 P=0.00 CFI=0.94 RMSEA=0.06

10.3.5.4.5 Interaction between goal-setting strategy, language anxiety and

language performance

Model 16 depicts the interaction between goal setting strategies, anxieties and

language performance. In my study goal setting strategies have a significant positive

correlation with foreign language test state anxiety and does not have any significant

interaction with language performance or classroom anxiety.

Figure 46 Model 16 Interaction between goal setting strategies, anxieties and language proficiency
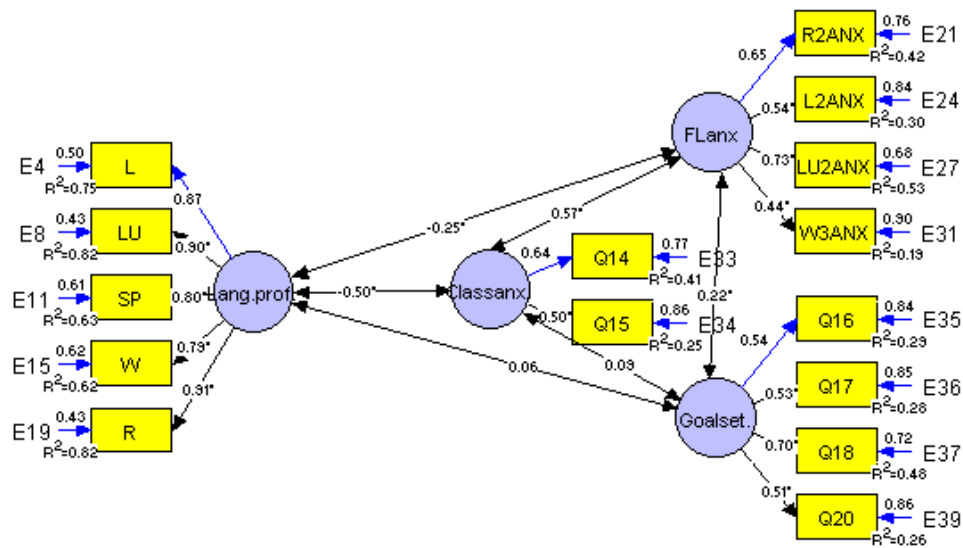
Chi sq.=123.56 P=0.00 CFI=0.97 RMSEA=0.04

If we compare the interaction between the three areas of meta-cognitive strategies (see Table 57), we can see that the assessment strategy area differs from the planning and goal setting strategy areas: its influence on language performance (factor Lang.prof.) is negative, it acts as a cause for both classroom anxiety (.41) and foreign language test state anxiety (.32) (see Model 11). This disagrees with Purpura's (1999) conclusion that meta-cognition is a uni-dimensional construct mainly composed of different assessment strategies. On the other hand my findings support his conclusion that assessment strategies are more influential than other meta-cognitive strategy areas.

Table 57 Comparison of the interaction between meta-cognitive strategy areas, anxieties and language performance

Strategy area	Interaction with classroom anxiety (factor Classanx)	Interaction with foreign language test state anxiety (factor Flanx)	Interaction with language proficiency (factor Lang.prof.)
Assessment	.43	.48	-.29
Planning	-.00	.29	.05
Goal-setting	.09	.22	.06

The model that best depicts the interaction between different individual variables and language performance using most of the data from year 2000 is Model 17. This model fits the measured data well (CFI is .97, p<0.001). It depicts the interaction between the meta-cognitive strategy areas and shows its effects on classroom anxiety (.67) and foreign language test performance (-.44) as well as the influence of classroom anxiety on language performance (-.22).

Figure 47 Model 17 Interaction between the meta-cognitive strategies, classroom anxiety and language proficiency

Chi sq.=138.66 P=0.00 CFI=0.97 RMSEA=0.04
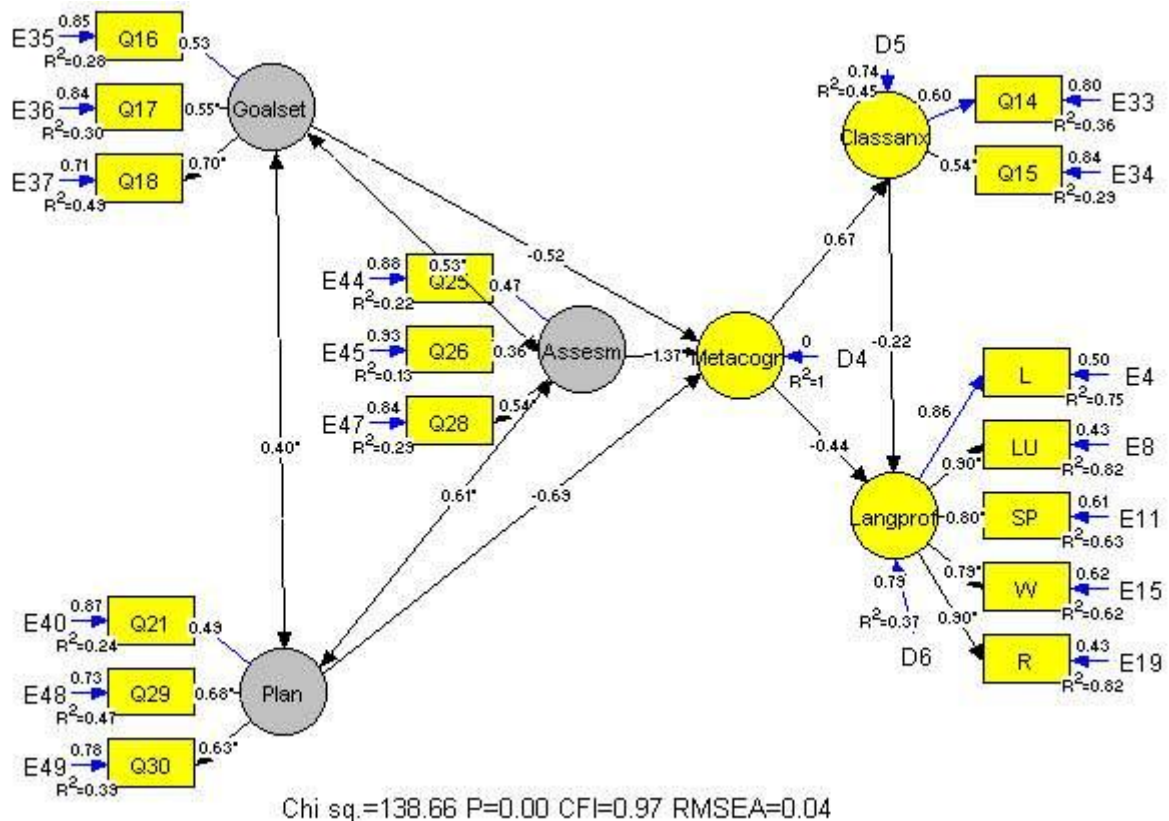
Both the regression path coefficients propose that meta-cognition, anxiety and language proficiency are all part of the process and have an effect on language test performance. I had to remove foreign language test state anxiety factor from this model as it had a very low variance coefficient (-.06) in interaction with language proficiency and as a result the CFI of the whole model was only .91. This suggests once again that foreign language test state anxiety has a minor influence when compared with the trait anxieties (classroom and test trait anxiety) and the meta-cognitive strategy area factors. Model 17 also agrees with Scherer's (2000) model of emotion production which suggests that appraisal (cognitive factor) is the basis of emotion, in this case, trait classroom anxiety. The negative influence of the meta-cognitive factor on language performance agrees with Purpura's (1999) finding that conscious use of meta-cognitive strategies during the test interferes with performance.

10.4 Findings of Study 3

The main findings of Study 3 are the following:

- Anxiety experienced during the foreign language test consists of three different dimensions: foreign language anxiety (formed by separate language skill anxieties (of which reading anxiety is the strongest); test anxiety (a trait acquired by previous experiences of tests in general) and classroom anxiety (which has a strong interaction with writing and language use anxiety). Both classroom and test anxieties (as traits) act as causal factors for foreign language test anxiety.

- The effect of classroom anxiety on language test performance is much stronger than that of foreign language test anxiety or test anxiety

- Meta-cognitive competence is a three-factor concept, consisting of goal setting, planning and assessment strategies. The interaction between the meta-

cognitive strategy areas is strong and bi-directional. This agrees with Bachman and Palmer's (1996) views on meta-cognitive strategies.

- Assessment strategy acts as a causal factor for both foreign language test and classroom anxiety, which agrees with Eysenck's (1992) view that anxiety has a cognitive basis. On the other hand classroom anxiety also activates both planning and assessment strategies. This supports Scherer's view that appraisal is at the basis of affect, but it is not enough; my study suggests, that in its turn, affect influences meta-cognition. This agrees with self-organisation theory that was proposed by Lewis (1996 and 2000).

- Meta-cognitive strategies, when they are consciously used during a test have a more negative influence on language test performance than classroom anxiety.

- Language use model that fits best incorporates meta-cognitive strategies, anxiety and language performance in a single network, thus supporting Bachman and Palmer's (1996) language use model of language use that sees strategic competence and affect as the basis of language production.

Chapter 11   Conclusion

In the final chapter of my research I will summarize the findings of my studies in order to answer the questions of research as they were posed in Chapter 6, discuss the limitations of my research and suggest implications for further research.

My research was largely based on Bachman and Palmer's (1996) theoretical model of language use and Purpura's (1999) research on the role of meta-cognition in language use as I compared these with the recent findings in psychology.

Having reviewed the theories on meta-cognition I found that in psychology the concept has become inactive (Brown 1987) since the role of meta-cognition both as knowledge about one's own cognition as well as processes of self-organisation has been taken over either by the concept of cognition or consciousness (see section 2.1). The situation in linguistics turned out to be different, as here meta-cognition contains a goal setting component. This differentiates the use of the concept in the two sciences and explains the longevity of the concept in applied linguistics. Furthermore, the existence of a concept that contains both cognitive and affective variables within the same framework provides basis for an investigation of interaction between cognition and emotion.

My research of existing theories of affect and meta-cognition and their functions provides evidence of interaction between meta-cognition and affect as some of the functions are attributed to both meta-cognition and affect. Interaction between meta-cognition and affect was also found in practical research (see the findings of Studies 2 and 3). Thus both my theoretical and practical investigation lead to the conclusion that our affective reaction depends on the assessment of the match between the task demands and our own abilities and anxiety is a reaction to the results of the assessment that is taken into account by our (meta) cognition. This agrees with the findings of, for example, Eysenck, who considers that anxiety has a cognitive basis.

11.1 Research questions

My research questions dealt with the two concepts explored in the theoretical part of the thesis: anxiety and meta-cognition as well as the interaction between the two concepts.

I investigated the existence of all the foreign language anxieties discussed in the previous chapters, their level, their observable signs, causes and effects on foreign language proficiency. Having found evidence of interaction between affect and cognition, I proceeded an exploration of the use of meta-cognitive strategies, its frequency, interaction between the different areas of the use of meta-cognitive strategies and their impact on foreign language performance.

Finally I explored the interaction between the two concepts (meta-cognition and anxiety) as well as their interaction with the foreign language performance. This section will summarize my findings on meta-cognition, anxiety and their impact on language use.

11.1.1 The role of meta-cognitive strategies in language use

In this section I present my main findings on meta-cognition, the concept itself and its interaction with language performance and anxiety. This part of research is based on the data provided by Purpura's (1999) meta-cognition questionnaire.

11.1.1.1   What is the construct of meta-cognition?

Although I used Purpura's (1999) questionnaire on meta-cognitive strategies, the findings of my research did not agree with his suggestion that meta-cognition is a single-factor concept consisting mainly of assessment strategies.

According to my research the concept of meta-cognitive strategies was found to consist of three different areas, goal setting, planning and assessment strategy areas, which have significant bi-directional interaction between them (see section 10.3.2), as

predicted by Bachman (1990). Apart from that, assessment strategies have a significant impact on both planning strategies as well as goal setting strategies.

11.1.1.2   What is the interaction between meta-cognition and language performance like?

Meta-cognitive strategy areas interact among themselves, at the same time affecting language performance. The influence of conscious use of goal-setting and planning strategies is insignificant, but the influence of conscious use of assessment strategies is significant and negative. The models produced to examine the impact of meta-cognitive strategy areas on each language skill suggest that the direct influence is insignificant. This agrees with Purpura's (1999) finding that meta-cognition does not have a direct impact on language performance.

11.1.1.3   What is the interaction between meta-cognition and anxiety like?

The influence is bi-directional; on the one hand, meta-cognition causes classroom anxiety (see Model 17, section 10.3.2), on the other hand, classroom anxiety activates all meta-cognitive strategy use (See Model 11 section 10.3.2).

11.1.1.4   How often are meta-cognitive strategies consciously used?

My findings suggest that test takers consciously use all three meta-cognitive strategy areas, but the most active are planning and monitoring and the least active is assessment area (see section 10.2).

11.1.1.4.1 How often are meta-cognitive strategies used in different proficiency groups?

Medium proficiency groups are consciously using all strategy areas more often than the high and the low proficiency groups. The planning and monitoring areas are the most active for these groups both during a study period (as they try to understand someone speaking English) and during the test-taking period (as they try to concentrate when taking a test).

11.1.1.4.2 How often are meta-cognitive strategies used in groups of different level of anxiety?

The most active area of meta-cognitive strategies is planning and monitoring for the medium anxious test-takers. If we look at the overall frequency of use of strategies, we find that the two highest anxiety groups use all the strategies more often but the two lower anxiety groups use all strategies less often. This tendency is especially evident in assessment strategies (see section 10.2.2). This provides another evidence of the link between meta-cognition and anxiety and assessment strategy as its basis.

11.1.2 The role of anxiety

The role of anxiety was examined with the help of observation, interview and questionnaire methods.

11.1.2.1 What is the level of anxiety during the Year 12 English language examination*?*

The overall level of foreign language test state anxiety fluctuates during the test and follows the actual and not the expected task-difficulty level, with some exceptions (for example, unexpected task type or topic, see section 10.2.5). The change in level of anxiety is more radical for the higher language proficiency groups; their reactions are stronger when they have an unexpected task or topic. The level of anxiety does not

depend on language skill or the task type although the listening anxiety level tends to be higher than it is for the other language skill tests.

The different-levels-of-anxiety groups are not the same as different-levels-of-proficiency groups: the test-takers with the highest levels of anxiety are distributed across all proficiency levels. Nevertheless the level of anxiety does depend to some extent on test-takers' foreign language proficiency level as the highest level of proficiency group has the lowest level of anxiety and the lowest level of proficiency group has the highest level of anxiety (see section 10.2.5).

11.1.2.2 What signs of anxiety can be observed during the examination?

All the signs listed in the anxiety sign framework (Oxford 1999) could be observed (general avoidance, physical actions, and physical symptoms) during the Year 12 examination of English.

The majority of test-takers avoided direct interaction with the administrators of the test or avoided touching the test materials although they were willing to examine the materials' packaging after other test-takers had volunteered. One test-taker avoided looking at the interlocutor and sometimes did not seem to be listening to what was being said.

Test-takers were constantly squirming and fidgeting (wringing their hands, tapping their feet, hitting their pens against their palms etc.) before the test and during the Listening test pauses. One test-taker started stuttering during the Speaking test, some had difficulty in breathing (they breathed in fits and starts and often sighed) many were red in the face. Both the interviewers also showed signs of anxiety (see Chapters 8 and 9).

11.1.2.3   What does the concept of anxiety consist of?

Anxiety during a foreign language test consists of foreign language test state anxiety (which in its turn consists of reading, listening, writing and language use anxiety), and two trait anxieties: classroom anxiety and test anxiety (see Model 2 in section 10.3.2.).

11.1.2.3.1 What does foreign language test state anxiety consist of?

In my research foreign language test state anxiety was formed by four separate language skill anxieties (reading, listening, language use and writing anxiety (speaking anxiety was not included in my questionnaire). All the four language skill/element anxieties interacted during the test. The interaction between language use and reading anxiety was the strongest (see section 10.3.2.3).

11.1.2.3.2 What is the relationship between foreign language test state and test trait anxiety?

General test anxiety that has become a trait causes foreign language test state anxiety and interacts with all four foreign language skill anxieties. The interaction between test anxiety and reading anxiety is the strongest (see section 10.3.2).

11.1.2.3.3 What is the relationship between foreign language test state and classroom anxiety?

Classroom anxiety that has become a trait is activated by the test situation and in its turn causes foreign language test state anxiety. During the test it interacts with all four language skill anxieties and has its strongest interaction with writing and language use anxiety (see section 10.3.2).

11.1.2.4   What are the self-reported causes of test anxiety?

The self-reported causes of anxiety can be divided into three groups: test situation, foreign language use and problems of test design or administration.

In the whole sample 'foreign language use' caused the largest number of complaints by students and contains criticism of practically all aspects of language use (vocabulary, grammar, spoken and written texts). It is the middle range of anxiety and proficiency level test-takers that complain about foreign language use the most often. Foreign language use complaints were mentioned least often in the lowest proficiency group.

In the highest anxiety group, 'test situation' is the most often mentioned cause of anxiety and it is usually connected to comments on the test difficulty level and to respondents' own reaction to this.

Among the comments on the causes of anxiety caused by the examination design or administration problems, unexpected task types or themes are mentioned 3 times more often than all the other complaints. All in all test administration and test design problems are mentioned as a cause of anxiety most often in the two highest proficiency groups (see section 10.2.5)

In addition to questionnaires I used also SEM to analyse the source of anxiety. The analysis suggests that anxiety has a cognitive basis: both classroom and foreign language test anxieties are to a great extent caused by meta-cognitive strategies (planning (regression path coefficient .22) and assessment (.48) areas) and foreign language test anxiety is indirectly caused by trait anxieties: classroom anxiety (.58) and test anxiety (.73) (see section 10.3.2).

11.1.2.5   What are the effects of anxiety?

Anxiety affects both language performance, the type and the frequency of use of meta-cognitive strategies.

11.1.2.5.1 What is the effect of foreign language test state anxiety on foreign

   language performance?

Foreign language test state anxiety has an insignificant effect on foreign language test performance. It is half the effect of test trait anxiety (see Model 9) and has nearly ten times less effect than classroom anxiety (see Model 7 in section 10.3.2)

11.1.2.5.2 What is the effect of classroom anxiety on foreign language test

   performance?

Classroom anxiety has a significant negative effect on foreign language test performance and affects performance in all language skills. The correlation between classroom anxiety and language performance is significant and negative for test-takers regardless of their level of anxiety. It was strongest for Reading and Language Use tests in the medium proficiency groups (see section 10.3.2).

11.1.2.5.3 What is the effect of anxiety on meta-cognitive strategy use?

Classroom anxiety affects all three areas of meta-cognitive strategies: it activates both planning (.31) and assessment strategies (.44). It has a negative influence on the use of goal strategies. Goal setting strategies are reportedly used more often in medium proficiency and medium anxiety groups, planning strategies are used regardless of test-takers' level of anxiety or proficiency but assessment strategies are used most often in the highest anxiety and medium proficiency level groups. The higher the level of anxiety, the more assessment strategies are activated and used (both during the study period and during the test) (see section 10.3.2)and the more they are consciously activated the lower the performance level.

11.1.2.6   What is the effect of classroom discussion of the impact

   ## of anxiety on test-performance?

For the whole sample, there was no significant correlation between the frequency of discussions of test-anxiety impact during the test preparation phase and language

performance, but for the highest proficiency group the correlations were significant mostly positive and moderately strong. For the lowest anxiety test-taker group the frequency of test-anxiety discussions had significant positive and moderately strong correlations with nearly all the tasks (see section 10.2.5).

11.1.3 What is the interaction between anxiety, meta-cognition and language performance during a language test?

According to the self-report questionnaire results meta-cognition, anxiety and language proficiency interact during a language test as separate factors. Test situation activates not only foreign language anxiety, but also general test trait anxiety and classroom anxiety. These in turn activate meta-cognitive strategy use, especially assessment strategy use that further escalates anxiety.

Thus the interaction is bi-directional, as during use of assessment strategies, students, having evaluated the test task demands and their abilities, may detect a lack of proficiency and this produces more anxiety. This leads to more vigorous assessment strategy application and as a result the test-takers' attention is divided which leads to deterioration in performance. This agrees with White's (1981) findings.

If, however, the test-taker's goals are not concerned with test situation, and the test-taker has a overarching goal of performing in the best possible manner (this can be caused for example, by the interviewer's personality), the test-taker is fully engrossed in the task, the test-situation is forgotten, the positive external feedback from the interviewer promotes spontaneous interaction, idea generation and initialisation. The model depicting the interaction between meta-cognition, anxiety and language performance suggests that conscious application of meta-cognitive strategies affects performance directly, as well as indirectly through classroom anxiety (see section 10.3). This supports Bachman and Palmer's (1996) theoretical model of interaction

between meta-cognitive competence and affect in language use and language test performance.

11.2 Limitations of the study

The conclusions presented above can be generalized only if we take into account the following limitations of the study:

1. population generalizability is possible only to test-takers who have received formal training in English as a foreign language and who are 17 to 18 years old

2. foreign language anxiety was measured during writing, listening, language use and reading tests, but not during the speaking test. Thus, speaking anxiety, which is often considered as the most important (see Horwitz 1986), was not measured. To compensate, I used an interview to discuss the level of anxiety and its causes during a speaking test. Nevertheless, all the conclusions from the statistical data concerning foreign language test anxiety could be affected by the fact that it did not contain the speaking anxiety variable.

3. although the study did not concern itself with differences between the genders in foreign language anxiety, I must point out that all the interviewees who volunteered to discuss their concerns about the Speaking test were girls

4. the models depicting the impact of anxiety and meta-cognition on foreign language performance explore neither the links with other affective variables (such as motivation and self-confidence) nor with cognitive strategies (such as summarizing, associating, transferring and inferencing)

which have also been reported to have an impact on foreign language performance

5. and finally, my research is based on self-report questionnaire results. Brown (1987) considers that meta-cognitive knowledge and processes are often subconscious. As a result, a self-report questionnaire cannot reveal the use of meta-cognitive strategies during automatic processing. Similarly, the self-report anxiety level questionnaires measure the level and the causes of anxiety that the test-takers have commented upon and were aware of. This is why my study reveals only the interaction between language performance and consciously used meta-cognitive strategies and consciously experienced anxiety.

11.3 *Implications of the research*

The theoretical implications section (11.3.1) will assess my main findings and methodological implications (11.3.2) section will evaluate the different methods used in this research.

11.3.1 Theoretical implications

The study was concerned with the exploration of meta-cognition and anxiety and their interaction with language performance. The findings in each of the areas of research have their own theoretical implications.

11.3.1.1 Meta-cognition in language use

Both my theoretical and practical research supported Bachman and Palmer's (1996) framework of meta-cognition, which suggests that meta-cognition consists of three different areas, that is, goal setting, planning and assessment strategy areas, that constantly interact and affect each other during a foreign language test.

Simultaneously these are also largely responsible for our affective reactions (anxiety being one of them).

Therefore O'Malley's (1989) definition of meta-cognition as

1. knowledge about cognition or

2. applying thoughts about the cognitive operations of oneself or others

3. and regulation of cognition (see Chapter 2)

has to be reformulated according to my research findings so that it includes affect. Therefore, I would re-define meta-cognition as

1. knowledge about cognition and affect or

2. applying thoughts about cognitive and affective operations of oneself or others

3. and regulation of cognition and emotion.

In this definition we can see firstly, how meta-cognition interacts with affect, secondly, how we can apply knowledge we have on our own emotions and those of others to achieve our aims, and finally, we have a mechanism that is responsible for regulating and applying emotions, instead of simply controlling them. This approach, to my mind opens up our view not only on affect and emotions, but also meta-cognition, and not only in language testing, but also language teaching. At the same time it agrees with Lewis' (2000) self-organisation theory, which considers that our decisions are not based on a one-off assessment of a situation and our reaction to it, but on a constant interaction between cognition and emotion that is constantly reassessed and readjusted for the needs of a situation. I think that a dynamic view of meta-cognition which suggests constant human development is not only more flattering to us, but also more realistic.

11.3.1.2   Affective variables in language use

The results of my findings on anxiety, as one of the affective variables, agree with Stevick's (1999) definition of affect. He says that affect towards a particular thing or action is how this action fits in with one's needs and purposes and its resulting effect on emotions (see section 3.3). According to my research the source of affective reaction (positive or negative) is based on an assessment strategy that examines the task demands and one's own abilities. If these match, the affective reaction is confidence, if they do not, the result of the assessment is negative and the affective reaction is anxiety.

Anxiety, according to my research is a multi-dimensional construct, consisting of trait anxiety (formed by classroom and general test anxiety) and foreign language test state anxiety (consisting of separate language skill anxieties). This agrees with Horwitz's (1986) and Cheng et al's (1999) findings.

Trait (classroom and general test) and state (reading, listening, writing and language use) anxieties interact during the test, but it is classroom anxiety that has the most significant negative impact on foreign language performance.

We can create anxiety both voluntarily and involuntarily as could be seen in both interview and comments on causes of anxiety in my research. For example, there were several students who said that they consciously used pre-test anxiety to motivate their better preparation for the examination (see Chapters 9 and 10).

The view of anxiety as a motivator agrees with Eysenck's (1992), Csikszentmihalyi's (1998), May's (1977) and Kierkegard's (1849) suggestions that the role of anxiety is to see that we develop our potential to its fullest. In the case of language learning, this means acquiring language as well as possible in our situation, and warning us if we have not learned language as well as we could have (hence the negative correlations between one's language proficiency and anxiety level in my findings in Chapter 10).

The finding that we are ourselves responsible in a way for creating our own anxiety (both teachers and students) is not new (see Lewis 1996 or LeDoux 1999). However, the myth of the 'test' as the essential cause of anxiety and anxiety in its turn as a means of further distorting the measurement allows teachers and test-takers to move responsibility onto a foreign agency. Teachers blame tests for upsetting their students and test-takers blame tests for their inability to perform as well as they could have. Thus the common myth flourishes and attracts new believers, until the test-taker is even worried about not being worried (see Karina's comment in Chapter 9).

Furthermore, test-developers use correlation coefficients for measuring the impact of test-anxiety on language performance and find that there is a low but significant negative correlation, which cannot answer the question whether anxiety is a cause or an effect of performance deterioration. As a result the concept of test-anxiety in test-taking population becomes even more menacing.

The findings of my research, that the test state anxiety effect on test-takers' language performance is negligible, and nearly ten times smaller than classroom anxiety that has become a trait, attempts to shift the attention from what takes place during the test to students' everyday experiences in the classroom. My research suggests that it is the foreign language classroom where the test performance is decided and not the examination hall.

I do understand that this finding will be popular neither with teachers (who will have to accept responsibility for creating classroom anxiety), nor with students (who will have to admit that they did not learn the language as well as they could have), but I hope that it will give an opportunity for test developers to share responsibility with society for the anxiety provoked by the tests.

11.3.1.3 Interaction between language performance, meta-cognition and anxiety

All my findings support Bachman and Palmer's (1996) theoretical model of language use that predicts interaction between meta-cognition and affect. The interaction between motivation and meta-cognition is already well researched and recorded (see section 3.4.2). Test and classroom anxiety, however, are usually researched only in an interaction with language performance and are usually found to cause deterioration in foreign language performance (see for example, Horwitz 1986). The interaction between meta-cognition and anxiety on the other hand seems to have been neglected in applied linguistics (I do not know of any research on interaction between anxiety and meta-cognition in applied linguistics). Therefore anxiety is often seen as a source of disruption and chaos and meta-cognition as a means of control of one's learning (see, for example, Cohen 1998).

This thesis also set out to measure the extent of damage test anxiety causes to test-takers' performance. Recent theories in psychology, however, suggest that an interaction between cognition and emotion is a source of emotion and the focus of my thesis therefore gradually became the interaction between anxiety and meta-cognition. The findings of my research suggest that two supposedly opposing mechanisms, regulating (meta-cognition) and disrupting (anxiety) are in reality intimately connected and even have a causal relationship, that is, assessment strategies cause anxiety, which in its turn activates planning and monitoring strategies.

Although theoretically this is nothing new, everybody knows that anxiety is one of the strongest affective variables (Oxford 1999), and affect interacts with meta-cognition (Bachman and Palmer 1996), nevertheless it is difficult to accept the finding that the strongest allies in the fight with unruly emotions, is, in fact, responsible for their creation.

If we accept the causal role of meta-cognition, we cannot blame emotion as something damaging and dangerous. On the contrary, according to May (1979), LeDoux (1999) and Lewis (2000), emotion assesses the match between our goals and if it discovers a conflict, we perceive it ads anxiety. We can strive to control it (as Cohen (1998) suggests or we can face the problem and look for a solution, thus combining emotion and cognition in order to solve the conflict instead of subduing one or the other.

Apart from assessing the compatibility of our goals, emotions also have other functions. Fifty years ago, when foreign language testing was concerned with our ability to reproduce what we had learned *about* language, affect could be seen as interference, now that language testing asks us to express our own thoughts and our own ideas, and then assess their appropriacy to the situation, we have to accept that other roles of affect (for example, control of access to memory, making associations with previous experiences, affective feedback and rhythm of our speech) has a significant role in language production.

I think that all the findings on meta-cognitive competence can be adjusted to include emotions and to elaborate theory and develop methodology that can help our students understand their own emotions and emotions of others in their own culture and that of the foreign language. Thus affect, with the help of meta-cognitive strategies can become an efficient tool in language acquisition.

11.3.2 Methodological implications

I found the area of interaction between language performance, anxiety and meta-cognition to be an exciting research subject, although complex, but at the same time rewarding. This was partly caused by the recent findings in consciousness research that have enabled researchers to re-examine existing theories in many sciences, including linguistics.

Nearly all the findings of this study are based on self-report data that were selected and interpreted in an equally subjective manner, but then, Damasio (2000) says that the human mind itself is a subjective phenomenon. The methods used in this research do not allow me to pretend that I have managed to analyse the interaction between the cognitive and emotional variables during a language test; all I have done is taken a few snapshots of the end result of what the test-takers feel while taking a test and how this affects their performance. In spite of all the limitations discussed in section 11.2, I hope that combined analyses of the results of external manifestations (observation: Study 1), comments on the internal experiences (interviews: Study 2 and questionnaires: Study 3) and the Structural Equation Modeling managed to verify objectively the consistencies of many individual subjectivities (Damasio 2000).

Of all the methods used in my research (observation, interview, questionnaire and SEM) I found mathematical modelling the most revealing. Firstly because it finally allowed me to separate causes from effects; secondly, because I found the graphical form of communication with a program easy to acquire and explicit for understanding complicated notions with many variables, and thirdly because it had such vigorous measures for testing the reliability of its models, that when I finally had a model with good fit, I felt I could trust the model and the relationships between the variables that the model suggested.

The method that I found most frustrating was the use of correlation coefficients, because of their ambiguousness: the fact that you never know which variable causes which. Nevertheless correlations did allow me to discover certain regularities that were later tested with the help of SEM.

The most pleasurable of all the methods, was interviewing test-takers and later analysing the transcripts of the interviews. During this stage I suddenly discovered answers to questions I had never asked. Then of course I wished I had asked other

questions (for example, concerning meta-cognitive strategies. Nevertheless, I think that the information I acquired during the interviews and observing the test-takers gave me a feeling of immediate involvement.

11.3.3 Suggestions for further research

I think that the recent findings in consciousness research in psychology will allow us to re-examine the theories in language production in general and phonetics in particular, for example to explain the role of affect in deciding whether we keep our native language pronunciation or acquire the foreign language pronunciation. However, before we can use the findings of psychology in applied linguistics we will have to resolve the problem of terminology.

As I did not meet the term 'meta-cognition' in recent articles or monographs in psychology I have to presume that the function of control and regulation of cognition and emotion in experimental psychology has been taken over by the concepts of cognition, consciousness and 'self' (Damasio 2000 and LeDoux 1999). This makes cross-referencing between the two sciences more complicated as every researcher has to compare the use of concepts and make the connection on his or her own. Nevertheless, the exchange of the terms of 'meta-cognition and affect' for 'consciousness', in a language use framework is not possible either as to my mind this would have both positive and negative consequences. On the one hand it would connect applied linguistics to the latest findings on human consciousness and open new possibilities in research of language production and acquisition, but on the other hand it would render all the research on meta-cognition out-dated thus leaving a void in our understanding of language use.

Evidently, the best solution would be a research on similarities and differences in the use of the terms of 'meta-cognition' and 'consciousness' in psychology and linguistics that would lead to a reformulation of the existing applied linguistics theories using

the new terminology. This would allow us to connect the findings of the two

sciences. Then we would say that the basis of language use is extended consciousness

that connects the external environment with the internal environment of the mind.

Damasio says that the work of extended consciousness can be assessed by assessing

1. recognition, recall, working memory,

2. emotion and feeling and

3. reasoning, decision making and planning over large intervals of time
   (Damasio 2000, p. 202).

All the three elements (cognition, emotion and meta-cognition) make language

acquisition possible and language use appropriate to the needs of the individual in a

definite situation (Bachman and Palmer 1996). Thus, by testing language

performance we are indirectly testing the work of our consciousness, that is our

ability to recognise, recall and use working memory, emotions and feeling, our ability

to reason, make decisions and plan. All these features are an inherent part of language

ability.

The role of affect changes depending on whether we use language to report our

experiences (ideational function), extend our knowledge of the world around us

(heuristic function) or *create and extend our environment for humorous or*

imaginative aesthetic purposes, where the value derives from the way in which

language itself is used (imaginative function, Bachman 1990, p. 94). This, however,

is just a suggestion for further research, which I hope and trust will be carried out

soon. Then we could move from admitting the crucial role of affect in language

acquisition (Oller 1983) and theories exploring the role of affect in language use

(Bachman and Palmer 1996 and Stevick 1999) to practical implications of the

findings on affect into language acquisition and testing processes. This would allow

language testers, teachers and students to resolve many of their anxieties concerning

language learning, teaching and testing, just as it allowed me to resolve my worries

concerning the effect of foreign language test anxiety on foreign language test

performance.

Bibliography

Aida Y. (1994) Examination of Horwitz, Horwitz and Cope's construct of foreign language anxiety: the case of students of Japanese. *Modern Language Journal* Vol.78, pp.155-168.

Alderson C.J., C. Clapham, D. Wall (1995) Language Test Construction and evaluation, Cambridge University Press.

Alderson J.C. (2000) Assessing Reading, Cambridge Language assessment series, Cambridge University Press.

Alderson J.Ch., Szollàs K. (2000) The context: the current School-leaving examination in Alderson J.Ch., Nagy E., Eniko (eds.) English Language education in Hungary (Part 2), the British Council, Hungary.

Alexander P.A., Schallert D.L., Hare V.C. (1991) Coming to terms: how researchers in learning and literacy talk about knowledge. *Review of Educational Research* Vol. 61, pp.315-343.

Allen G.J., M.J. Elias, S.F. Zlotow (1980) Behavioural interventions for alleviating test anxiety in Sarason I. (Ed.) Test anxiety: Theory, research and applications, Hillsdale, Erlbaum, New York.

Arnold M.B. (1960) Emotion and personality, Columbia University Press, New York.

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

Bachman L.F., A.S. Palmer (1996) Language testing in Practice, Oxford University Press.

Bargh J.A. (1992) Being unaware of the stimulus versus unaware of its interpretation: why subliminality per se matter to social psychology. R. Bornstein and T. Pittman (Eds.) Perception without awareness, Guilford, New York.

Barlow D.H. (1988) Anxiety and its disorders, Guilford, New York.

Bartlett F.C. (1932) Remembering , Cambridge University Press, UK.

Beck A.T. and G. Emery (1985) Anxiety disorders and phobias: a cognitive perspective, Basic Books, New York.

Becker P. (1982) Fear reactions and achievement behaviour of students approaching an examination in H.W. Krohne and L. Laux (Eds.) Achievement, stress and anxiety, Hemisphere Publishing Corporation.

Bentler P.M., (1998) EQS, A Structural Equation Program, Multivariate Software, INC, Version 5.7b.

Bialystok E. (1978) A theoretical model of second language learning, *Language learning*, No 28, pp.69-83.

Bourhis R.Y. (1990) Social and individual factors in language acquisition in B. Harley, P.Allen, J. Cummins and M. Swain (Eds.) Development of second language proficiency, Cambridge University Press.

Bower G.H. (1981) Mood and memory, *American Psychologist* No.36, pp.129-148.

Bower G.H. (1994) Some relations between emotions and memory in P.Ekman, R.J.Davidson (Eds.) The nature of emotion, Oxford University Press.

Brown A. (1987) Meta-cognition, executive control, self-regulation and other more mysterious mechanisms, in F.E. Weinert, R.H. Kluwe, Lawrence (Eds.) Meta-cognition, motivation and understanding, Erlbaum Associates.

Brown A.L. (1978) Knowing when, where and how to remember: a problem of meta-cognition. In R.. Glaser (Ed.) Advances in instructional psychology (Vol. 1) Hillsdale, Lawrence Erlbaum associates, New York

Brown A.L. (1980) Meta-cognitive development and reading. In R. Spiro, B.Bruce and W. Brewer (Eds.) Theoretical issues in reading comprehension. Hillsdale, Lawrence Erlbaum associates, New York

Brown H.D. (1994) Teaching by principles, Englewood Cliffs, Prentice Hall, New York.

Burkovec T.D., E. Robinson T Pruzinsky and J. A. DePree (1983) Preliminary exploration of worry: some characteristics and processes. *Behaviour research and therapy* No 21, pp.9-16.

Cacioppo J.T., W.L. Gardner, G.G. Berntson (1999) The affect system has parallel and integrative processing components: form follows function, *Journal of Personality and Social Psychology*, Vol. 76 No5, pp. 839-855.

Canale M., Swain M. (1980), Theoretical bases of communicative approaches to second language teaching and testing in *Applied Linguistics* 1/1 pp.1-47.

Cannon W.B. (1929) Bodily changes in pain, hunger, fear and rage. Appleton, New York.

Cassidy T. (1999) Stress, cognition and health, Routledge.

Cheng Y., E.K. Horwitz, D.L. Schallert (1999) Language Anxiety: differentiating writing and speaking components, *Language Learning* No 49:3, September, p.417-446.

Chomsky N. (1980) Rules and representations, Basil Blackwell, Oxford.

Clark D.M. (1988), A cognitive model of panick attacks. In S. Rachman and J.D. Masser (Eds.) Panic: psychological perspectives Lawrence Erlbaum associates, Hillsdale, New York.

Cohen A.D. (1998), Strategies in learning and using second language, Longman.

Cohen A.D. (1984) On taking tests: What the students report. Language Testing Vol. 1 (1), pp. 70-81.

Combs and Marzano (1994) Self-system in Motivation: theory and research, H.F.O'Neil, Jr, M.Drillings (Eds.) Lawrence Erlbaum Associates, New Jersey.

Coulter (1986) Affect and social context: Emotion definition as asocial tasks in The social construction of emotions Ed. R. Hare, Basil Blackwell, New York.

Csikszentmihalyi M., Nakamura J. (1989) The dynamics of intrinsic motivation: a study of adolescents. In C. Ames Research on motivation in education, Academic Press San Diego.

Csikszentmihalyi M. (1998) Living well: the psychology of everyday life, Phoenix, UK.

Culler R.E. and C.J. Holahan (1980) Test anxiety and academic performance: the effects of study-related behaviour, *Journal of educational psychology* Vol.72, pp.16-20.

Damasio A. (2000) The feeling of what happens. Body, emotion and the making of consciousness, Vintage, UK.

Damasio, A. (1994) Descartes' Error: Emotion, Reason and Human brain, Avon, New York.

Davies A., A.Brown, C. Elder, K. Hill, T.Lumly. T McNamara (1999) Dictionary of Language testing, Cambridge University Press.

Deci E.L., R.M. Ryan (1985) Intrinsic Motivation and self-determination in Human behaviour, Plenum Press, New York.

Deffenbacher J.L. (1978) Worry, emotionality and task-generated interference in test anxiety: an empirical test of attentional theory. *Journal of educational psychology*, Vol.70, pp. 248-254.

DiTomasso R.A. (1981) A meta-analyses theraphy outcome research with test anxious college students, Dissertation Abstracts International, Vol.41.

Dodge K.A. (1991) Emotion and social processing. In J.Garber and K.A. Dodge (Eds.) The development of emotion regulation and dysregulation, Cambridge University press, Cambridge.

Dyer C. (1997) Beginning research in psychology, Blackwell publishers, Oxford.

Eckland G.A. (1981) Conceptual framework for cognitive motivational processes. Scheme Theory, Academic Press Inc.

Ehrman M.E. and R.L. Oxford (1991) Affective survey, Foreign service institute, Arlington.

EQS, A Structural Equation Program, Bentler P.M., Multivariate Software, INC, Version 5.7b (1985-1998).

Erdelyi M.H. (1992) Psychodynamics and the unconscious. *American Psychologist* No.47, pp.784-787.

Ericsson K.A. and H.A. Simon (1980) Verbal reports as data, *Psychological review* Vol. 87, pp. 215-251.

Eysenck H.J. (1979) The conditioning model of neurosis. *Behavioral and brain sciences,* No 2, pp. 155-199.

Eysenck M.W. (1992), Anxiety: the cognitive perspective, Lawrence Erlbaum Publishers: Hove UK.

Faerch C and G. Kasper (1983) Strategies in Interlanguage communication, Longman, London.

Flannagan J.C. (1954) The critical incident technique *Psychological Bulletin* No 51, pp.327-358.

Flavell J.H. (1979) Meta-cognition and cognitive monitoring: a new area of cognitive –developmental inquiry, *American Psychologist*, No.34, pp. 906-911.

Fodor J.A. (1987) Psychosemantics, Bradford books, Cambridge.

Fransson A. (1984), Crammimg or understanding, Effects of intrinsic and extrinsic motivation on approach to learning and test performance in J.C. Alderson and A.H. Urquahart (Eds.) Reading in Foreign language, Longman , London.

Frijda N. (1986) The emotions, Cambridge University Press, Cambridge.

Frijda N.H. B. Mesquita B. (2000) Beliefs through emotions, in N.H. Frijda and A.s. Manstead, S.Bem (Eds.) Emotions and Beliefs, Cambridge University Press, Cambridge.

Ganschow L. and R. Sparks (1994) Differences in anxiety and language performance among high, average and low anxious collage foreign language learners. *Modern Language Journal*, 78, pp. 41-55.

Gardner R.C. (1983) Social psychological aspects of second language learning, Edward Arnold, London.

Gardner R.C. (1985) Social psychology and second language learning: the role of attitudes and motivation. Edward Arnold, London.

Gardner R.C. (1990) Attitudes, motivation and personality as predictors of success in foreign language learning, in Language aptitude reconsidered. Centre for Applied Linguistics by Prentice Hall Inc.

Gardner R.C. (1997), Individual differences and second language learning, Encyclopaedia of language and education, Vol.4, Kluwer Academic Publishers, Canada.

Gardner R.C. and P.D. MacIntyre (1991) An instrumental motivation in language study: who says it isn't effective? *Studies in Second language acquisition*, No 13, pp. 57-72.

Goldstein K. (1938) A further comparison of the Moro reflex and the startle pattern, *Journal of Psychology*, No 6, pp. 33-42

Goldstein M.D. and M.J. Strube (1994) Independence revisited: the relation between positive and negative affect in naturalistic setting. *Personality and social psychology bulletin* Vol. 20, pp. 57-64.

Goleman D. (1995) Emotional Intelligence, Bantam, New York.

Gombert J.E. (1992) Metalinguistic development, Harvester Wheatsheaf.

Gray J. A. (1990) Brain Systems that Mediate both Emotion and Cognition *Cognition and Emotion*, 4, pp.269-288.

Hatch and Lazarton (1991) The research manual design and statistics for applied linguists, Heinle and Heinle Publishers.

Heckhausen H. (1991) Motivation and action. Springerverlag, New York.

Heider F. (1958) The psychology of interpersonal relations, John Wiley and Sons Inc.

Heinrich D.L. C.D. Spielberger (1982) Anxiety and complex learning in Krohne H.W., Laux L. (Eds.) Achievement, stress and anxiety, Hemisphere Publishing House.

Hilleson (1996) I want to talk with them… in K.M. Bailey and. D. Nunan (Eds.), Voices from the classroom, Cambridge University Press.

Hollandsworth J.G., R.C. Glazeski, K.Kirkland, G.E.Jones, L.R. Van Norman (1979) An analyses of the nature and effects of test anxiety: cognitive, behavioural and physiological components. *Cognitive therapy research* Vol.3, pp. 165-180.

Horwitz E.K. (1986) Preliminary evidence for the reliability and validity of a foreign language scale, *TESOL Quarterly* Vol.20, pp.559-562.

Horwitz E.K., D.J. Young (Eds.), (1991) Language anxiety: from theory and research to classroom implications, Prentice Hall.

Horwitz E.K., M.B. Horwitz, J.Cope (1986) 'Foreign language Classroom anxiety' in *The Modern Language Journal*, 70 (2), pp.125-132.

Houston B.K. (1977) Dispositional anxiety and the effectiveness of coping strategies in stressful laboratory and classroom situations in Spielberger C.D., Sarason I. (Eds.) Stress and Anxiety vol. 4, Hemisphere Publishing Corporation.

Hull C.L. (1952) A behaviour system: an introduction to behaviour theory concerning the individual organism, Yale University Press, New Haven.

Humphreys G.W., and W. Revelle (1984) Personality, motivation and performance: a theory of relationship between individual differences and information processing, *Psychological review* No.91, pp.153-184.

Isen A.M. (1984) Toward understanding the role of affect in cognition. In R. Wyer and T. Scrull (Eds.) Handbook of social cognition (vol.3) Lawrence Erlbaum Associates Inc., Hillsdale, New York

ITEMAN (1995) Conventional Item Analyses Program, Assessment systems Coorporation.

Izard C.E., B.P. Ackerman, K.M. Schoff and S. Fine (2000) Self-organisation of discrete emotions, emotion patterns and emotion-cognition relations, in Lewis M.D, I. Granic (Eds.) Emotion, development and self organistion, Cambridge studies in social and emotional development, Cambridge University Press.

Kelly G. (1955) The Psychology of personal constructs, Norton, New York.

Kierkegaard S. (1941). Sickness unto death, Translated by Walter Lowrie, Princeton University Press, (Original published in 1849).

Kierkegaard S. (1944). The concept of dread, Translated by Walter Lowrie, Princeton University Press, (Original published in 1844).

Kirkland K., J.G. Hollandsworth (1980) Effective test taking: skills-acquisition versus anxiety reduction techniques. *Journal of consulting and clinical psychology*, Vol.48, pp.431-439.

Kunnan A.J. (1995), Test Taker characteristic's and test performance, Cambridge University Press.

Lazarus R.S., J.C. Coyne, S. Folkman (1984) Cognition, Emotion and Motivation: The Doctoring of Humpty Dumpty in K.R. Scherer, P. Ekman (Eds.) Approaches to Emotion, Lawrence Erlbaum Associates, London.

LeDoux J. (1999) The Emotional Brain. The mysterious underpinnings of emotional life, Phoenix.

Levi L. (1987) Definitions and the conceptual aspects of health in relation to work. In R.Kalimo, M.A. El_Batawi and C.L. Cooper (eds.) Psychological factors at work and their relation to health, WHO, Geneva.

Lewis M.D, I. Granic (2000) A new approach to the study of emotional development in Lewis M.D, I. Granic (Eds.) Emotion, development and self organistion, Cambridge studies in social and emotional development, Cambridge University Press.

Lewis M.D. (1996) Self-organising cognitive appraisals, *Cognition and emotion*, No10 (1), pp.1-25.

Lewis M.D. (2000) Emotional self-organisation at three time scales in Lewis M.D, I. Granic (Eds.) Emotion, development and self organistion, Cambridge studies in social and emotional development, Cambridge University Press.

Liddell H.S. (1949) The role of vigilance in the development of animal neuroses in P. Hoch and J Zubin (Eds.). Anxiety, Grune and Stratton, New York.

Liebert R., L.Morris (1967), Cognitive and emotional components of test anxiety: a distinction and some initial data, *Psychological reports,* 20, pp.975-978.

Locke E.A., G.P. Latham (1994), Goal setting theory in H.F. O'Neil Jr. and M. Drillings (Eds.), Motivation: theory and research, Lawrence Erlbaum Associates, Publishers, Hove UK.

MacIntyre P.D. (1995) How does anxiety affect second language learning? A reply to Sparks and Ganschow, *The Modern Language Journal* 79(1), pp. 90-99.

MacIntyre P.D. and R.C. Gardner (1994) The effects of induced anxiety on three stages of cognitive processing in computerised vocabulary learning. *Studies in Second Language Acquisition* No.16, pp.1-17.

MacIntyre, P.D. and R. C. Gardner (1991) Methods and results in the study of anxiety and language learning: a review of the literature, *Language Learning*, Vol.41 (1), pp.85-117.

MacIntyre, P.D., R. C. Gardner (1991) Language Anxiety: its relationship to other anxieties and to processing in native and second languages, *Language Learning*, Vol.41 (4), pp.513-534.

Madsen H.S. (1982) Determining the debilitative impact of test anxiety, *Language Learning*, Vol. 32, (1) pp.133-143.

Madsen, H. S., Brown L.B. and Jones L. R., 1991 Evaluating Student Attitudes Toward Second Language Tests in Horwitz E.K., D.J. Young (Eds.), Language anxiety: from theory and research to classroom implications, Prentice Hall.

Mandler G. and S.B. Sarason (1952) A study of anxiety and learning. *Journal of abnormal and social psychology*, No.47, pp.166-173.

Mandler G. (1984) Mind and body: psychology of emotion and stress, Norton, New York.

Maslow A.H. (1968), Toward a psychology of being, Van Nostrand Reinhold, New York.

Mathews G., D.R. Davies, S.J. Westerman and R.B. Stammers (2000) Human performance: cognition, stress and individual differences, Psychology press, Sussex.

May R. (1979), The meaning of anxiety, Pocket books, Gulf and western Corporation, New York

May R. (1981) Value conflicts and anxiety, in I.L. Kutasch, L.B. Schlessinger (Eds.) Handbook on stress and anxiety, Jossey-Bass Publishers.

McLelland D.C. (1961) The achieving society, Van Nostrand, New York.

McLaughlin (1990), The relationship between first and second languages: language proficiency and language aptitude, in B. Harley, P. Allen, J. Cummins and M. Swain (Eds.), Development of second language proficiency, Cambridge University Press.

McLaughlin B. (1990) Restructuring, *Applied Linguistics*, No 11, pp.113-128.

McNamara T. (2000) Language Testing, in H.G. Widdowson's (Series ed.) Oxford introductions to language study, Oxford University Press.

McNamara T. F. (1996) Measuring Second Language Performance, Longman.

Morris L., W. Engle (1981) Assessing various coping strategies and their effects on test performance and anxiety. *Journal of Clinical Psychology* Vol. 37 pp.165-171.

Morris L.W. and R.M. Liebert (1969) Effects of anxiety on timed and untimed inteeligence tests, *Journal of consulting and clinical psychology*, No.33, pp.240-244.

Morris L.W., M.A. Davis and C.H. Hutchings (1981) Cognitive and Emotional components of anxiety: Literature review and revised worry-emotionality scale, *Journal of Educational psychology* Vol. 73, No. 4, pp.541-555.

Morris L.W., N.R. Brown, B.Halbert (1977) Effects of symbolic modeling on the arousal of cognitive and affective components of anxiety in preschool children in Spielberger C.D., Sarason I. (Eds.) Stress and Anxiety vol. 4, Hemisphere Publishing Corporation.

Morris L.W., R.S. Fulmer (1976) Test anxiety (worry and emotionality) changes during academic testing as a function of feedback and test importance. *Journal of Educational Psychology*, Vol.68, pp.817-824.

Mowrer O.H., A.D. Ullman (1945) Time as a determinant in integrative learning, *Psychology Review*, Vol. 52:2, pp. 61-90.

O'Malley, Chamot, Stewner-Manzanares, Kupper and Russo (1985), Learning strategies used by beginning and intermediate ESL students, *Language learning,* Vol. 35, No.1 pp.21-45.

Oatley K. and P.N. Johnson-Laird (1987) Towards a cognitive theory of emotions. *Cognition and emotion* No.1, pp. 29-50.

Oatley K.and J.M. Jenkins (1996) Understanding emotions Blackwell, Oxford.

Okada M., R.L. Oxford, S. Abo (1996) Not all alike: motivation and learning strategies among students of Japanese and Spanish in an exploratory study, in R.L. Oxford's (Ed.) Language learning motivation: pathways to new century, Second language teaching and curriculum center.

Oller J.W.Jr. (1983) A consensus for the eighties? In Issues in Language Testing research. Rowley, Mass.:Newbury House.

O'Malley J.M., A.U. Chamot (1990) Learning strategies and second language acquisition, Cambridge University Press.

O'Malley J-M., A.U. Chamot, G. Stewener-Manzares, R.Russo and L.Kupper (1985) Learning strategy applications with students of English as a second language, *TESOL Quarterly*, Vol. 19, pp. 285-296.

Onwuegbuzie A.J., P Bailey. and C.E.Daley (2000) The validation of three scales measuring anxiety at different stages of the foreign language learning process: the Input anxiety scale, the Processing anxiety scale and the Output anxiety scale, *Language Learning,* Vol.50 No.1, pp. 87-117.

Ost L.G. (1987) Age of onset in different phobias, *Journal of abnormal psychology*, No.96 pp.223-229.

Oxford R. (1990) Language learning strategies: what every teacher should know, Newbury House, Rowley.

Oxford R.L. (1990) Styles, strategies and aptitude: connections for language learning in 'Language aptitude reconsidered'. Centre for applied Linguistics by Prentice Hall Inc.

Oxford R.L.(1990) Language learning motivation- pathways to the new century. Second language learning and curriculum centre, Unversity of Hawai.

Oxford R.L. (1999) Anxiety and the language learning: new insights in J. Arnold (Ed.), 'Affect in Language Learning', Cambridge University Press.

Paulman R.G., K.J.Kennely (1984), Test anxiety and ineffective test taking: different names, same construct?, *Journal of educational psychology* Vol.76, No.2, pp.279-288.

Piaget J. (1976) The grasp of consciousness: action and concept in the young child, MA: Harvard University, Cambridge.

Piaget J. (1978) Success and understanding, MA: Harward University Cambridge.

Porter D. (1991) Affective factors in language testing, in Ch. Alderson and B.North (Eds.) 'Language Testing in the 1990s', Phoenix.

Power M., T.Dalgliesh (1998) Cognition and Emotion, Psychology Press, UK.

Price M. L. (1991) The Subjective Experience of Foreign Language: Interviews with Highly anxious Students Language anxiety: from theory and research to classroom implications Horwitz E.K. and Young D.J. (Eds.) Prentice Hall.

Purpura J.E. (1999) Learner strategy use and performance, Studies in language testing, Cambridge University Press.

Pylyshyn Z.W. (1984) Computation and cognition: toward a foundation for cognitive science. Cambridge MIT Press.

Raven J.C. (1965) Advanced progressive matrices, London H.K. Lewis.

Reeve J., E.L Deci. (1996) Elements of the Competitive Situation That Affect Intrinsc Motivation, *Personality and Social Psychology Bulletin*, Vol.22, Nr.1.

Richardson F.C., H.F. O'Neil, R.D. Grant (1977) Development and evaluation of an automated test anxiety reduction program for a computer-based learning environment In J.E. Sieber, H.F. O'Neil, S.Tobias (Eds.) Anxiety, learning and instruction Hillsdale, Erlsbaum, New York.

Rolls E.T. (1990) A Theory of Emotion, and its Application to Understanding the Neural Basis of Emotion, *Cognition and Emotion*, 4, 161-190.

Rubin J. (1981) Study of cognitive processes in second language learning *Applied Linguistics* No 2(2) pp.117-131.

Ryle G. (1949) The concept of mind, Penguin, London.

Saito Y., Th.J. Garza, E. K. Horwitz (1999) Foreign language reading anxiety, *The Modern Language Journal*, 83 (2), pp.202-218.

Sarason I. (Ed.) 1980 Test anxiety: Theory, research and applications Hillsdale, NY: Erlbaum

Sarason I.G. (1958) Interrelationships among different individual difference variables, behaviour in psychoteraphy and verbal conditioning. *Journal of abnormal and social psychology*, N.56. pp. 339-344.

Sarason I.G. (1960) Empirical findings and theoretical problems in the use of anxiety scales. *Psychologal Bulletin*, Vol. 57, p.403-415.

Sarason I.G. (1978), The test anxiety scale: concept and research. Examination stress and test anxiety in Spielberger C.D., Sarason I. (Eds.) "Stress and Anxiety" Vol. 5, Hemisphere Publishing Corporation.

Sarason I.G. and R.Stoops (1978) Test anxiety and passage of time, *Journal of consulting and clinical psychology*, No. 46, pp.102-109.

Sarason S. and G.Mandler (1952) Some correlates of test anxiety, *Journal of abnormal and social psychology* No 47, pp. 810-817.

Sassenrath J.M. (1964) A factor analyses of rating scale items on the test anxiety questionnaire, *Journal of consulting psychology*, Vol. 28, pp. 371-377.

Schachter S. and J.E. Singer (1962) Cognitive, social and psychological determinants of emotional state. *Psychological review* No. 69, pp.379-399.

Schank R.C., R.P.Abelson (1977) Scripts, plans, goals and understanding, Lawrence Erlbaum Associates Inc. Hillsdale, New York.

Scherer K.R. (2000) Emotions as episodes of subsystem synchronisation driven by non-linear appraisal processes in Lewis M.D, I. Granic (Eds.) Emotion, development and self organistion, Cambridge studies in social and emotional development, Cambridge University Press.

Schlenker B.R., M.R. Leary (1985) Social anxiety and communication about the self. *Journal of Language and Social Psychology*, 4, pp.171-192.

Schmidt R., D. Boraie, O. Kassabgy (1996) Foreign language motivation: internal structure and external connections in Oxford R.L (Ed..) "Language learning motivation" Second Language Teaching and Curriculum Centre.

Shohamy E. (1982) Affective considerations in language testing, *Modern language journal*, Vol. 66, pp.13-17.

Schoonen R., Hulstijn J., Bossers B. (1998) Metacognitive and Language specific knowledge in native and foreign language reading comprehension: an empirical study among dutch students in grades 6, 8 and 10, *Language Learning* 48, 1, pp.71-106.

Schuman J.(1999) Affect in Language Learning, *Language Learning* Vol. 1, 1999

Scovel, T., 1991 The Effect of Affect on Foreign Language Learning Language anxiety: from theory and research to classroom implications (Eds. Horwitz E.K. and Young D.J.) Prentice Hall.

Selye H. (1956) The stress of life, McGraw-Hill, New York.

Selye H. (1981) The Stress Concept Today, I.L. Kutash, L.B. Schlessinger (Eds.) Handbook on Stress and Anxiety, Jossey Bass Publishers.

Shohamy E., 1982, Affective considerations in Language testing, *The Modern Language Journal*, 66, pp.13-17.

Skehan P.   (1998) A cognitive approach to language learning, Oxford University Press, Oxford.

Skinner B.F. (1953) Science and human behaviour, Macmillan, New York.

Smith C.A., L.W. Morris (1976) Effects of stimulative and sedative music on two components of test anxiety. *Psychology reports* Vol. 38 pp.1187-1193.

Smith J.A., R.,Harre, L.Van Langenhove (1995) Rethinking methods in psychology, Sage, London.

Snow R.E, D.N. Jackson III (1994) Individual differences in conation: selected constucts and measures in H.F. O'Neil Jr. and M. Drillings (Eds.), Motivation: theory and research, Lawrence Erlbaum Associates, Publishers, Hove UK.

Sparrow S.S., Davis S. (2000), Recent advances in the assessment of intelligence and cognition, *J.Child Psychology and Psychiatry Journal*, Volume 41, pp.171 to 131.

Spielbeger C.D., Gonzalez H.P., Taylor C.J., Algaze B., Anton W.D. (1978) Examination stress and test anxiety in Stress and Anxiety vol. 5, Spielberger C.D., Sarason I. (eds.) Hemisphere Publishing Corporation .

Stevick E. (1999) From chemistry to alchemy in J. Arnold (Ed.) Affect in Language Learning, Cambridge University Press.

Stevick E.W (1996), Memory, meaning and method, Newbury House, Heinle and Heink Publishers.

Swain, M. (1984) Large scale communicative testing: a case study. In Savignon S.J. and Berns M.S. (Eds) Initiatives in Communicative Language Teaching, Addison Weasley Reading MA, pp.185-201.

Tache, H. Selye (1978) On Stress and Coping Mechanisms in C.D. Spielberger, I.G. Sarason (Eds.)., The Series in Clinical and Community Psychology. "Stress and Anxiety", Vol.5.

Tarone E., G. Yule (1989), What learners do and do not know, Oxford University Press.

Trylong V.L (1987) Aptitude, attitudes and anxiety: a study of their relationship to achievement in foreign language classroom, Purdue University.

Tsui A.B.M., (1996) Reticence and Anxiety in Bailey K.M and D.Nunan (Eds.) 'Voices from the Language Classroom', Cambridge University Press.

Van Ek (1990), Threshold Level, Council of Europe, Strasbourg.

Vogley A.J. (1998) Listening comprehension anxiety: students' reported sources and solutions, *Foreign language annals*, Vol. 31:1, pp.67-80.

Weiner B. (1986) An attributional theory of achievement motivation and emotion. New York, Springer.

Weinert F.E. (1987), Metacognition and motivation as determinants of effective Learning and understanding in F.E. Weinert, R.H. Kluwe (Eds.) Metacognition, motivation and understanding. Lawrence Erlbaum Associates.

Weir C., 1995 Understanding and developing language tests, Phoenix.

Wenden A. (1991) Learner strategies for learner autonomy, Englewood Cliffs, Prentice Hall, New York.

Wenden A.L. (1998), Metacognitive knowledge and language learning, *Applied Linguistics* Vol.19/4, pp.515-537.

Wesche M. (1987) Second language performance testing: the Ontario Test of ESL as an example *Language Testing* 4, pp. 28-47.

Wine J.D. (1971) Test anxiety and direction of attention. *Psychological Bulletin*, Vol.76, pp.92-104.

Wine J.D. (1980) Cognitive-attentional theory of test anxiety in I.G. Sarason (Ed.) Test axiety: theory, research and applications, Erlbaum Hillsdale, New York

Wine J.D. (1982) Evaluation anxiety, a cognitive-attentional construct, in H.W.Krone and L.Laux (eds.), Achievement, stress and anxiety, Hemisphere Publishing.

Yerkes R.M. and Dodson J.D. (1908) The relation of strength of stimulus to rapidity of habit formation in *Journal of comparative neurology and psychology*, Vol. 18, pp. 459-82.

Young, D. J. (1986) Relationship Between Anxiety and Foreign Language Oral Proficiency Ratings on Foreign Language Learning in Horwitz E.K. and Young D.J. (Eds.) Language anxiety: from theory and research to classroom implications Prentice Hall.

Zajonc R.B. (1984) On the primacy of affect. *American psychologist* No.39, pp. 151-175.