

University of Warwick institutional repository: <http://go.warwick.ac.uk/wrap>

A Thesis Submitted for the Degree of PhD at the University of Warwick

<http://go.warwick.ac.uk/wrap/58287>

This thesis is made available online and is protected by original copyright.

Please scroll down to view the document itself.

Please refer to the repository record for this item for information to help you to cite it. Our policy information is available from the repository home page.

Measuring Language Learner Autonomy in Tertiary-Level Learners of English

David Dixon

**A thesis submitted in partial fulfilment of the
requirements for the degree of Doctor of Philosophy**

University of Warwick, Centre for Applied Linguistics

March 2011

TABLE OF CONTENTS

TABLE OF CONTENTS	I
LIST OF TABLES	V
LIST OF FIGURES	VI
ACKNOWLEDGEMENTS	VII
DECLARATION	VIII
ABSTRACT	IX
ABBREVIATIONS	X
1 INTRODUCTION	1
1.1 Background: Initial motivations	1
1.2 Direction	2
1.3 Overview	3
1.4 Why the research is important	5
2 LITERATURE REVIEW: MEASURING LEARNER AUTONOMY IN ADULT SECOND LANGUAGE LEARNING	6
2.1 Introduction	6
2.1.1 Aims of the Literature Review	6
2.1.2 Use of the term “Measure”	8
2.1.3 Overview of the structure of the Literature Review	10
2.2 Is there a need for a measure of autonomy?	11
2.3 Should autonomy be measured?	15
2.4 Aspects of autonomy	17
2.4.1 Autonomy is a multidimensional concept	18
2.4.2 Autonomy is variable	25
2.4.3 Autonomy is a capacity	35
2.4.4 Autonomy is demonstrated	35
2.4.5 Autonomy requires metacognition	37
2.4.6 Autonomy involves responsibility	39
2.4.7 Autonomy involves motivation	41
2.4.8 Autonomy involves social interaction	45
2.4.9 Autonomy is political	51
2.5 Previous autonomy measuring research	57
2.6 Discussion	75
3 METHODOLOGY	83
3.1 Overview of chapter	83
3.2 Clarifying the research aims	84
3.2.1 Aims	84
3.2.2 The broad construct of autonomy to be used in the research	85
3.2.3 Construct validity checking	86
3.2.4 The measuring aspect	88
3.2.5 Qualitative aspect	89
3.3 Research questions	90
3.4 What would the instrument be like?	92
3.4.1 Questionnaire type	94
3.5 Methodological stance/paradigm	95
3.5.1 Mixed methods research	96

3.6	Other researchers	98
3.7	The research design	101
3.7.1	The plan for the research.....	102
3.7.2	Smaller scale data collection.....	105
3.7.3	Sampling procedures.....	106
3.7.4	Subject groups.....	111
3.8	Reliability and validity	111
3.8.1	Reliability.....	111
3.8.2	Validity.....	113
3.9	Item selection.....	114
3.9.1	The construct of autonomy used in the present research	115
3.10	Questionnaire development	118
3.10.1	Putting items into a draft questionnaire	119
3.10.2	The parts of a questionnaire	123
3.11	Administering the questionnaire	125
3.11.1	Groups or individuals.....	125
3.11.2	Medium	126
3.11.3	Gathering feedback	127
3.12	Data analysis.....	128
3.12.1	Standardisation.....	129
3.12.2	Coding of responses	130
3.12.3	Factor analysis.....	131
3.13	Translation	141
3.13.1	Introduction.....	141
3.13.2	Reasons for translating.....	142
3.13.3	Translation equivalence	142
3.13.4	The choice of translation procedure.....	144
3.13.5	Cultural differences in questionnaire response	149
3.14	Overview of eventual research stages.....	151
3.15	Summary.....	152
4	THE LONG LIST STAGE	155
4.1	Overview	155
4.2	Writing items	155
4.2.1	Item wording guidelines.....	155
4.2.2	The choice of closed items.....	162
4.2.3	Items in the Long List	166
4.3	Subject groups	167
4.3.1	The UAE data gathering.....	167
4.3.2	English language Internet data gathering.....	171
4.3.3	Mandarin Chinese Internet data gathering.....	173
4.3.4	The BNU Mandarin Chinese data gathering.....	174
4.4	Chapter summary.....	176
5	THE SHORT LIST STAGE.....	177
5.1	Item selection process for Short List	178
5.1.1	Respondents' feedback.....	179
5.1.2	Range of response	180
5.1.3	Standard deviation.....	181
5.1.4	Polarisation of response	181
5.1.5	Selection table	182
5.1.6	Discrimination indexes.....	182

5.1.7	Combined selection procedure.....	184
5.1.8	Items in the Short List.....	187
5.2	Small scale research.....	187
5.2.1	Estimating autonomy levels.....	194
5.2.2	Student interview.....	196
5.2.3	Teacher interview.....	197
5.3	ELTCS year 3 data gathering.....	198
5.3.1	Description.....	198
5.4	Preessional ELTCS group data gathering.....	199
5.5	Other Warwick preessional students data gathering.....	200
5.6	Taiwan group data gathering.....	201
5.7	Ethics.....	202
5.8	Limitations.....	204
5.9	Conclusion.....	205
6	DATA ANALYSIS.....	206
6.1	Introduction.....	206
6.2	Larger scale data.....	207
6.2.1	Introduction.....	207
6.2.2	Factor analysis procedure.....	207
6.2.3	Groupings emerging from the factor analysis process.....	220
6.2.4	Comparison with literature.....	223
6.2.5	Comparison with Short List categories.....	226
6.2.6	Instrument performance indications.....	229
6.2.7	Summary results for the full sample.....	232
6.3	Smaller scale data.....	234
6.3.1	Introduction.....	234
6.3.2	Correlating teacher estimates with the questionnaire.....	234
6.3.3	ELTCS year-3 teacher interview and confidence-related categories.....	241
6.3.4	Movement of teacher estimates.....	244
6.3.5	Two students.....	245
6.4	Validity and reliability.....	265
6.5	Formative benefits of the questionnaire.....	267
6.6	Conclusion.....	268
7	DISCUSSION.....	270
7.1	Introduction.....	270
7.2	Teacher estimates.....	270
7.3	The questionnaire's construct.....	272
7.3.1	Introduction.....	272
7.3.2	Technical factors.....	273
7.3.3	Psychological Factors.....	277
7.3.4	A Model from the questionnaire.....	284
7.3.5	Autonomy and motivation.....	289
7.4	The questionnaire and its model's relation to themes in the autonomy literature.....	291
7.4.1	Autonomy and responsibility.....	292
7.4.2	Autonomy as capacity and behaviour.....	292
7.4.3	Autonomy and social interaction.....	293
7.4.4	Autonomy is variable.....	294
7.4.5	Self-evaluation in self-report autonomy questionnaires.....	297
7.4.6	Summary.....	300

7.5	What the questionnaire is measuring.....	300
7.6	Suitable uses for the questionnaire	304
7.7	Translating questionnaires	308
7.8	Conclusion.....	311
8	CONCLUSIONS	313
8.1	Introduction	313
8.2	Answers to research questions.....	313
8.2.1	Research question 1.....	313
8.2.2	Research question 2.....	315
8.3	Contributions to knowledge	317
8.3.1	In the area of language learner autonomy theory.....	317
8.3.2	Innovative approach to researching autonomy	319
8.3.3	Translation of questionnaires	323
8.4	Limitations.....	324
8.5	Implications	325
8.6	Further research	327
8.7	Concluding remarks.....	329
9	REFERENCES	332
10	APPENDICES	350
10.1	Full 256 items with back translation.....	350
10.2	Long List areas covered.....	358
10.3	Long List format C with translation	369
10.4	Selection Table	371
10.5	Movement of teacher estimates	382

LIST OF TABLES

Table 2.1: Three-stage model of autonomy (Breen & Mann 1997: 143)	31
Table 2.2: Three-stage model of autonomy (Macaro 1997: 170–172)	32
Table 2.3: Levels of implementation of autonomy (Nunan 1997: 195)	33
Table 2.4: Levels of autonomy (Littlewood 1996: 429-430)	34
Table 2.5: Locus of Control (based on Williams and Burden 1997: 101)	43
Table 2.6: The Aspects of Metacognitive Awareness (Sinclair 1999: 102)	73
Table 2.7: Linguistic evidence for metacognitive awareness (Sinclair 1999: 104)	74
Table 2.8: Frequent concepts in descriptions of autonomy, with references	79
Table 3.1: Top 20 countries of origin for non-UK students at UK higher education institutions, 2008-2009	110
Table 3.2: Countries of origin of respondents	110
Table 3.3: The sample groups involved in the research and the format of items used	112
Table 3.4: Questionnaire development stages	118
Table 3.5: The formats of the instrument used in the research	119
Table 3.6: Comparison of Likert item responses	131
Table 3.7: Overview of the actual stages of research as carried out	151
Table 3.8: Original outline plan for the research and subsequent changes	154
Table 4.1: Comparison of item wording guidelines	158
Table 4.2: The general areas for questionnaire items with thesis sections	167
Table 4.3: Participants in HCT UAE data gathering	168
Table 4.4: Participants in English Language Internet group	171
Table 4.5: Participants in Mandarin language Internet group	173
Table 4.6: Participants in BNU group	176
Table 5.1: Summary of all items queried by respondents	180
Table 5.2: Short List's coverage of areas of autonomy	191
Table 5.3: Comparison of coverage of the areas of autonomy in the Long List and Short List	192
Table 5.4: Participants in ELTCS third year group	198
Table 5.5: Participants in ELTCS preessional group	200
Table 5.6: Participants from other preessional groups	201
Table 5.7: Participants in Taiwan group	202
Table 6.1: The 12 runs of factor analysis	208
Table 6.2: Loadings from factor analysis run eight	210
Table 6.3: The rejected items and scales with reasons for exclusion	211
Table 6.4: Final groups for factor analysis run eight with suggested identifications	212
Table 6.5: Combining the 12 factor analyses	216
Table 6.6: Identification of factors	219
Table 6.7: Possible matches of questionnaire factor groupings with Cotterall (1995)	223
Table 6.8: Comparison of input and output groupings	226
Table 6.9: Comparison of items in Short List and factor analysis groupings	229
Table 6.10: Significant correlations of categories to Categories Average in descending order (standardised pooled data)	230
Table 6.11: Significant item Correlations with Categories Average in all data combined in descending order of Pearson's correlation	231
Table 6.12: Item scores (%) in each category (combined standardised data)	233
Table 6.13: Significant correlations between Teacher Estimate and the categories and Categories Average (ELTCS year-3 and ELTCS preessional P2)	236
Table 6.14: Significant correlations of categories to Categories Average arranged in descending order (groups ELTCS year-3 and ELTCS preessional P2)	237
Table 6.15: Significant correlations of individual items to Categories Average arranged in descending order (groups ELTCS year-3 and ELTCS preessional P2)	238
Table 6.16: Significant correlations of items to Teacher Estimate using Spearman's r	239
Table 6.17: Comparison of ELTCS year-3 students' questionnaire results with teacher estimates of autonomy and confidence	242
Table 6.18: Confidence categories in ELTCS year-3 teacher's interview	243
Table 6.19: Questionnaire results for students T1-St-b and P-St-a	246
Table 6.20: Average questionnaire results for ELTCS year-3 students with T1-St-b's results	247

Table 6.21: Common terms associated with learning English from T1-St-b’s interview, with frequencies and frequency order	248
Table 6.22: Internal reliability of scales	266
Table 6.23: Respondent feedback from data gatherings	267
Table 7.1: Inter-Category correlations	272
Table 7.2: Possible equivalences between the questionnaire, Fazey & Fazey, and SDT	291
Table 7.3: Classification of translation problems in the Long List with examples	309
Table 10.1: Full 256 items with back translation	357
Table 10.2: The Long List items with areas covered.	368
Table 10.3: Long List Format C with translation	370
Table 10.4: Selection table	381

LIST OF FIGURES

Figure 2.1: A “spectrum” of autonomy	22
Figure 2.2 The major influences on language learning autonomy (Benson 2001)	22
Figure 5.1: Process for combining lists	186
Figure 7.1: Self-Concept	281
Figure 7.2: Autonomy seen as an electric circuit	285
Figure 7.3: A model of Questionnaire Autonomy based on the questionnaire category correlations (N.B. Arrow thicknesses indicate strength of correlation)	287
Figure 7.4: Littlewood’s (1996: 430) model of the components and domains of autonomy	288
Figure 7.5: Example of the hypothesised dip in student questionnaire results	298
Figure 10.1: Categories Average scores plotted against Teacher Estimates	382
Figure 10.2: Social Comparison scores plotted against Teacher Estimates	383
Figure 10.3: Linguistic Confidence scores plotted against Teacher Estimates	384
Figure 10.4: Information Literacy scores plotted against Teacher Estimates	385
Figure 10.5: Locus of Control scores plotted against Teacher Estimates	386
Figure 10.6: Metacognition scores plotted against Teacher Estimates	387
Figure 10.7: Self-Reliance scores plotted against Teacher Estimates	388

ACKNOWLEDGEMENTS

This thesis is dedicated to the memory of my father Derek Dixon.

Many thanks go to all those who have helped in this research, some of whom are named below.

My supervisors:

- Dr Ema Ushioda
- Dr Richard Smith
- Professor Hilary Nesi

The translators:

- Xinjie Guo
- Dr Shi Lijing
- Dr Jiang Xiaoli
- Wei Yi-Chun (Sherri)
- Dr Fei-Yu Chuang

Those who helped with supplying data:

- Anna Calvi
- Professor Wang Qiang
- Wei Yi-Chun (Sherri)
- Dr John Grainger
- Dr Bill Lex
- Lucy Cooker
- Heather Baba
- Rod Revell
- Virginia Robson
- Dr Gerard Sharpling
- Tilly Harrison
- Anny King
- Sian Khouri
- Professor Helen Spencer-Oatey
- Liu Wei

I would also like to thank all the students who took part in this study.

DECLARATION

I declare that this thesis represents my own work, except where due acknowledgement is made. It has not been previously included in a thesis, dissertation, or report submitted to this university or any other institution for a degree, diploma, or other qualifications.

Signed _____

David DIXON

ABSTRACT

The thesis aims to explore the viability of using a quantitative instrument to measure language learner autonomy and investigate whether such an instrument has a function in supporting teachers and learners in the development of learner autonomy.

The research developed into a critical reflexive approach which probed the theoretical and design issues surrounding the development of a quantitative autonomy-measurement instrument by actually attempting to produce such an instrument. This approach means that I could experience and examine first-hand the theoretical and practical issues which the quantified measurement of autonomy would involve.

The main conclusions of this research were, firstly, that the aim of measuring learner autonomy needs to be recast in the light of the research which indicated that it is necessary to understand autonomy as a quality which has only an abstract existence if it is not instantiated in a context. This means that the aim of producing an instrument which measures an abstract universal learner autonomy cannot be achieved. However, such an instrument can be used to monitor learners in autonomy-relevant areas and can serve a useful purpose in scaffolding the learners in their environment in order to facilitate the dialogue which enables a teacher to support the learners better in the development and maintenance of their autonomous learning. Secondly, teacher estimates of their learners' autonomy can be complemented and assisted by using the data provided by the quantitative instrument developed in this research.

Another outcome was that the translation of instruments in second language teaching research is an issue which needs to be given more serious consideration and should be carried out in a more principled way than it is currently.

ABBREVIATIONS

BA	Bachelor of Arts
BNU	Beijing Normal University
CA	Cronbach's alpha
CEPA	Common Educational Proficiency Assessment
CFA	Confirmatory Factor Analysis
CUP	Cambridge University Press
DI	Discrimination Index
EFA	Exploratory Factor Analysis
ELT	English Language Teaching
ELTCS	English Language, Translation, and Cultural Studies
ESL	English as a Second Language
GNVQ	General National Vocational Qualification
HCT	Higher Colleges of Technology
HE	Higher Education
HESA	Higher Education Statistics Authority
HKUP	Hong Kong University Press
ILC	independent Learning Centre
KMO	Kaiser-Meyer-Olkin measure of sampling adequacy
KSA	Kingdom of Saudi Arabia
L1	First Language
L2	Second Language
OUP	Oxford University Press
P	Per cent
P1	Preessional ELTCS group first administration

P2	Preessional ELTCS group second administration
R	Reverse Coded
SAC	Self-Access Centre
SDT	Self-Direction Theory / Saudi Development and Training
SILL	Strategy Inventory for Language Learning
SPSS	Statistical Package for the Social Sciences
T1	Third year ELTCS group
UAE	United Arab Emirates
UK	United Kingdom
ZPD	Zone of Proximal Development

1 INTRODUCTION

1.1 *Background: Initial motivations*

The roots of the initial motivation for this research go back to the 1990s when I was working for Saudi Development and Training (SDT) as the self-access centre (SAC) coordinator. SDT management required a quantification of the functioning of the SAC so that it could be given targets and its performance monitored. Gains in the learning skills and abilities of the students would be recorded using a General National Vocational Qualification (GNVQ) in the key skill area of Improving Own Learning and Performance. This proved to be very labour-intensive, requiring much time for the training of teachers as assessors and for the paperwork required for a formal certificate. This experience suggested to me that it would be advantageous to find an alternative method. Later, working at the Fujairah Women's College, part of the Higher Colleges of Technology (HCT) in the United Arab Emirates (UAE), I was called upon to justify the independent learning centre by showing a gain in the autonomy of the learners. There was not an instrument available to do this in the quantified way that was requested. For such an instrument to be useable at the HCT it would need to be suitable for use by non-experts in autonomy, would need to be suitable for use in existing courses without upheavals, be quick to complete for learners, and not make unreasonable demands on teacher time. It would need to be easy to distribute around the widely scattered campuses of the HCT and would also need to be reliable and valid.

Benson (2001: 186) states that “there is surprisingly little empirical evidence available for the effectiveness of any particular approach” and that there is no practical tool with which to contribute evidence. Perhaps, I thought, the lack of a tool

was due to practical difficulties in measuring autonomy or perhaps it was because the measurement of autonomy is inappropriate conceptually. When I began to look into the possibility of designing such an instrument it became clear that it was not an easy task as there were many issues around the idea of measuring autonomy, and I became interested in exploring the possibility of a convenient autonomy measure in order either to produce one, or to satisfy myself that it was not practicable.

1.2 Direction

Initially the main aim was to investigate whether a closed-item autonomy-measuring instrument was possible, and in the course of the research as my knowledge and understanding of the issues developed the nature of the investigation became clearer and the research became more reflexive and critical. The aims developed into a focus on whether a quantitative instrument could emulate teachers in their estimates of their learners' autonomy. This was fruitful in two main ways: it provided a way of indicating the practical value of the instrument as an alternative or complement to what teachers were already doing (which was estimating their learners' autonomy in an informal way); and it eventually led me to a new understanding of what the purpose of an autonomy measure should be, in sum, I realised that what was really valuable in the research was not the autonomy measuring aspect but the aspect of helping teachers to help learners with the development of their autonomous learning.

Difficulties were encountered with finding sufficient subjects to provide the data necessary for quantitative analysis, but the more qualitative and small scale aspect of the research proved fruitful. A difficulty also emerged in the area of translation of the instrument but this, however, led to an interesting consideration of the role of translation in questionnaire research. The journey was thus complicated and

exploratory, with practical problems and “dead ends” along the way. I have tried to represent this journey in the thesis by making it broadly chronological and developing the ideas to reflect how my own thinking developed. The researcher at the outset of this project is not the same person as the researcher at the end.

1.3 Overview

This thesis is presented in primarily chronological order, beginning with the concrete but undeveloped idea presented in this introduction, moving to consideration of theory and then to attempting to develop the practical measure and analyse its performance, followed by the interpretation of the evidence, but with review and development of my understanding of the problem feeding back into the process.

I begin by putting this investigation into its wider context with a review of the literature on measuring language learner autonomy (Chapter 2) which raises the issues connected with the idea of autonomy measurement and discusses how other researchers have approached the problem. Elements of autonomy are discussed in order to establish the areas which the autonomy-measuring instrument should cover and this informs the choice of questionnaire items. I also consider criticisms of the idea of measuring autonomy.

Chapter 3 discusses the methodological and theoretical underpinnings of the research. Here I present my research aims, i.e. to investigate the viability of an instrument by attempting to design one and examine its validity. In the chapter I also address the issue of translation. I present an overview of the stages of the research, both as initially envisaged and in its final form. I present the methods of statistical

analysis which would be necessary to probe the construct represented in the questionnaire and to establish its reliability and validity as a model of autonomy.

Chapter 4 concerns the first stages in designing the instrument under investigation in the research, including writing the items to address the areas of autonomy established in Chapter 2.

Chapter 5 looks at the item reduction process and why it differed from the originally intended method (factor analysis). This chapter describes the small scale data collection which took place, and explains how large scale research was not possible due to delays caused by slow returns and translation problems. The research had to adapt to these problems, which meant that gathering statistically significant data was limited and indications from smaller-scale samples were used for illustrative and qualitative purposes. The data gathering using the shorter list of items is described.

Chapter 6 presents the data and analyses the patterns emerging the comparison of questionnaire data with teacher estimates and a more detailed examination of two respondents in relation to their questionnaire returns. In Chapter 7 the indications of the teacher estimates (which involved two teachers, including myself) are discussed, and the construct embodied in the questionnaire is presented and examined. The progress of the research is discussed and questions about the function of the questionnaire are addressed.

Finally, in Chapter 8 the contribution of the research is considered, the research questions are answered, possible future uses of the instrument are suggested, and directions for future research are proposed. The implications of the research are discussed.

The thesis is exploratory, and describes the journey which I took in developing the initial idea of a practical measure of autonomy, from initial motivations, through researching the field, investigation of possible approaches to measurement, developing an instrument, and gathering data to probe the instrument. It also presents the problems encountered and the limitations which they imposed.

1.4 Why the research is important

The thesis is worthwhile because it addresses, in innovative and flexible ways using mixed methods, a question which is much asked: can autonomy be measured? It is an area which is discussed theoretically but less often are attempts made to establish empirically the viability of a quantitative instrument. Through the critical appraisal of such an instrument it was hoped to be able to explore the problem from an original perspective which could shed new light on the question in a fresh way. The account of my own development in the course of this research will I hope be useful for other teachers and researchers involved in the challenging area of language learner autonomy.

2 LITERATURE REVIEW: MEASURING LEARNER AUTONOMY IN ADULT SECOND LANGUAGE LEARNING

2.1 Introduction

The subject of the present research is measuring learner autonomy in second language learning and therefore it is essential to review the relevant literature in order to establish clearly where the present research fits in the context of previous thought and research. It is necessary also for defining what the research will and should aim to do and to clarify the meanings of the concepts which will be used in the methodology and in the subsequent stages of analysis and discussion.

2.1.1 Aims of the Literature Review

The main purposes of this literature review are:

- to review the literature for evidence to establish that there is a need to measure autonomy;
- to review the literature to establish that it is desirable to measure autonomy;
- to establish whether it is theoretically possible to measure autonomy;
- to examine possible ways to measure autonomy;
- to describe and discuss previous attempts at measuring autonomy, and;
- to prepare an initial selection of the areas to be covered in the “Long List” of candidate items for a future autonomy measuring instrument.

In order to arrive at a conclusion as to whether measuring autonomy is theoretically possible arguments both for and against will be presented and evaluated in the context of the aims of the present research.

The findings of this Chapter will be used in the decisions regarding the design of the instrument to measure autonomy (whose feasibility is being investigated in the present research) and also in the initial selection of items to populate a preliminary long list (see Section 4.2.3) of items intended to represent all the relevant elements of autonomy as a starting point for subsequent data reduction by statistical methods to obtain the most important items to retain for an eventual autonomy measurement instrument. The findings will also inform the discussion of the eventual results of the present research and the conclusions about the possible uses and limitations of an instrument to measure autonomy.

In this Chapter I will not include detailed discussion about measuring possible learning gains resulting from autonomy or on the literature concerning the effectiveness of differing types of independent learning schemes, such as the research into the effectiveness of self-access centres (e.g. Morrison 2005; Reinders and Lázaro 2007; Gardner and Miller 1999: 205-240). This is because the aim of the present research is to establish whether measuring autonomous learning is feasible, and not to establish the effectiveness of individual autonomous learning schemes compared to other schemes or other ways of learning.

The answers to four primary questions which are essential for shaping the nature and direction of the present research need to be found in the literature. The questions are:

1. Is there a need for a measure of autonomy?

2. Should autonomy be measured (i.e. is it desirable)?
3. Can autonomy be measured?
4. Does the literature provide clues as to how autonomy can be measured?

To provide the answers to these questions it will be necessary to consider the views from the literature concerning the nature of autonomy.

2.1.2 Use of the term “Measure”

At the outset it is necessary to clarify how the term *measure* (and its derivatives: “measuring”, “measurement” etc.) will be used in this thesis. A key point which needs to be emphasised here is that the term “measure” (and its other forms) has been very deliberately chosen for use in the present research. I make a distinction between “measure” and other related terms such as: “test”, “assess”, “judge”, and “evaluate”. The word “measure” is the most appropriate for the present research because it expresses the idea of quantification, but with a more neutral connotation than the other words which suggest determining value, how good or bad something is, or how satisfactorily it is performing. I am aiming to design an instrument which is a quantification tool as distinct from a test. It is hoped that it will have many useful functions, but testing is assuredly not one of these. As will be seen in Section 2.3 below, the distinction is crucial to the possibility, and indeed desirability, of measurement in the area of autonomy.

Very often the literature on the measurement of autonomy assumes a situation involving formal high stakes *testing* where the results will be used to grade students. Seldom considered are situations of, for example, self-measurement by learners, measurement for research purposes, or measurement for needs analysis.

There are objections that testing autonomy is implicitly un-autonomous (for example, Benson 2010; Champagne, Clayton, Dimmitt, Laszewski, Savage, Shaw, Richmond, Thein & Walter 2001) because it opposes learner choice, but with the clarification of the term “measure” distinguishing it from “test”, “evaluate”, etc. these do not actually apply. This is a key point for the present research which is focused on the measurement of autonomy but not with testing or evaluating it. When the distinction between measurement, evaluation, testing etc. is made it becomes clear that different research reported in the literature can be characterised differently as concerning predominantly measurement or testing or evaluation. For example, in Dam’s (2000) paper on evaluating autonomy there is no focus on either measurement or testing, but it is aimed entirely at evaluation. At what might be seen as the other extreme Ravindran (2000) is concerned with testing and certification of autonomy based on evaluations, but with no attempt at measurement.

In the field of autonomy in language learning both more qualitative (e.g. Dam 2000) and more quantitative (e.g. Cotterall 1995) techniques have been used to investigate learners’ levels of autonomy.

Testing and measuring are associated with quantitative techniques (Dörnyei 2007: 32-34). *Measuring* and *testing* both suggest quantitative techniques, whereas *evaluating* can suggest more qualitative means, including the involvement of someone who makes judgements regarding the subject of the evaluation. In the present research I aim to investigate the measurement of autonomy, which will therefore involve using quantitative measurement techniques. However, though both are quantitative, I do not see testing and measuring as the same. Measuring is the collecting of quantitative data, but formal testing is specifically the gathering of data

for purposes of judgement. In this thesis the focus is on the measurement of autonomy and therefore testing and its specific issues are not a central concern. To avoid any confusion this distinction between the terms *test* and *measure* will be observed throughout the present thesis.

2.1.3 Overview of the structure of the Literature Review

The four key questions introduced in Section 2.1.1 above will underlie the Literature Review. The first two questions regarding the need for a measure and whether it should be measured are dealt with in Sections 2.1.3 and 2.3. The conceptual issues regarding whether autonomy can be measured (question 3) are dealt with in Section 2.4 where the differing concepts of autonomy found in the literature will be examined through discussion of the key concepts found in the different conceptions of autonomy. This will serve as the basis for the “Long List” of items which will be candidates for inclusion eventually in an instrument to measure autonomy (see Section 3.9). The more practical challenges regarding the question of how to measure (question 4) are discussed by presenting previous research which has a direct bearing on the present project (Section 2.5). Here I will attempt to highlight the lessons to be learnt from the literature concerning levels of autonomy, practical problems of autonomy measurement, and how autonomy has been measured previously. The Discussion (Section 2.6) will summarise the issues and attempt a synthesis which will lead to my stance being stated and the four questions posed above being answered. Here I will highlight the points which will be important for the present research and especially for the methodology.

2.2 Is there a need for a measure of autonomy?

In this section I will review the literature and establish that there is an expressed need for a measure of autonomy which is not being fulfilled by presently available means (such as the measurement of learning gains). I will show that an autonomy measure would provide solutions to a number of common requirements.

There is a perception that autonomy lacks evidence to support its claims to provide educational advantages. This can be found among both supporters (such as Dickinson 1987; Sinclair 1999), and detractors (such as Hand 2006). Benson (2001: 54) makes the point that “If we aim to help learners to become more autonomous, we should at least have some way of judging whether we have been successful or not”.

There is therefore a perception that a measure of autonomy would be useful to measure changes in learners’ levels of autonomy.

If autonomy produces better learning then measuring learning gains using achievement tests would appear to offer an indication of autonomy level which would make a direct measure unnecessary. A major problem with this, however, is that achievement tests do not measure autonomy directly and therefore it cannot be assumed that results of tests are not influenced by other influences unconnected with autonomy. La Ganza (2002) sought to investigate the effect of autonomy on learning outcomes but found problems of attribution. He (La Ganza 2002: 47) makes the point that control groups cannot be used since if strict rules were laid down for the activities of two groups then it would no longer be a situation of autonomy. Morrison (2005) also encountered this problem of securely isolating the reasons for learning gain in his study on the evaluation of self-access learning.

Some views of autonomy have it as at least partly a disposition of the learner rather than entirely a product of a specific context (Little 1997; Carr & Claxton 2002). In this case autonomy would be transferable (to some degree) between settings and consequently transferability would be an important element of autonomy. A problem with methods which focus on learning gains alone is that they do not address the transferability (see Section 2.4.2 below for a discussion of transferability). A more direct autonomy measure would therefore be useful in avoiding the problems inherent with using learning gains as a measure of autonomy. According to Benson (2010: 78) it would also potentially allow researchers to investigate how autonomy interacts with different contexts of teaching and learning, and how it is transferred from one situation to another (Benson 2010: 85).

Benson sees a measure as potentially allowing researchers to identify the developmental processes of autonomy acquisition (2001: 51) and how it develops over time (2010: 78). This is an important area about which there are a number of differing theoretical models (e.g. Littlewood 1996; Nunan 1997; Breen & Mann 1997) which have great importance for learners and practitioners, but which have yet to be empirically tested and hence a measure, if possible, would have benefits.

Sinclair (1999: 100) makes the point that the lack of a measure or recognised framework for autonomy means that “Teachers, course planners and materials writers are left to do what they think is best, to rely on their own beliefs about learning, their values, experience and intuition”. Evidence for the presence of autonomy and to what degree it is present in individual learners and classes would help teachers to make better estimates of their learners’ autonomy. Benson (2001: 51) also states that “For the purposes of research and the evaluation of practice, it

would indeed be convenient if we had a reliable method of measuring degrees of autonomy”. A measure would also be of use in evaluating self-study systems or learner training initiatives, as proposed by Ravindran (2000).

Rivers & Melvin (1981: 90) maintain that it can be useful to know the “average” learning style for a class to match instructional activities to fit the strengths, weaknesses or “bias” of the class, and also the profile of the instructor. This will allow teachers to become aware of their own biases with respect to mode and style of presentation, and could prevent a drastic mismatch between the instructional method and the style of the ‘typical’ student in the class. While I do not intend to propose that autonomy is a learning style, the principle that knowing the average or bias of a class can prevent a “drastic mismatch” is an important idea in support of knowing the autonomy level of students as it would serve a needs analysis function. A measure of autonomy would therefore provide data which could be used for needs analysis, potentially providing evidence of learner beliefs and learner readiness for self-study (Cotterall 1995). Depending on the nature of an eventual instrument it may be able to provide different types of data, from a very basic indication of overall level to a detailed breakdown of levels in specific areas of autonomy. With a measure providing a basic level of information the data would serve as an indication of whether there was cause for concern and would, for example, enable a teacher to identify at-risk learners and devote time to establishing the precise nature of the problem through counselling, leading to the formulation of remedial action. If the instrument is capable of greater resolution then specific areas of strength or weakness would be indicated thus providing useful indications of areas on which the learner needs to focus. For example, metacognition is vital for autonomy (Flavell 1987; Victori & Lockart 1995; Wenden 1995; Vickers & Ene 2006) and hence appropriate

remedial learner training activities could be initiated. Likewise, if motivation was revealed to be a problem area the learner would be counselled and possible solutions discussed. In the case of a self-administered measure feedback and advice could be provided directly to the learner. Self-assessment is fundamental to metacognition and so to autonomy (Champagne et al. 2001; Rivers 2001; Reinders 2007). The feedback provided by an autonomy-measuring instrument could thus be used to support learners in their self-assessment. It appears, then, that an instrument to measure autonomy should not be aimed at a rather abstract concept of autonomy but should if possible be a functional and useful complement to currently available autonomy-supporting techniques.

A suggested way of evaluating learner autonomy is the learner diary or logbook Dam (2009: 139). Nunan sees these as providing “insights into processes of learning which would be difficult, if not impossible, to obtain in any other way” (Nunan 1992: 123). However, logbooks are not always clear, they depend on the level of the learner, the learner does not automatically cover all the relevant areas, and indeed may not be able to (Dam 2009: 139). Tsang (2005) found that learners were divided about how helpful they found logbooks, and Chan, Spratt & Humphreys (2002: 10) found that 80.8% of students rarely or never write a diary to help their studies. It appears, therefore, that to use diaries or logs to evaluate autonomy it would be necessary to make them compulsory; Dam (2009: 134) says that her students are expected to use the logbook. Blin (2005: 101-103) has actually used diaries to assess learners’ levels of autonomy, though, again it was compulsory and in addition had the drawback of being a long-term commitment taking 12 weeks, which would make it impossible to use on shorter courses and would make it impractical for use in initial needs analyses. There is, then, a need for some other type of measurement tool

which can aid evaluation (or measurement). Dam (2000) also uses questionnaires to contribute to evaluations, and this would be quicker than the logs. It would also be formative, i.e. the learners will be thinking about areas they may not have considered before and this will potentially feed back into their reflections and logbook considerations.

There is, therefore, evidence from the literature that an autonomy measure would be useful for researching the nature and development of autonomy, for assisting teachers, planners and materials developers, and for supporting learners.

2.3 Should autonomy be measured?

The previous section has shown that there is evidence of a need for an autonomy measure, but some feel that autonomy should not be measured. This can be based on feelings of the type that “traditional assessment has been a major force in retarding educational reform” according to Reeves & Okey (1996: 192); or that testing interrupts learning (Zimmerman 1995); or Benson’s (2010: 78) more vague “nagging feeling that this was not perhaps the ‘right’ way to think about autonomy”. In this section I will look at arguments against measuring it.

Benson (2010: 95) considers the “likely consequences of ‘autonomy testing’”. He feels that learners faced with a test of autonomy would try to achieve high grades. This would clearly be an external motivation rather than the internal one appropriate to autonomy. There could be a “mask of autonomous behaviour” (Breen & Mann: 1997) rather than true autonomy. This argument is specifically aimed at autonomy *testing*, and as such is not relevant to the present research which (as discussed in Section 2.1.2 above) is concerned with measuring autonomy and specifically

excludes testing. Another argument is based on autonomy being seen as fundamentally involving free choice. Baumann (2007) finds that in his situation the development of autonomy is constrained by the imposition of learning outcomes. If an autonomy measure were imposed on learners it would limit autonomy, and Champagne et al. (2001: 49) do see it as anti-autonomous to measure autonomy. This argument assumes, however, that the measurement will be imposed on the learner. It would be, in effect, to try to measure something while at the same time taking that very thing away. However, this is not an argument against measuring per se but rather against a situation where the learner does not have a free choice to have or not to have the measurement, as might be the case with an institutional test. As discussed previously (Section 2.1.2 above), the distinction between *measure* on the one hand and *test/assess/evaluate* on the other is crucial to the aims of the present research. When the measurement is not a test and where the learner has chosen to measure/have measured his or her autonomy this type of objection is no longer relevant. This clarification avoids many of the objections to the measurement of autonomy found in the literature.

Another objection to measuring autonomy is that it is more for the benefit of the global economy than for the individual learner (Benson 2010). This is attached to a fear that autonomy is becoming popular not for its intrinsic benefits but for its perceived benefits to employers who want a flexible workforce with the ability to learn new skills rather than merely having a fixed body of knowledge (Pemberton 1996: 1). However, as discussed in the previous section (Section 2.1.3 above) there are arguments in favour of measurement, and so to prevent autonomy measurement evidence of disadvantages linked to measurement would need to be presented, and it would need to outweigh the advantages to the learner.

There is also a suspicion that the desire to measure autonomy is motivated not by the desire for any real benefits this may offer, but because there is an attitude that if something cannot be measured it is not a worthwhile goal (Benson 2010: 78). In a similar vein, Biesta (2009) feels that values must not be lost sight of with “The rise of the measurement culture in education” (2009: 34). Clearly, motivations for measuring autonomy will differ between individuals, but the present research is motivated by the potential benefits and hence would count as “measurement of what we value”. Potential beneficiaries of a measure would be, for example: learners wishing to measure their autonomy for self-evaluation, teachers wishing to measure their learners’ autonomy for needs analysis, and researchers wishing to measure autonomy to find out more about its development.

In conclusion, I would argue that the key to maintaining autonomy while measuring it is to ensure that the measure is not imposed on the learner.

2.4 Aspects of autonomy

A key aim of the present research is to explore the question “Can autonomy be measured?”. In order to proceed with this it has been necessary to clarify the term *measure* (Section 2.1.2 above), and in this section I will review the literature concerning the definition of autonomy and extract from this the basis of a rationale for measurement which provides the elements a measure will have to cover. This section of the Literature Review provides the areas which will be used to inform the selection of the putative autonomy-measuring questionnaire’s items.

This section highlights and discusses nine key recurring ideas and themes found in the language learner autonomy literature. These areas can be seen as the areas which

are most commonly found across the spectrum of the autonomy debate and which are the most important areas for an understanding of the issues current in the field. The nine ideas are:

- Autonomy is a multidimensional concept
- Autonomy is variable
- Autonomy is a capacity
- Autonomy is demonstrated
- Autonomy requires metacognition
- Autonomy involves responsibility
- Autonomy involves motivation
- Autonomy involves social interaction
- Autonomy is political

In the following subsections each of these areas will be discussed in turn (though they are in fact all interrelated).

2.4.1 Autonomy is a multidimensional concept

In order to measure autonomy a definition would appear to be necessary. However, despite numerous attempts (for example Holec 1981; Dickinson 1987; Little 1991; Littlewood 1999), a single agreed definition remains elusive. Authors still write of it as being a concept which is “fuzzy” (Dam 2000: 59) or “seemingly abstract” (Smith 2008: 395). Problems of definition led La Ganza (2002: 51) to change his research focus away from quantifying autonomy, and others have found it necessary to formulate their own definitions to allow their research to proceed (e.g. Cotterall 1995).

There is a range of definitions of autonomy displaying “notable semantic variations” (La Ganza 2002: 47-48) which reflect different stances and theoretical camps, and as

yet there is no universally established definition, taxonomy, or terminology for autonomy in the field of language learning. At one extreme (influenced by constructivist approaches to learning) there is the view that all successful learning is by definition autonomous (Benson & Lor 1998: 12), while at the other end of this spectrum are ideas that autonomy is an unattainable ideal, which prompts Sinclair (2000: 6) to speak of teachers being “put off by what seems to them to be a highly unrealistic, and unachievable goal in their own contexts”.

One reason for the difficulty of definition is that autonomy is considered to be multidimensional (e.g. Hurd 2004; Benson 2001: 51; Blin 2004). Shaw (2008: 188) writes of a “conflation of means and goals”, where he feels that a set of techniques has become attached to the pursuit of autonomy which has become confused with autonomy itself and which leads to what he views as a lazy way of speaking, for example, “putting learner autonomy into practice”. Another reason is that there are many autonomy-related terms and they have been used inconsistently. For example Holec essentially defines autonomy as “the ability to take charge of one’s own learning” (1981: 3) and he sees autonomy as a capacity of the learner, but for Dickinson (1987: 11) autonomy is “the situation in which the learner is totally responsible for all of the decisions concerned with his learning and the implementation of those decisions”. Most authors follow Holec’s (1981: 3) definition of autonomy, but this still leaves room for “a good deal of interpretation” (Martinez 2008: 105).

Definitions of autonomy can vary because they reflect the author’s stance on autonomy, providing a plethora of dimensions to autonomy. Some authors have attempted to identify the different stances or types of autonomy with labels such as:

psychological, technical, sociocultural, political/ideological, liberal progressive, humanistic, behavioural, economical, etc. Crabbe (1993: 443) for example finds in the literature three ways of viewing autonomy: psychologically, economically, and politically/ideologically. La Ganza (2008: 65) distinguishes four dimensions of autonomy:

- Political – learners taking control of learning (Benson 1997; Pemberton et al. 1996)
- Liberal progressive – learners taking responsibility (Holec 1979; Kohonen 1992)
- Behavioural notion of strategy development (Wenden 1987; Dickinson 1992)
- Humanistic self-initiation or self-direction (Rogers 1961; Kenny 1993; Savage & Storer 1992)

Benson (1997: 25) found three basic versions of definition: the technical version (where autonomy is the act of learning on one's own and the technical ability to do so); the psychological version (where autonomy is the psychological capacity to self-direct) and; the political version (where autonomy is control over the content and processes of one's own learning).

Oxford (2003) sees the situation as “far from coherent” and as “beset by conflicting ideologies, roiling inconsistencies, and fragmentary theories” (Oxford 2003: 75). In response to this she attempts to organise and even integrate the different types of definition by seeing them not as conflicting definitions but as perspectives which can coexist. The four classifications she finds are: the technical perspective (the physical situation); the psychological perspective (the characteristics of the learners); the socio-cultural perspective (mediated learning); and the political-critical perspective (an ideological view concerning empowerment).

In Benson's technical variety there is the view that skills, techniques, and strategies (which can be taught) can be used to pursue autonomy. Oxford (2003) has a similar perspective included in her classification; however, this includes a strong emphasis on the physical situation being one in which the reins of power have been handed over to the learner.

Benson's second variety, the psychological version, embodies constructivism, the theory of learning where knowledge is seen as personally constructed. Oxford's (2003) Psychological Perspective corresponds to this. Learning is seen at the level of the learner where motivation, learning styles, and learning strategies are normally found. There is a strong affinity with work in psychology, such as self-efficacy (Bandura 1997) and attribution theory (Weiner 1979). Oxford does not include in this the idea of mediated learning, but includes it in her "Sociocultural Perspective".

Benson's political version sees the construction of knowledge as dependent on prevailing political and social ideologies, with issues of power relationships and rights (see Section 2.4.9). There are strong connections with cultural issues such as the appropriateness of learner autonomy to non-Western contexts. What is appropriate in the West should not be assumed to be equally appropriate worldwide, and beliefs about the nature of autonomy may need to be changed in the light of differing views in different cultural contexts. Oxford (2003) differs from Benson by distinguishing and giving a separate section to what she calls the "Sociocultural Perspective", where she places ideas, largely influenced by Vygotsky, that autonomy involves socially mediated learning.

Thus, there is a spectrum of versions of what autonomy is which are based on differing theoretical arguments (see Figure 2.1 below).

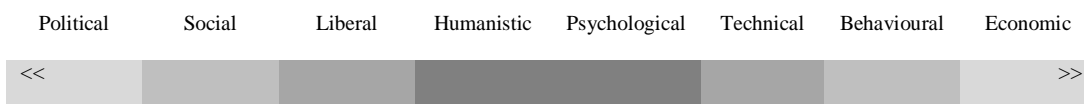


Figure 2.1: A “spectrum” of autonomy

In this spectrum political and economic are presented as being on the margins and humanistic and psychological as being the core ideas, but other “colours” could equally well be placed more centrally. The positioning of the “colours” will naturally reflect the author of the spectrum’s favoured view of autonomy.

Gremmo and Riley (1995: 152) stress the eclectic origins of autonomy, and Benson (2001: 22) gives five major sources of theory which have influenced the thinking on autonomy in language learning (see Figure 2.2 below).

From field	Concept/Notion	
Psychology of Learning	Constructivism	} Autonomy in Language Learning
Educational Reform	Freedom in Learning	
Adult Education	Self-Directed Learning	
Political Philosophy	Personal Autonomy	
Language Learning	Focus on Learner	

Figure 2.2 The major influences on language learning autonomy (Benson 2001)

In addition to the problems of defining autonomy due to the different underlying philosophies there are terminological problems. For example, Holec distinguished between *autonomy* and *self-directed learning*; autonomy is the capacity goal, the ability to self-direct, and self-direction is the way of learning produced by having, or being on the way to having, autonomy (Pemberton 1996: 2-4). However, this distinction is not uniformly observed in the literature. If there are different versions of autonomy, what does it mean to use the term “autonomy” as if there were one underlying concept? Although hard to define, autonomy has come to be seen as an unquestioned universal moral good and Shaw (2008: 188) says it is a goal of education; however, to take only one small example, Holliday (2003) sees autonomy

as a pre-existing social phenomenon. It seems that autonomy is just the pedestal on to which can be placed a large number of different sculptures. Lamb & Reinders (2006: vii-xi) believe that it is not important to be constrained by a particular definition of autonomy in language learning, but rather to look at individual contexts. Holec (2008: 4) believes that it is necessary to stop searching for “monolithic and stable answers”.

Benson (2001: 44) writes that “People value personal autonomy for its own sake, and for this reason, it is not simply instrumental in the achievement of well-being, but an aspect of well-being deserving of protection in its own right”. Both political and more scientific advocates will be able to agree with a view that autonomy has intrinsic value. The differences between the poles would be paradigmatic, i.e. the more technical side *aims* to be detached and the political side *aims* to be involved. Both sides would, however, be basing their views on values. This might be expressed as a cline with at one end those whose values lead them to feel that autonomy is more usefully seen in terms of the capacities and psychology of individual learners, to at the other end those who feel autonomy has to be overtly seen as the right to challenge and struggle for democracy.

The problem for measurement can be illustrated by considering an example drawn from Nicolaides (2008), who reports on an ethnographic research project carried out on future English teachers into learners’ perceptions of their roles. The example of one of the subjects, Otávio, illustrates an area of possible difficulty for the present project (Nicolaides 2008: 154-155). The problem is that depending on which dimension of autonomy is considered the measure would give different, even contradictory, indications. Using Benson’s (1997) technical and psychological

versions of autonomy, Otávio's independence is high: he has the ability, will and motivation to learn well by himself. Viewing autonomy from the social interaction perspective, i.e. joining in with the group, his autonomy is quite low as he shows a lack of social responsibility by not getting organised for required class activities. According to the political view his autonomy is low as he does not adapt to the context (the class) and therefore does not "exercise his most important right – the right of learning also in the classroom" (Nicolaidis 2009: 154). This appears to have discouraging implications for a search for a single instrument to measure autonomy. It appears inevitable that such an instrument would need to find a definition of autonomy which is measurable, but this would be to confuse what is measurable with what is valuable (as discussed in Section 2.3 above).

Some researchers attempting a measure provide their own definitions of autonomy which they then use to operationalise the concept. For example Cotterall (1995: 195) characterises autonomy as "the extent to which learners demonstrate the ability to use a set of tactics for taking control of their learning". This strategy of pre-defining autonomy reduces the abstraction of the concept and can also focus on the aspects of autonomy which are more accessible to measurement, making it possible to support a particular research design. It has the disadvantage of not fully characterising the concept to everyone's satisfaction and of redefining for pragmatic rather than theoretical reasons.

Multidimensionality may mean that autonomy is not one thing and may not be accessible by means of a single quantitative measure. The multidimensionality as seen in the example of Otávio means that a learner's autonomy may be manifested differently for different dimensions.

2.4.2 Autonomy is variable

In addition to variations due to types of autonomy, individual learners' levels of autonomy may vary with circumstances, such as different tasks. Benson (2001: 47), for example, has stated that autonomy is “a multidimensional capacity that will take different forms for different individuals, and even for the same individual in different contexts or at different times”. Little (1991: 4) points out that an individual's level of autonomy can also “take numerous different forms, depending on their age, how far they have progressed with their learning, what they perceive their immediate learning needs to be, and so on”. Levels of autonomy may also depend on the nature of different tasks and “the learner who displays a high degree of autonomy in one area may be non-autonomous in another” (Little 1991: 4). There are many other variables which may influence the degree of autonomy, such as:

affective factors (e.g., mood), environment (e.g., noise, temperature), physiological factors (e.g., tiredness, hunger), motivation (e.g., attitude towards the task, the subject matter, the teacher, materials, co-learners) and so on. (Sinclair 2000: 8)

Considerations such as these would appear to make attempts at measuring autonomy inherently unreliable at best and at worst reductivist.

A further problem according to Benson (2001: 53) is that we know very little about the stages in the development of autonomy except that they are highly variable and uneven. Breen & Mann (1997) for instance theorise that learners react against the introduction of autonomous learning as part of its development. Consequently “a snapshot of the learner's performance at any given moment in time may give a misleading picture” (Benson 2001: 54). This is a question relating to the reliability of tests or assessments, especially high stakes ones, and is not peculiar to attempts to

measure autonomy. The reliability of any instrument will affect the conclusions that can safely be drawn based on its data. Benson seems to be referring to more formal testing or assessing, but I feel if the instrument is not to be used for testing but, for example, as a formative spur to reflection, this problem is much less relevant, as long as the limitations of snapshots are not ignored.

Learners' levels of autonomy may vary with motivation which will affect their willingness to take on the responsibility of autonomy. As Sinclair points out (2009: 185) a learner may have ample capacity for autonomous learning but not have the will to operationalise it, since:

The willingness to take control varies from time to time and task to task, depending on a range of variables, including psychological (e.g. depression, irritation), physiological (e.g. headaches) and contextual factors (e.g. too much noise, not enough resources) which can influence learners at any time.

If autonomy is viewed as a capacity, the measurement of autonomy would be inconsistent and would misrepresent the learner's capacity if it were based on the observation of the learner when not willing to deploy the capacity (see Section 2.4.4 below for a discussion of the question of the observation of autonomy). This indicates that measurement of autonomy would require the active volition of the learner, which would be problematic in the case of a test (see Section 2.1.2 above), but would appear far more achievable if the measure were voluntary.

Context is an often cited variable affecting autonomous learning and hence can introduce variations in apparent levels of autonomy. If autonomy is dependent on context the learner's level of autonomy may be more closely linked to the learning context rather than to the qualities of the learner, as Carr & Claxton (2002: 12) note

“the manifestation of learning dispositions will be very closely linked to the learning opportunities, affordances and constraints available in each new setting”. Dickinson (1987: 11) defines autonomy with reference to situation. Consequently an instrument designed to measure the learner’s autonomy may in fact be measuring the learning environment. However, autonomy is also seen as something the individual carries between different situations, i.e. autonomy is transferable; in fact this is frequently cited as one of its defining characteristics.

Holec (1981) sees autonomy as a capacity of the learner, and Carr & Claxton (2002: 12) describe it as “A tendency to respond or learn in a certain way that is somewhat, but incompletely, ‘disembedded’ from particular constellations of personal, social and material detail”. This limited element of transferability ascribed to autonomy has, however, been amplified by others to make it one of the key features, for example Cotterall (1995) sees it as what she terms “readiness for autonomy” (1995: 196). Little sees autonomy as necessarily involving transferability:

Human beings are autonomous in relation to a particular task when they are able to perform that task (i) without assistance, (ii) *beyond the immediate context in which they acquired the knowledge and skills on which successful task performance depends*, and (iii) *flexibly, taking account of the special requirements of particular circumstances* [emphasis added] (Little 1997: 94)

Boud also sees autonomy as an ability which is transferable, as:

It implies a responsiveness to one’s environment and the ability to make creative and unique responses to situations as they arise rather than patterned and stereotypical responses from one’s past (Boud 1988: 23)

He points out that:

It is not likely that students who are dependent on their teachers are going to be as effective in the world of learning or subsequent employment as those who have developed strategies which enable them to find and use their own resources for learning. (Boud 1988: 21)

That is, Boud is referring to the likelihood of the transfer of autonomy to subsequent study or employment. The Bergen Definition asserts that an autonomous learner “knows how to learn and can use this knowledge in any learning situation she/he may encounter at any stage in her/his life” (Dam, Eriksson, Little, Miliander & Trebbi 1990: 102). This again puts transferability at the centre of autonomy.

Authors often indirectly imply that autonomy is transferable when describing aspects of autonomy. Dam has written that “Active involvement facilitates awareness of the different elements involved in, and when, learning – an awareness to be made use of in *other learning contexts* (lifelong learning)” (2009: 134) and “The learners, for their part, are expected to engage actively in their own learning in order to become fully aware of the different elements involved in, and when, learning – an awareness to be made use of in *other contexts*” (2000: 49) (emphases added). Little (2009: 151) gives two reasons for wanting autonomous learners: they are efficient and effective because motivated and reflective, and they have knowledge and skills which they gained in the classroom but which can be applied beyond it. Sinclair (2009: 185) writes of metacognition (a key component of autonomy; see Section 2.4.5 below) that it is necessary for transferring learning know how and transcending the classroom.

This concept of the transferability of autonomy is interesting partly because it brings support back to the idea of measuring an individual’s autonomy (as opposed to a situation), but also because it raises the question of the nature of generalisation in

autonomy i.e. can autonomy be generalised (transferred), and if so, how and within what kind of limits. The literature of language learner autonomy (for example Benson 2001) has raised the problem of generalising or transferring autonomy by stipulating that it is situation specific, but the boundaries of a situation are not specified; some situations are similar to others, but at what point does the autonomy stop being transferable? If autonomy is very specific to certain situations and tasks, is it still to be seen as autonomy rather than a limited ability in one very specific and confined area? Lamb (2009: 84) reports on research he carried out with young teenagers in a UK high school. He found they had clear levels of ability in speaking about their learning. Those with “more sophisticated language and a broader metacognitive knowledge” were better able to describe and discuss their learning and had “a better chance of feeling more in control of what they are doing”. The capacities of these learners appear to make them more autonomous language learners. Language learning may be seen as one broad context or domain (Littlewood 1996) which would suggest a considerable degree of transferability is possible among tasks within this domain. If this is the case then the localised, situated nature of autonomy may in fact be quite broad and an instrument aimed at the measurement of autonomy within the domain of language learning may therefore not be as limited by the variability of autonomy as it at first appears.

An aspect of autonomy’s variability is the concept of degrees of autonomy. Degrees or levels are frequently mentioned in the literature, and this suggests that, whether they are actually accessible to measurement or not, autonomy is something which can range from low to high. Most authors either state or imply that autonomy is a matter of degree, and that it has levels. Nunan (1997: 193) for example says that autonomy is not an absolute but has degrees. Sinclair (2001: 8) also sees “degrees of

autonomy” and describes the idea of a continuum from no autonomy at one end, to ideal maximum autonomy at the other. In this characterisation learners will be at various points along this scale. Holec also sees autonomy as a scale from the lower levels of dependence to the higher levels of autonomy. He (1981: 22) holds that autonomy has to be acquired, and this is achieved through two parallel processes:

A. a gradual deconditioning where the learner sheds misconceptions about language learning (e.g.: there is only one ideal method and teachers possesses it; knowledge of the L1 is of no use for learning the L2 and experience gained from other subjects cannot be transferred; learners cannot assess their own learning).

B. a gradual acquiring of the knowledge and know-how needed to assume responsibility for learning.

Through these two processes the learner can gradually proceed from dependence to independence “from a non-autonomous state to an autonomous one” (Holec 1981: 22). If the deconditioning process (A) can be observed or gauged in some other way, perhaps by means of a self-report questionnaire, then this may offer one strand in a multidimensional package aimed at the quantification of autonomy. Holec’s second process (B) also indicates a possible avenue to explore for a method of measurement, i.e. quantifying the knowledge and know-how which the learner has.

The idea, then, that autonomy has degrees is well represented in the literature. A number of authors have in fact gone further and attempted to describe the levels of autonomy. Since levels of autonomy may offer clues to a route to measuring autonomy I will look in more detail at the ideas of a number of authors who have proposed models of autonomy which include descriptions of its different levels.

2.4.2.1 Breen & Mann (1997)

Breen & Mann (1997: 143) look at the learner and the group and see three broad phases of the development of autonomy described below (see Table 2.1).

	The Learner	Classroom Group (Including the teacher)
Phase 1	Dependent or Counter-dependent	Autocratic
Phase 2	Independent or Individualistic	Anarchic, uncertain and fragmented
Phase 3	Interdependent	Collaborative learning community

Table 2.1: Three-stage model of autonomy (Breen & Mann 1997: 143)

In Phase 1 the individuals will probably have been socialised to be dependent and the class will be teacher-led. Both teacher and learners expect and accept this. In Phase 2 the teacher encourages autonomy and so there is a shift in the classroom towards autonomy. An uncertainty of roles can result in anarchy and individualistic, non-cooperative, or competitive behaviour. The class may revert to Phase 1 or move on to Phase 3. This apparently worse situation may be a necessary step towards fuller autonomy. It may be a feature which a measure of autonomy, or users of the measure, will need to accommodate, i.e. that the autonomy level given by a measure may appear to dip – but may still be a sign of progress – before it again rises. In Phase 3 an interdependent relation emerges with the group, genuine collaboration occurs and roles are agreed.

The simple picture of autonomy levels derived from this is: No autonomy; Exposure and reaction to autonomy, and; Acceptance of autonomy. A (highly simplistic) scoring scheme could be applied, e.g. 0 for Phase 1, 1 for Phase 2, and 2 for Phase 3. This could be a starting point for a more detailed description of levels within the phases. However, the adjustment level (Phase 2) is a warning that the development of autonomy may not be a smooth rise.

2.4.2.2 Macaro (1997; 2008)

Macaro's (1997: 170–172) model divides autonomy into three domains within each of which there are progressions from lower to higher (see Table 2.2). One of the domains, Autonomy of choice and action, overtly refers to autonomy beyond the area of language learning (cf. Littlewood 1996).

The description of the abilities of learners is more specific than in Breen & Mann's model and provides hints of areas that may be measurable, such as learner strategies and transference of learning skills. Macaro's inclusion of language competence as an area of autonomy suggests that familiar measures of language ability may have a part to play in the overall measurement of autonomy, if autonomy is viewed as necessarily socially situated. However, learning gains are notoriously difficult to associate causally with levels of autonomous learning (Morrison 2005). If level of target language was safely associated with autonomy then there would be no call for a separate measure.

Area of Autonomy	Development in the learner
Autonomy of language competence	<p>Ability to communicate having acquired a reasonable mastery of the L2 rule system.</p> <p>Able to operate by and large without the help of a more competent speaker of the target language (in most classroom cases, the teacher)</p> <p>Progression from formulaic output to freer, individualised and extended output</p>
Autonomy of language learning competence	<p>Reproduction and transference of learning skills to other situations</p> <p>Learner strategies</p>
Autonomy of choice and action	<p>Opportunity to develop autonomy of choice in order to develop skills</p>

Table 2.2: Three-stage model of autonomy (Macaro 1997: 170–172)

2.4.2.3 Nunan (1997)

Level	Learner Action	Content	Process
1	Awareness	Learners are made aware of the pedagogical goals and content of the materials they are using	Learners identify strategy implications of pedagogical tasks and identify their own preferred learning styles/strategies
2	Involvement	Learners are involved in selecting their own goals from a range of alternatives on offer	Learners make choices among a range of options
3	Intervention	Learners are involved in modifying and adapting the goals and content of the learning programme	Learners modify/adapt tasks
4	Creation	Learners create their own goals and objectives	Learners create their own tasks
5	Transcendence	Learners go beyond the classroom and make links between the content of classroom learning and the world beyond	Learners become teachers and researchers

Table 2.3: Levels of implementation of autonomy (Nunan 1997: 195)

Nunan's (1997: 195) model is of five levels of "learner action" which increase in degree of autonomy from level 1 to level 5 (see Table 2.3 above). The "actions" of the learners and the degree of control they use, whether shown in overt behaviours or carried out internally as decisions, appear to be of a nature which could potentially be probed by an instrument, perhaps asking learners about aspects of their learning, for example: "Do you create your own materials?" (Level 4); "Do you know your learning style?" (Level 1); "Have you formulated your own goals?" (Level 4).

2.4.2.4 Littlewood (1996)

Littlewood's levels of autonomy are defined by the choices which are available to the learner, from low-level choices to high-level ones. He sees any number of levels as possible depending on how detailed a description one wishes; he gives an example with seven levels (see Table 2.4).

Level	Degrees of choices
1	Learners are able to make their own choices in grammar and vocabulary (e.g. in controlled role-plays and simple tasks involving information exchange).
2	Learners choose the meanings they want to express and the communication strategies they will use in order to achieve their communication goals
3	Learners are able to make more far-reaching decisions about goals, meanings and strategies (e.g. in creative role-playing, problem-solving and discussion)
4	Learners begin to choose and shape their own learning contexts, e.g. in self-directed learning and project work
5	Learners become able to make decisions in domains which have traditionally belonged to the teacher, e.g. about materials and learning tasks
6	Learners participate in determining the nature and progression of their own syllabus
7	Learners are able to use language (for communication and learning) independently in situations of their choice outside the classroom.

Table 2.4: Levels of autonomy (Littlewood 1996: 429-430)

Littlewood believes it is possible to speak of autonomy not only in a global sense but also in specific domains, such as professional or task specific. For language learning Littlewood (1996: 429-431) sees domains of:

- Autonomy as a communicator, which is the central domain of foreign language teaching e.g. choosing and using communication strategies
- Autonomy as a learner e.g. choosing and using appropriate learning strategies
- Autonomy as an individual, which is relevant as the two domains of autonomy as a communicator and autonomy as a learner also contribute to the individual's ability to make choices in life more generally

As with Nunan's model, Littlewood's appears to provide evidence of the potential to operationalise autonomy levels, for example, by formulating questionnaire items which address the points in the levels.

2.4.3 Autonomy is a capacity

A key point in Holec's view of autonomy is that of "capacity". Autonomy is a potential or ability for self-directed learning which the learner has. Thus, autonomy is "the ability to take charge of one's own learning", and the skills that this involves including determining objectives, selecting methods, and evaluating what has been acquired (Holec 1981: 3). Littlewood (1996: 428) agrees that autonomy is a capacity, but sees two distinct elements in it, ability and willingness. Dickinson has also agreed with the idea of autonomy as a capacity rather than being defined by action, and sees this as necessary if the concept is to be applied in teacher-led situations and also to situations such as self-access centres (Dickinson 1995: 167). Measurement of autonomy would need to probe beyond the behaviour of the learner and measure their ability and willingness (Littlewood 1996), and their abilities to control their learning (Holec 1981). This can be seen as a dimension from the internal or capacity to the external or demonstrated behaviour (see Section 2.4.4). Another dimension sees capacity as the more psychological and individual property contrasting with more social views of autonomy (see Section 2.4.8).

The view of autonomy as a capacity which is not necessarily demonstrated contrasts with the view that autonomy is present only when there is action by the learner which is the next key area to discuss.

2.4.4 Autonomy is demonstrated

Benson (2010: 83) feels that autonomy is not considered to be observable. However, many authors have included forms of behaviour in their definitions, for example Cotterall (1995: 195) defines autonomy (with my emphases added) as "the extent to which learners *demonstrate* the ability to use a set of tactics for taking control of

their learning” and goes on to say “Learners will *display* these tactics to varying degrees” and speaks of “learners’ readiness for the changes in beliefs and *behaviour* which autonomy implies”. The idea that autonomy has to be manifested for it to exist is, in fact, very much in evidence in definitions of autonomy, often bringing in the idea of observation of the deployment of learning skills and strategies (see for example Wenden 1991: 15; Ellis 1994: 516).

However, Sheerin (1997: 57) points out that a learner may be disposed to act autonomously, but not have the skills to do so, and there is a sense in which the disposition can be seen as constituting autonomy, as can be seen in Holec’s (1981: 3) description of autonomy as “a potential capacity to act in a given situation – in our case – learning, and not the actual behaviour of an individual in that situation”.

There are possible problems with the observation of autonomy. It is not “a single, easily described behaviour” Little (1991: 3-4) and it can be manifested in many different ways (Esch 1996: 37). Sinclair (1999: 95-96) recognises that autonomy is not the same as behaviour, and makes the point that behaviour can be observed but not its rationale: “the tutor cannot see this process, only the outcome” (1999: 101). Further to this, in cases where there is no observed autonomous behaviour learners may in fact have very good autonomous reasons for not manifesting it. Sinclair concludes that it is not useful to assess learner autonomy on the basis of observation.

Benson (2010: 79) also sees serious problems attached to using behaviour to assess autonomy. Firstly, there is the problem of determining what the key observable signs of autonomy are. There is also the probability that autonomy has non-observable components which may be important, possibly too important to ignore, and he suggests that it would be problematic to determine whether they are, in fact, vital

parts or not. Little (1991: 4) admits that in fact “we recognize autonomous learners by their behaviour”, but he is not advocating it. He points out that observable behaviour by autonomous learners:

... can take numerous different forms, depending on their age, how far they have progressed with their learning, what they perceive their immediate learning needs to be, and so on.

Autonomy, in other words, can manifest itself in very different ways. (Little; 1991: 4)

If autonomy is understood as a capacity that may not be manifested at all by a learner even though he or she does possess it, the learner may be autonomous but not behave in an observable way that could reliably be used to measure this (Sinclair 1999: 101; Benson 2001: 52; Confessore & Park 2004). It is important not to assume that autonomy will be demonstrated, or how it will be demonstrated, even when the potential is there. It is therefore important not to rely on apparent demonstrations of autonomy or its lack to measure autonomy. Hence a measure of autonomy should not be based on learner behaviour (see Section 2.5 below for a discussion of the methods of some published attempts to measure autonomy).

2.4.5 Autonomy requires metacognition

Breen & Mann (1997: 135) say that “The autonomous person is able to step back from what they are doing and reflect upon this in order to make decisions about what they next need to do and experience.” Flavell (1979: 908) sees metacognition as necessarily conscious, and Sinclair (2000: 9-10) interprets it as “conscious awareness of the learning process” which Chan (2001: 508-509) sees as essential, as “without such meta-cognitive awareness, the learner will find it difficult to exploit the learning resources at his/her disposal”. Lai (2001: 40) sees the alternative to metacognition as “robot learners who mechanically carry out all designated activities” without much

awareness of their overall learning process. Gan, Humphreys, & Hamp-Lyons (2004: 231) see metacognition as including “control or management of cognitive processes through planning, monitoring, and evaluating activities or strategies, or both”. Thus Conscious control is seen as an essential element of metacognition.

In the literature metacognition is very often interpreted as a set of skills or knowledge (for example Oxford 1990; Wenden 1991). These skills fall into three main areas: person, task, and strategy (Flavell 1979: 907). Wenden (1998) describes learners’ person knowledge as relating to knowledge and beliefs about for example aptitude or motivation and their ability as learners, both in general and for particular tasks. Task knowledge relates to knowledge of the purpose of a task and how it will serve their language learning needs, the type or purpose of the task, and its demands. Strategic knowledge is awareness of strategies in general and when and how to use them (Wenden 1998: 518-519). Sinclair (1999: 102) characterises these as awareness of the learner him/herself as a learner, awareness of the subject matter (i.e. the English language), and awareness of the processes of learning.

Cotterall (2009: 87-88) maintains that it is only possible for learners to begin to develop autonomy once they have metacognitive abilities, specifically:

- a. awareness of their strengths and weaknesses in relation to the tasks;
- b. an understanding of the tasks they are engaged in; and
- c. knowledge of strategies which can help them undertake such tasks.

Bailey & Onwuegbuzie (2002) found that the learners with the poorest performance in language learning usually had a lack of metacognitive skills shown by: poor note-taking, not seeking help when needed, not reviewing notes, not being able to manage

their moods, losing concentration, and not checking words they do not understand. Lamb (2009: 84) in his study of high school learners reports that learners with a greater metacognitive knowledge had a better chance of feeling more in control of what they were doing. The theme of control in relation to metacognition is frequent in the literature and can be seen as one of the reasons why metacognition is stressed as being conscious. Benson believes the mastery of learning skills is necessary but not sufficient for autonomy; he stresses control as being fundamental: learners have to be free to choose to learn what they want to learn or their learning may not be “authentically self-directed” (Benson 2001: 99). Little, Ridley, & Ushioda (2002: 15) agree that learners need control so that they can choose their own goals and accept responsibility. Giving learners a significant measure of control is empowering and “In motivational terms the importance of this step can hardly be overestimated” (Little et al. 2002: 15).

It can be seen therefore that there is a prevailing feeling in the literature that metacognition is essential for autonomy and is necessary for any meaningful taking of responsibility and thus for controlling learning, though the support it gives the learner to control learning is not in itself sufficient for truly self-motivated autonomy.

2.4.6 Autonomy involves responsibility

Another key aspect of autonomy found in the literature is that of responsibility. Scharle & Szabó (2000: 4) state that “in order to foster learner autonomy, we clearly need to develop a sense of responsibility”. However, it is a confused area because there are different senses of the words autonomy (see Section 2.4.1 above) and responsibility. According to Holec (1981: 3), to learn autonomously the learner needs “to have, and to hold, the responsibility for all the decisions concerning all aspects of

this learning”. Little (1996: 203-204), Boud (1988: 23), and Dickinson (1987: 15) all highlight the importance of learners taking responsibility for their own learning.

The sense of responsibility is a property of the learner not the situation. Hence, it is more closely linked to a conception of autonomy as residing in the individual learner, for example Holec’s (1981) definition. However, Dickinson (1987: 11) defines autonomy as “the *situation* in which the learner is totally responsible for all of the decisions concerned with his learning and the implementation of those decisions” (emphasis added). Here autonomy is seen as residing in the situation, though responsibility is also seen as central. Holec and Dickinson may be using different senses of responsible; Dickinson for where the situation requires it of learners, and Holec for where learners (are motivated to) seek it out.

Responsibility also implies seeing oneself as having some significant element of control or influence over what one is responsible for, a sense of agency. There are implications and connections with the area of motivation since recognising one’s ability to be an agent and so take responsibility can be seen as leading to motivation (Ushioda 2003). Alternatively, motivation can be seen as leading to responsibility (Spratt, Humphreys & Chan 2002). Even if the direction of causation is not agreed, clearly the two are closely linked. (See Section 2.4.7 which follows for discussion of the closely related area of motivation and agency.)

Learner responsibility as an aim for education is not entirely uncontroversial. Benson (2009: 25) is concerned that stressing the importance of responsibility is linked to a view of education as the encouragement of desirable behaviours expected by the institution or society. This is a concern, but it seems to me that it is part of a much broader issue of how the aims of education are implicitly linked to, and in tension

with, society's values (Biesta 2009), and is therefore not a problem specifically with the concept of responsibility.

2.4.7 Autonomy involves motivation

Explicit links between autonomy and motivation are frequently mentioned in both the literatures of language learning (e.g. Dickinson 1995; Ushioda 1996; Benson 2001) and psychology of education (e.g. Deci, Vallerand, Pelletier, & Ryan 1991; Ryan & Connell 1989). Williams & Burden (1997: 120) give their definition of motivation as:

- a state of cognitive and emotional arousal,
- which leads to a conscious decision to act, and
- which gives rise to a period of sustained intellectual and/or physical effort
- in order to attain a previously set goal (or goals).

A strong link between autonomy and motivation is found in the notion of control, especially when the learner's conscious perception is that he or she is making the decision to act based on their own intrinsic desires rather than for externally-controlled reasons. Ryan & Deci (2000: 54) illustrate the difference using the example of a student who can either be "highly motivated to do homework out of curiosity and interest or, alternatively, because he or she wants to procure the approval of a teacher or parent". Intrinsic motivation is when an action is done because it is "inherently interesting or enjoyable" (Ryan & Deci. 2000: 55) and extrinsic motivation is when an action is taken because of a "separable outcome" or consequence (ibid) where the consequence is not the inherent satisfaction of doing

something for its own “reward”. A key feature of the idea of control is that it be intentional (a conscious decision).

Deci et al. (1991: 327) maintain that there are three basic human psychological needs: competence (i.e. knowing how), relatedness (connecting well with other people) and autonomy (self-determination, being self-initiating and self-regulating). Motivation will be related to satisfaction of needs in one or more of these three areas. However, for intrinsic motivation to be maintained or enhanced there must be a sense of the act being self-determined or autonomous (Ryan & Deci. 2000: 58). There is a cline of perceived control which can be expressed in terms of locus of causality.

The concept of locus of causality is a refinement of the theory of locus of control. Williams & Burden (1997: 101) describe this as concerning one’s perception of personal control over events. According to this theory people can be placed on a continuum between those who see the control of events as internal (“internalisers”) and those who see it as external (“externalisers”). Table 2.5 shows the characteristics associated with the two extremes of the scale.

Williams & Burden (1997: 102-103) cite studies which have reviewed ways of changing a learner’s locus of control, and they suggest that it can be done, especially by teaching learners to assume control of their own learning, e.g. by practising and carrying out self-managed tasks, planning, finding and organising information, setting goals and so on.

Weiner (1979) built on the locus of control theory to allow for the fact that a learner can vary in how he or she makes attributions regarding their successes or failures.

There are three dimensions to attributions.

Internalisers	Externalisers
Feel responsible for everything that happens in their lives	Everything that happens in their lives is due to fate, luck or other people
More academically successful	Less academically successful
Active	Passive
Exploratory	Non-exploratory
Assertive	Compliant
Seek information	Inattentive
Excited about learning	
Persistent	
Problem solve	
Delay to increase rewards	

Table 2.5: Locus of Control (based on Williams and Burden 1997: 101)

The first dimension is “Locus of Causality”. People tend to attribute success or failure to ability or effort (originating inside the individual), and luck and perceived task difficulty (originating outside the individual). The locus of causality can be perceived as internal or external.

The second dimension is “Stability”. An achievement can be seen as due to an area which is fixed or permanent (“stable”), for example IQ might be seen as stable; conversely a success may be attributed to a cause which is seen as varying or subject to change (“unstable”), such as luck.

The third dimension is “Controllability”. The individual may see success as within their control or beyond their control. Mood or illness might be seen as uncontrollable. Relating the theories of locus of control and attributions back to Deci & Ryan’s theories, Deci et al. (1991: 327) say:

When a behavior is self-determined, the person perceives that the locus of causality is internal to his or her self, whereas when it is controlled, the perceived locus of causality is external to the self.

Deci & Ryan's theory (Self-Determination Theory) entails that the optimum motivation is intrinsic, and the best conditions for intrinsic motivation are when the individual feels competent, related, and autonomous. Likewise "Autonomous language learners are by definition motivated learners" (Ushioda 1996: 2) because they have the intention and the competence to take control of their learning.

Fazey & Fazey (2001: 345-346) give a description of the key features of autonomy, and these have much in common with the picture of motivation given in this section:

Autonomous people are intrinsically-motivated, perceive themselves to be in control of their decision-making, take responsibility for the outcomes of their actions and have confidence in themselves.

There is a problem with measuring autonomy which is related to a view that autonomy must be self-initiated. According to Deci et al.'s (1991) view of autonomy and motivation, autonomy requires intrinsic motivation, and intrinsic motivation requires that learners have both the will to make their own choices and the freedom to exercise that will. As Lamb (2009: 71) points out "intrinsic motivation can be stifled if a person is not allowed to be actively self-determining". There are clear implications for measuring autonomy as it may restrict the learners' freedom. Benson (2001: 52), also arguing that autonomy must be self-initiated, says:

the essence of genuinely autonomous behaviour is that it is self-initiated rather than generated in response to a task in which the observed behaviours are either explicitly or implicitly required

He is referring to situations where a researcher or teacher requires a learner to "perform" some task so that he/she can be observed and assessed for the autonomy displayed. Importantly, however, he does not look at the possibility that a learner

may actively seek such a “test” in order to self-measure (with a view to self-assessment). This clearly would be self-initiated and, if such an instrument were available, a self-administered instrument would not be open to his objection. His point, though, is a caveat for the present research, and is applicable to some previous research where a role has been imposed on learners (as will become apparent when previous attempts to measure autonomy are examined in Section 2.5 below).

If autonomy is self-initiated and stems from an internal intrinsic motivation then imposing a test on a learner will inhibit the very autonomy on which observation is being attempted. However, Lamb (2009: 71), drawing on empirical evidence from learners he interviewed, found that giving learners a real choice can overcome the problem of teachers inhibiting learners’ freedom. I argue, therefore, that a measure of autonomy which a learner has freely chosen to undertake is not subject to the criticism that it is anti-autonomous.

2.4.8 Autonomy involves social interaction

In this section I consider the view that autonomous learning is essentially social and interdependent and contrast this with the view that it is primarily concerned with independent learning. There is a tension between individual and social views of autonomy. It can be seen as a quality of the individual which is affected by his or her psychology (e.g. Little 1991) and skills (e.g. Holec 1981), or conversely it is argued that second language learning is a process situated in a social context (e.g. Pavlenko & Lantolf 2000).

My original perception as a practitioner in the 1990s was that autonomy focussed on the individual learner; it appeared to have developed from a constructivist view of

how individuals learn, which Gremmo & Riley see as a reaction against behaviourism, which they call “the sterile hubris of a mechanistic psychology which dared to extrapolate from dumb animals to human” (1995: 152). This reaction converged from at least two directions to the notion of learner-centredness and autonomy involving humanistic and cognitive psychologies.

Humanistic psychology (Maslow 1968; Rogers 1969) sees self-actualisation, or the growth of the individual as a complete person, as a human need which is the source of motivation. It led to the growth of the humanistic curriculum (Dubin & Olshtain 1986: 75). This view of education reacted against traditional ideas where learners were encouraged to:

develop rote abilities and depend upon being able to give back what is expected rather than to make it into something that relates to the rest of their cognitive life (Bruner 1974: 406)

This type of learning, where the learner was characterised as a recipient of knowledge, did not engage what humanistic psychologists saw as the innate human desire to learn. Humanistic theories claim that learning should involve the learners more, making them active participants and having them take on personal responsibility for the process of learning. Rogers (1969: 162) says:

Learning is facilitated when the student participates responsibly in the learning process. When he chooses his own directions, helps to discover his own learning resources, formulates his own problems, decides his own course of action, lives with the consequences of each of these choices, then significant learning is maximised.

The other major psychological strand leading to the notion of individual autonomy in learning was work in the field of cognitive psychology. Piaget’s constructivist view of learning emphasised the importance of cognitive processes in the individual

(Williams and Burden, 1997: 21-24). In this view an individual learns by constructing knowledge for him/herself. Learning is seen as the incorporation of new information into the learner's mental framework and may necessitate the learner actively reorganising the way the framework is configured, which involves an active participation in the process of learning. As Page (1992: 83) puts it "Every learner in every situation is, strictly speaking, autonomous because only the learner can learn, no-one can do her learning for her".

Both humanists and cognitivists emphasised learning as a process resulting in an extension of the individual's capabilities. Learning was something learners did rather than something which was done to them and this led to more learner-centred approaches to language teaching. The communicative approach, for example, grew from ideas of learner centredness and the view of language as a tool for communication in social groups, combined with the constructivist cognitive psychology reacting against behaviourism. According to Gremmo & Riley (1995: 153), autonomy is a "logical entailment" of the communicative approach and Littlewood (1996: 427) saw autonomy as a concept that fitted well with learner-centred teaching methods.

This was how I saw autonomy in the 1990s, concerned with bringing out and developing the self-reliance of the individual learner. However, since then there has been a growing belief that it is better to treat autonomy as socially situated. Esch (1996: 37) says that autonomy "is not self-instruction or learning without a teacher" and similarly Little (1991: 3) says it is neither "synonymous with self-instruction" nor "essentially a matter of deciding to learn without a teacher". Smith & Ushioda (2009: 244) note that:

autonomy is now seen to develop out of interaction with others; it benefits from interdependence, and classrooms and teachers are no longer peripheral but at the centre-stage of practical concern

With the change of focus from autonomy as independence to autonomy as social there has been a growth in interest in autonomy in the classroom as opposed to the previous emphasis on individuals in self-access centres. Smith & Ushioda (2009: 248) maintain that the social view of autonomy means that it:

is not seen as an abstract set of discrete skills, attitudes or behaviours to be developed, but a historically and socially situated process that evolves and is mediated and instantiated through relations among persons-in-action in specific contexts of practice

A major psychological strand leading to the notion of social autonomy is work in the field of cognitive psychology by Bruner et al. (1966) using the idea of social scaffolding, which elaborated on Vygotsky's work. Vygotsky (1994: 116) pioneered ideas of learning as a social process:

The entire history of the child's psychological development shows us that, from the very first days of development, its adaptation to the environment is achieved by social means.

From childhood more competent others help the learner to move to the next level in what Vygotsky (1978: 86) calls the zone of proximal development (ZPD):

It is the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers.

Although this theory of scaffolding relates specifically to children's development, the belief is that the pattern continues and is even reflected in adults' internal "dialogue"

(Ohta 2000: 53-54). This model emphasises that autonomous learning should not imply solitary learning.

Authors sometimes polarise the social and independence views of autonomous learning, arguing for its social nature by attacking the straw man of radically isolated learning. Little (1990: 27) wrote that “total detachment is a principal determining feature not of autonomy but of autism”; this may be true of “total detachment” but this does not have the same meaning as the terms “individual” or “independent”. The degree of detachment suggested by Little is not found in the literature, though Little has been quoted by other authors (e.g. Barnett 1993; Hurd 1998) as if it were. This type of polarising may lead to an idea that autonomy cannot be found in independent settings.

One of the often-quoted definitions of autonomy is the “Bergen definition” (Dam et al. 1990: 102) which specifies that “An autonomous learner is an active participant in the social processes of classroom learning”. This appears to be saying that, by definition, an autonomous learner cannot be one who learns outside the classroom. Empirical, logical, or conceptual grounds for this assertion are not given; it appears, rather, to be a statement of how the authors will use the term autonomy based on their own conceptions (and/or those of others). Seen in the context of Gabrielsen (1990: 96-101), who contributes the section immediately preceding Dam et al.’s in the same 1990 Nordic Workshop report (in fact it forms part of the same article as Dam et al.), the Bergen Definition can be read as a reaction to a prevailing focus on the individual learner, which Gabrielsen (1990) outlines, and to a role of the teacher which is seen as too prescriptive to allow learners to be involved in the control of their own learning. If the Bergen Definition is read in this light it appears not so

much to be a “definition” of autonomy but rather a statement of how formal education should be carried out. In the two-part article the first part by Gabrielsen briefly explains the view of the situation and outlines the policies to improve education. In the second part of the article Dam et al. provide a kind of executive summary of the implications of the preceding section in the form of a description of how education should be. This has, misleadingly, become labelled the Bergen *Definition*. In fact it is not a definition, but should be read as a call for educational reform, with some specific aims in mind to counter a prevailing situation which is seen as having negative points (e.g. concentration on individual learners, and the overly-controlling nature of teaching). Part of the “definition” (Dam et al. 1990: 102) reads “[Learner autonomy] entails a capacity and willingness to act independently and in cooperation with others” (underlining in original). The emphasis on cooperation with others is not underlining that autonomy is inherently social as a concept but is emphasising the point that education must in future not *only* focus on individuals and their needs but also on the group. It appears, then, that autonomy is being allowed both individual and cooperative, social, aspects.

Cooperation with others can be seen as integral to autonomy from social and psychological perspectives. Kohonen (1992: 19) writes that:

Personal decisions are necessarily made with respect to social and moral norms, traditions and expectations. Autonomy thus includes the notion of interdependence, that is being responsible for one’s own conduct in the social context: being able to cooperate with others and solve conflicts in constructive ways

Little (1996: 210) sees collaboration as essential for autonomy as a psychological capacity:

the development of a capacity for reflection and analysis, central to the development of learner autonomy, depends on the development of an internalization of a capacity to participate fully and critically in social interactions

There appear to be two closely linked but separable concepts regarding the social nature of learning: (i) that learning is *inherently* and unavoidably social because that is the origin (in children) of how humans learn (e.g. Vygotsky 1994; Lantolf 2000), and (ii) that interaction is *beneficial* to learning (e.g. Littlewood 1981: 93-94; Legutke & Thomas 1991: 150-151).

If autonomy is necessarily socially situated (e.g. Murphey 2003; Little 1990) it would mean that measuring an individual's autonomy would be to attempt to measure something which does not exist. Carr & Claxton (2002: 12) see assessment in the social environment as being "concerned with the process of participation". If so, this would mean that a quantitative questionnaire may not be appropriate for measuring autonomy. The highly complex interrelation of influences in a social situation would not appear to be amenable to measurement by a closed-item quantitative autonomy-measuring instrument. The opposing poles of individual and social, however, do not exclude a synthesis and according to Sinclair (2000: 11) "Autonomy has a social as well as an individual dimension". My investigation was aiming to be open-minded and it would be interesting to see how the outcomes of my research related to social views of autonomy.

2.4.9 Autonomy is political

In this section I will focus on the more political end of the autonomy spectrum, as distinct from the more technical and psychological end. These ends can (crudely) be represented by qualitative and quantitative research paradigms respectively. Authors

from the more political end of the spectrum sometimes write of their fears that the quantitative paradigm is influencing, misrepresenting, or taking over the understanding of autonomy. Many authors see autonomy as starting out politically, then becoming steadily more “technologized”. Pennycook (1997) and Benson (1997) felt that the political was inherent in autonomy but was being ignored due to the prevalence of the positivist paradigm. Pennycook wrote (1997: 41):

The idea of autonomy has therefore moved rapidly from a more marginal and politically engaged concept to one in which questions are less and less commonly asked about the larger social or educational aims of autonomy. Broader political concerns about autonomy are increasingly replaced by concerns about how to develop strategies for learner autonomy. The political has become the psychological.

One suggested reason for this is given by (Crabbe 1993 444), “the psychological is the most appealing to educationists in that it is pedagogical rather than political”. Another is given by Benson (2001 46), “In the context of language education, the more convincing arguments for autonomy are likely to be pedagogical rather than political or philosophical” (though he is not himself convinced). It is also suggested that the technologized pole is being favoured at present due to autonomy being taken up more popularly around the world, but with differing understandings (Esch 2009).

Gremmo & Riley (1995: 152-154) examine the reasons underlying the development of autonomy in language learning, and they see much of its impetus as being inspired by the desire to make changes in society. Holec (1981:1) sees in Western countries in the late 1960s a rise in social awareness regarding improving quality of life (e.g. civil rights movements). The expectation of greater freedom and equality led to a focus on bringing ideas of political autonomy into education. Much of the initial impetus for the rise of learner autonomy in education can therefore be seen as political.

Partly due to these origins, many authors feel that autonomy necessitates learners being more aware of the power relations implicit in learning so that they can take more control over their own learning (e.g. Pennycook: 1997; Benson: 1997). Kenny (1993: 440) states that:

... it can be said that only when autonomy is being allowed to function is education taking place at all. For where autonomy is repressed or ignored--in other words where the learner has no say and no being—then what we have is not education but some sort of conditioning procedure; the imposition and reinforcement of dominant opinion

According to Little (1991) autonomy has an implication of social awareness, and there is not a strong distinction between learning and life in general. The attitudes autonomy promotes should result in “more useful members of society and more effective participants in the democratic process” (Little 1991: 8). For Benson (2001: 46) education is a matter of concern to the whole community, “authoring the social realities that constitute our collective lives”. In this critical pedagogy learner autonomy is seen as “socially situated agency”; Toohey (2007: 241-242) writes:

if we are interested in education for democracy, we must ask critical questions of if and/or how specific practices, resources and identity roles for teachers and students mirror other (actual or desired) social arrangements in larger social worlds beyond the classroom

Ideas of democracy are frequently used to distinguish the aim of this view of autonomy:

Many advocates of autonomy, despite their national and/or cultural situations, seem interested in promoting the power of individual students and teachers to determine their futures, to participate in democratic communities that recognize teachers, as well as learners, as simultaneously involved in learning and critical social practice (Toohey 2007: 242).

This democracy and autonomy can only happen, according to Moreira (2007: 70), when “learners have their voices heard, are able to participate in pedagogic decisions, and are able to decide on the course of their learning”. Autonomy and democracy are frequently linked with struggle, for example Morieira (2007: 58) says

democratic transformation in the classroom is achieved through a shared struggle to promote students’ autonomy as learners, and, in the process, increase the democratic nature of the teaching and learning process

Phillipson (1992) sees the political nature of language learning as due not to the nature of autonomy specifically but to the nature of education. He says (1992: 67) “The belief that ELT is non-political [...] assumes that educational concerns can be divorced from social, political, and economic realities” which would be impossible.

A related strand which promotes the political interpretation of language learning relates to the perception that language is fundamentally embedded in society. The language learner should not be seen, to use Atkinson’s (2002: 525-526) analogy, as a single cactus in the middle of a lonely desert but rather as a plant in a tropical rainforest with a lush ecology of complex relationships. He says (2002: 538) “language and its acquisition are not radically disconnected from the rest of the world” and language learning has “real potential for changing the world”.

The points made so far can be seen as supporting a political interpretation of autonomy in that they are pro-social and seek to show what follows from the socially embedded nature of the learner, learning, and language. There is also, however, a political strand which takes a more individualistic perspective on the learner. For example Dearden (1975: 7 quoted in Boud 1981: 22) writes of the qualities of an autonomous person as:

wondering and asking, with a sense of the right to ask, what the justification is for various things [...] refusing agreement or compliance with what others put to him when this seems critically unacceptable [...] defining what he really wants [...] forming purposes and intentions of his own independently of any pressure to do so from others [...]. In short, the autonomous man has a mind of his own and acts according to it

Boud states that the concept of autonomy in education “refers to the capacity of an individual to be an independent agent, not governed by others” (Boud 1981: 22).

There is a non-conformist and pro-individual liberty feel to this. It is possible that autonomy has been a successful idea because it appeals to two different political views, i.e. there is the struggle against oppression in order to create a more equal society, and there is the struggle for the individual’s freedom within society.

One argument for the political view of autonomy is that all human life is political and therefore try as we might, it is impossible not to be politically engaged (Benson 1997). Taking a psychological stance on autonomy would in this case itself be a political stance. The argument then is not that autonomy is political or not, it is which stance is the “right” one for autonomy, the critical (political) or the psychological (“political”). He is saying, in effect, that autonomy is a belief and believers should do what they believe is right despite the absence of one agreed definition of autonomy.

The important debate for Benson can thus be seen as regarding how autonomy *should* be seen, rather than trying to establish how it “really is”. He says “it is difficult to establish or defend any particular definition of autonomy against any other definition through logical or reasoned argument alone” (Benson 2009: 21).

Benson uses this to justify the approach of seeing autonomy in terms of what kind of world we want to live in, of how things should be, and this means seeing autonomy

in political terms, particularly as a counter to the negative effects of globalisation on education (Benson: 2009: 23).

The quantification of autonomy required for a measure suggests a scientific conception of autonomy and the associated (“political”) stance of aiming to separate data regarding the technical considerations from political-critical values. This would be with the belief that data of this nature will be useful for learners, and may assist them in their personal and/or collective struggles for autonomy.

This concludes the section on the key conceptual considerations in autonomy and its measurement. Some challenges to a measure have been discussed and progress has been made in narrowing down the theoretical stance implied by attempting a quantification of autonomy when it is very often seen as occupying a qualitative dimension.

Each of these concepts has been seen as key to autonomy, but in reality they will tend to be present together and will interact to varying degrees. For example, taking responsibility implies having the ability to control which implies metacognition; motivation provides the necessary energy to take up and wield control; social interaction may enhance intrinsic (and extrinsic) motivation; behaviour is the tangible end result, from the perspective of others, which can again feed back into motivation. It has to be assumed that a putative measure of autonomy will need to take account of all of these areas and this is reflected in the item selection process for the initial long list described in Section 3.9.

2.5 Previous autonomy measuring research

Measuring autonomy is not a large area of the literature, but there has been much literature looking at evaluation or assessment, especially learner self-evaluation, for example: Little (2005) on formal self-assessment; Karlsson, Kjisik & Nordlund (2005) on self-evaluation for needs analysis; and Wenden (1998) on self-evaluation as necessary for metacognition. There has not, though, been much exploration of the measurement of autonomy of the type under investigation in the present research. In this section I will first look at a qualitative approach which focuses on evaluation and discuss its usefulness for the present research. Following this I will look at four examples of work relevant to measurement which highlight issues for the present research. I will use lessons learned here to explain and justify my methodological decisions (see Section 3.6).

2.5.1.1 Dam (2000)

Dam aims to evaluate a particular type of autonomous learning, which she defines as:

... what takes place in situations in which the teacher is expected to provide a learning environment where the learners are given the possibility consciously to be involved in their own learning and thus become autonomous learners. The learners, for their part, are expected to engage actively in their own learning in order to become fully aware of the different elements involved in, and when, learning – an awareness to be made use of in other contexts
(Dam 2000: 49)

In an institutional context the aim is not *autonomy*, maintains Dam, but *autonomous learning*, and this aim is to be achieved by means of autonomous learning itself. Consequently evaluation must be both of the process and the outcome.

Dam divides periods of learning into three phases:

- Phase 1. Teacher-initiated and directed activities. In this phase evaluation consists of awareness raising and introducing the evaluation tools.
- Phase 2. Learner-initiated and directed activities. In this phase learners are supported in the use of the evaluation tools.
- Phase 3. Together session. In this phase both learners and teachers are in charge.

The evaluations in each phase can be applied to all aspects of the learning process.

The evaluation methods are:

- Self-evaluation. This is always done before teacher evaluation to show learners they do not need to depend on the teacher.
- Oral spontaneous. This involves peer evaluation or teacher with learner evaluation which is not planned.
- Written. This utilises diaries and/or questionnaires which can be learner or teacher-produced. The teacher will give questionnaires to help learners realise and focus on new areas of their learning process. The teacher can collect in the questionnaires, analyse the data, and present the results to the class for discussion. Learners can also be asked to produce their own questionnaire items in order to help them reflect and to provide a window into how learners perceive their learning process.
- Oral planned. This can be in groups or as a whole class, and are based on the questions/questionnaires answered by learners.
- Combinations of the above.

In all phases the same basic procedure is followed: evaluation, followed by discussion, then revision. This raises awareness and hence has a developmental function. Dam also carries out peer evaluations based on presentations of learner projects. She asks individuals to write down their opinions which can then be discussed, or collected for class discussion (Dam 2000: 53).

Dam provides her students with feedback (evaluation) on the learning process in three ways: orally, to individuals or groups; via comments in their learner diaries; and via written evaluation given to the whole class, e.g. at the end of term. To evaluate the outcome of a period of learning from the learners' point of view she gives them two open ended questions and they have to write their answers. For example (Dam 2000: 54-56):

- Which issues in our English lessons would you describe as being important – and why?
- What have you learned in your English lessons that you feel you might make use of in other learning situations?
- Why have some of the ways we have worked in the English lessons been good?
- What are – in your – opinion the main differences between the way we work now and the way you worked before I took over the class?

The idea of evaluation (self-evaluation, evaluation between peers, and between learner and teacher) which is Dam's focus is a central and familiar part of learner autonomy, being necessary for the ability to manage one's own learning. This is rather different from the *measurement* (as opposed to evaluation) of autonomy being probed in the present research (as discussed in Section 2.1.2 above). In her paper

Dam is not addressing the need for a *measure* of autonomy, but is concentrating on *evaluation*. She describes methods of teaching she uses in her autonomous learning classroom utilising reflection and feedback to help learners develop their autonomous learning abilities.

Dam's perspective is that of a teacher evaluating her own classroom and practice rather than that of a researcher investigating the possibility of a measurement instrument. Consequently, she does not focus on a number of issues which would be important for measuring autonomous learning. She does not for example address questions of validity or reliability. The feedback which Dam presents from learners and parents shows that the learners express satisfaction with the classes, but she does not present this as an indicator of the degree of autonomous learning. Dam describes classes which appear to have a tight structure controlled by the teacher with the aim of developing the learners' autonomous learning. For example she describes how periods of learning are divided into phases by the teacher, with all phases following the same basic procedure of evaluation, then discussion, then revision. An autonomy measuring instrument on the other hand would need to justify its approach in the light of the points made by Champagne et al. (2001: 49) and others (discussed in Section 2.3) regarding the tension between autonomy and the imposition of a measurement instrument. The specifications for the autonomy measure under investigation in the present research (see Introduction Chapter), for example speed, usability in contexts of independent learning and large classes alike, accessibility for the inexpert user, etc., are not addressed by Dam's approach, which requires much time, and is integrated into lessons.

There is, in summary, a need for a measure of autonomous learning which is distinct from a procedure for evaluation. These are not mutually exclusive and could be complementary (Dam even recommends the use of teacher-produced questionnaires in her description of evaluation).

In the following four subsections I will consider practical examples from the literature of four researchers' attempts to design and apply measures of autonomy which are more quantitative in nature, or have the most potential relevance for the quantitative measurement of autonomy. The four examples described here demonstrate different approaches to the problem. I will assess each relative to the aims of the present research to find a practical measure. Review of this literature is also useful for discovering pitfalls and weaknesses which I may face in the present research. It will be seen that none of them provides a solution which suits the requirements of the present research, but each provides valuable lessons which can be incorporated into the hoped for simple yet effective measure of autonomy.

2.5.1.2 Ravindran (2000)

Temasek Polytechnic, Singapore, is perhaps unique in having a credit-based course aimed at assessing autonomy which leads to a formally recognised award, known as the Certificate in Independent Language Learning (CILL). In a course lasting up to three years, learners studying for the CILL “learn and demonstrate the skills of independent learning in a systematic manner” and are trained “to take responsibility for their own learning using language learning as a vehicle” (Ravindran 2000: 64).

The course begins with an orientation module on “the foundation skills of independent learning” to prepare learners for work in the self-access centre.

Following this learners progress through modules which require them to demonstrate

their abilities. The guidance given for the modules is progressively reduced and the credits available are accordingly increased. The course culminates in a compulsory summative project.

The researchers sought to make the CILL a “possible, workable, reliable and valid” assessment method (2000: 65) in the context of the Temasek Polytechnic self-access centre. Learners were expected to display their levels of awareness and provide demonstrations of their ability to apply the range of skills which had been selected by the team (2000: 66) as necessary for autonomy. The criteria for the award of the CILL consist of twenty items derived from Knowles’s (1975) “key skills of self-directed learning”. Assessment was continuous: CILL helpers made profiles of learners’ strengths and weaknesses; observations of the learners were carried out; CILL helpers read learners’ learning logs, learning reviews and contracts, and assessed the quality of learners’ reflections on their learning and on the tasks they had carried out; the quality of learners’ language in the work they submitted was also assessed.

For the final assessment, decisions were carried out as a team, as many people had been involved in judging and observing many students. For this reason the CILL team engaged in regular training, monitoring, and feedback sessions and carried out inter- and intra-rater calibration sessions which had at the time of going to press shown minimal discrepancies (Ravindran 2000: 66). Detailed records had to be kept by the CILL helpers (including records of consultations, absences, and cancellations at short notice) to aid in reporting and assessment, and for the quality control requirements of formal certification (2000: 69).

I will discuss the CILL project and consider what lessons can be learned from it with relevance to the present research. It is very different in conception from the kind of instrument envisaged in my aims, being based on a summative assessment after three years' of work, and involving formal testing and certification. A disadvantage of CILL is that a sustained effort and commitment were required to set up and maintain it. A dedicated and trained team was needed, which means this is not something an individual student, teacher, or researcher could hope to carry out. It is also not quick; it requires three years of work in a self-access centre, and is therefore not something which could be achieved quickly, for example, for a needs analysis. It is striking the extent to which it was necessary to commit time and resources to administration, training of staff, keeping records, and establishing and maintaining inter- and intra-rater reliability. A drawback of CILL is its complexity, which means that without the drive of the individual/s who initiated it there is a danger that it will collapse. My research, however, aims to investigate the feasibility of a simple, quick, and easy-to-use measure which will also be potentially usable for self-assessment. CILL is clearly none of these, but it may present useful lessons to be incorporated into the present research.

One such lesson is drawn from the impression that the CILL often appears to be un-autonomous; choice is not handled in an autonomy-conducive way, for example “The certificate programme requires learners to demonstrate their level of awareness and the ability to apply the skill of self-directed learning” (Ravindran 2000: 66). The word “requires” implies that learners do not have the freedom to manage their own learning in their own way. It seems to be an awkward combination to have a closely controlled training course as the basis of a measure of autonomy, as autonomy is supposed to involve the freedom to choose and control one's own learning path (as

discussed in Section 2.3). This issue is clearly of relevance to the present research and will need to be addressed so that the instrument is not imposed on learners.

A further important issue raised by Ravindran's paper is that, despite the great efforts that were made to establish its reliability, CILL is not explicitly construct-validated to establish that it is measuring autonomy. This aspect of establishing a measurement instrument, which appears to be of central importance, is an issue which the present research will need to address in relation to its aim of establishing whether such an instrument is viable. The choice of elements to include in the CILL was based primarily on one author's work, Knowles (1975). Other elements are from the author's and others' own experiences, but the paper does not detail the procedure used in the choice, and this suggests that the present research will need to present this area clearly and aim to be as objective as is practicable.

2.5.1.3 Cotterall (1995)

Cotterall defines autonomy as "the extent to which learners demonstrate the ability to use a set of tactics for taking control of their learning" (Cotterall 1995: 195). By "tactics" she means to include setting goals, choosing materials and tasks, planning practice opportunities, and monitoring and evaluating progress. The learners will have these to different degrees. This is partly because learners have different beliefs about language learning prior to interventions to encourage autonomy. Cotterall's aim is to try to see if learners are ready for these changes because autonomy requires changes in beliefs and behaviour, and she believes that the learner's beliefs profoundly influence their learning behaviour. If the learner has mistaken beliefs this may hinder their learning progress. Cotterall gives the example of "making mistakes is bad". If learners believe this, they will be inhibited, they will not practise, and

therefore they will not learn the language as effectively as they could. She says “All behaviour is governed by beliefs and experience. It follows that autonomous language learning behaviour may be supported by a particular set of beliefs or behaviours” (ibid 196). Cotterall wants to identify areas in students’ beliefs and probe how these categories relate to autonomy. She uses the statistical technique of factor analysis to examine the questionnaire data to form scales of similarly-behaving items which may shed light on students’ beliefs. (Factor analysis is used in the present research and is described in detail in Section 3.12.3.) These groups, called factors, can then be named by the researcher according to what appear to be their common characteristics. Cotterall wants to show with this technique what the beliefs are which autonomous language learners hold, and that these can indicate the learner’s “readiness for autonomy”.

Cotterall carried out interviews and used the data to make the questionnaire on learner beliefs about language learning. The questionnaire had 26 items with a 5-point Likert scale, and eight items using a forced choice format. The factors Cotterall found were:

- Role of the teacher
- Role of feedback
- Learner independence
- Learner confidence in study ability
- Experience of language learning
- Approach to studying

Cotterall’s Factor 1 is “Role of the teacher”. This suggests that preferences about the degree of teacher control are a key area in describing an autonomous learner. She describes factor analysis saying how it defines connections between items which seem to behave in the same way and she sees advantages to “empirically identifying

dimensions underlying the construct of language learning autonomy” (ibid 197).

Cotterall looks at the factors revealed by the study and discusses what this may reveal and the implications. Some of them are areas that tend to support autonomy and others seem to be blocks to it (ibid 200):

Autonomous learners are likely to be individuals who have overcome the obstacles which educational background, cultural norms, and prior experience may have put in their way. The degree of independence with which learners feel comfortable will be a key indicator of their readiness for autonomy

Cotterall’s Factor 4 is “Learner confidence in study ability”. She believes it is central to diagnosing readiness for autonomy and finds in the literature (ibid 201):

general agreement that learner confidence correlates with academic success [which] supports the view that confidence is a defining characteristic of autonomous learners.

This concurs with Littlewood (1996) and also with motivation as characterised in Self-Direction Theory (see Section 2.4.7) and makes confidence one of my choices of area for the items in my questionnaire to cover.

Cotterall sees Factor 5 “Experience of language learning” as being metacognitive knowledge and says (ibid 202):

Learner beliefs about language learning will profoundly influence their approach to language learning. Learners need to be aware of the role of cognitive and affective variables in language learning

Factor 6 is “Approach to studying”. Cotterall admits that the link with autonomy is not clear (ibid 203). By “approach to studying” she means learning style, and she

warns that “It may be unhelpful to suggest that a particular approach to studying characterises autonomous learners” (ibid 203).

A weakness is the small number of items included in the questionnaire, only 34, and it seems that this was the total number used in the development, though she says only that the items “were developed from a series of interviews with ESL students about their experience of language learning” (ibid 196), without giving the criteria for selection.

Cotterall does not propose this as a tool to be used by classroom teachers for measuring autonomy. However, the questionnaire could be used to help the teacher and learner explore the learner's beliefs and therefore help the learner with the reflection and awareness-raising which are recognised as essential for learner autonomy.

Cotterall does not explore the construct of autonomy she uses in her questionnaire; rather, it is presented as a starting point. The use of many more items in the development of the questionnaire combined with the factor analysis she uses may have produced more interesting empirical results and this idea has led to the inclusion of this technique in my own research into a viable closed-item questionnaire.

Cotterall’s paper highlighted for me that the interpretation of the underlying concept of factors is a subjective process. Occasionally her interpretations seem to go beyond what can be concluded, for example in Factors 4 and 5 there are only two items and it seems unsafe to draw conclusions based on this; I have consequently been cautious in this respect in my research (see Section 3.12.3.3.5).

It might be a problem that Cotterall assumes that the learners will have beliefs about language learning. In Cotterall's questions she starts with "I believe...", or "I am confident...", or "I am willing..." rather than making them simple statements. For example one item reads "I believe I know how to find my own ways of practising" which has a rather awkward sound. The more literal minded student may wonder whether they believe they know or whether they know they know; in a way it is a double-barrelled question. The Likert scale should provide the degree to which they agree or disagree with the statement and therefore it would not be necessary to add words into the statement which make it stronger or weaker. It may be that Cotterall included words such as "I believe" in her questions because she wanted them clearly to relate to beliefs. For the present research the wording of items and their relation to the Likert scale will be given great attention in an attempt to avoid this type of drawback which may have an impact on the validity and reliability of the instrument (see Section 4.2).

Interestingly motivation does not overtly appear among the factors. Perhaps this is because including motivation would have led to awkwardly-phrased items such as "I believe I am motivated". Since motivation is clearly important to autonomy (as has been seen in Section 2.4.7) it may be present in items or factor groupings but not clearly manifested. This is an issue which I will return to at a later stage in the thesis (see Section 7.3.5).

I gained useful lessons from the analysis of Cotterall's work; factor analysis is a very promising technique, especially when it is used in a genuinely exploratory way, without imposing one's own preconceptions about what factors should emerge from the data, and this approach was adopted in the present research (see Section 3.12.3).

2.5.1.4 Lai (2001)

Lai aims to demonstrate a validated scale for assessing learner autonomy which will have universal utility in contexts where learner training is being carried out. She divides autonomy into two areas: process control (which operates at the micro level of a task), and self-direction (which operates at the macro level). Lai (2001: 35) defines “process control” as:

a learner's ability to set realistic task aims for her chosen piece of material or activity; identify problems; employ relevant strategies to tackle the problems; and conduct self-assessment of the learning experience with an aim to set future challenges

She defines “self-direction” as (ibid 39):

the learners' ability to set realistic goals for their learning, identify scope of learning [...], relevant materials to work with and related activities to engage in, and skilfully employ them for monitoring their own learning, set their own pace for learning, and conduct self-assessment

For process control, Lai decided to evaluate the extent to which learners increased their control in two areas: “setting aims” and “carrying out self-assessment”. Lai asked them to decide aims and choose tasks to address them. The aims were assessed regarding: whether they were appropriate to the task chosen; and whether they were conducive to improving the learner’s listening skills/strategies (the context is a listening course).

For the self-assessment category Lai had two criteria for assessing how well the students had done: whether the self-assessment is related to the learner-specified aims; and whether the self-assessment is related to the learner’s listening process and/or performance. A 5-point rating scale (from 0 to 4) was used.

Lai stresses the importance of metacognition as “a necessary condition for conducting self-directed learning” (ibid 40) and this is a central part of her rationale for the assessment of the learners’ self-direction. Lai had the learners design a personal course of self-directed language learning and gave them a list of all the item headings that they were supposed to cover in their plan, which included setting their own criteria for self-assessment. Lai gave their plans to raters to assess the learners’ metacognitive awareness and their planning ability. She did not look at whether they could actually carry out the plan, but she believes:

we can nevertheless infer from a conceptual representation of the course, in the form of a course design, whether the learner has grasped the rationale behind it, or has the potential and/or ability to do so (ibid 39)

Lai had three raters to assess the plans and she prepared notes for each rater and also trained them beforehand. The plans were scored by how they compared to 17 statements, using a scale from zero to six. Validation of the measurement scale was carried out using the internal consistency of the items. Lai checked the inter-rater reliability of the three independent raters, and she then calculated the reliability coefficients among the three raters using Spearman rank-order coefficients. Lai concludes that the two rating scales were both valid and reliable and therefore the scores which they calculated based on the total mean scores of various raters were meaningful and reliable.

Lai’s research follows a format which is quite intuitive: a working definition of autonomy is given which involves elements which are measurable and an instrument or procedure is designed to gather data on these dimensions. A stage which is missing, I feel, is checking the construct validity by, for example, seeing whether

other suitably qualified people would formulate the same definition independently. There is a risk in her approach that the definition may not be acceptable to others, and this will limit the universality which Lai gives as one of her main aims. As with Cotterall's approach (Section 2.5.1.3) I feel that construct validity is an issue with autonomy measuring instruments which the present research will need to confront (see Sections 3.2.3 and 6.4).

Lai feels that a course design produced by a learner is in itself sufficient to show "whether the learner has grasped the rationale behind it, or has the potential and/or ability to do so" (ibid 39). I would agree that it shows potential as the learner is showing metacognition, but it does not address the different skills, capacities, attitudes, or motivations necessary for successfully carrying out a plan. In my own research I am investigating the possibility of a closed-item questionnaire for measuring autonomy and am, as with Lai's course design, not looking at the manifestation of autonomy in action; however, I will endeavour to minimise this feature of questionnaires by probing a broad range of autonomy-related dimensions by principled selection of items to cover such areas (see Sections 3.9 and 4.2.3).

The instrument used has a very similar approach to the General National Vocational Qualification (GNVQ) in the Key Skill area of Improving own Learning and Performance in which learners make personal study plans, including self-assessment plans, which are then rated by an assessor using a scale of band descriptors. This has the disadvantage of being a lengthy process making heavy demands on the teacher's time if a whole class has to be assessed (as I know from personal experience) and it also requires the teacher to be trained before using it. This means that most teachers are not able to use it without substantial preparation and commitment. Lai's approach

is similarly not a quick way of achieving the aim of measuring learner autonomy, and in the present research one of the aims is to explore whether speed and convenience can be increased.

Finally to assess their autonomy Lai requires the students to act in a certain way as she gives them explicit instructions and guidelines on what to do and then uses the data obtained to measure autonomy. This can be criticised for limiting autonomy by not allowing it to be self-initiated (Benson 2001: 52), and this is one reason why I do not wish to attempt to test autonomy but find a tool to help teachers support their learners in developing autonomy.

2.5.1.5 Sinclair (1999)

Since autonomy is a capacity (Holec 1981: 3) Sinclair believes it is this capacity which needs to be assessed, and this cannot be done reliably by observing learners for a short space of time. The key to doing this is metacognition. Sinclair (1999: 102) says that “The link between the development of metacognitive awareness and learner autonomy is clear”. The three areas which the learner should have metacognitive awareness of are: the learner him/herself as a learner; the subject matter (i.e. in this case, the English language); and the processes of learning.

Sinclair (ibid 102) gives aspects of metacognitive awareness for each of these areas (see Table 2.6), and gives the criteria for assessing metacognitive awareness as questions about ability, such as, can students:

- Provide a rationale for their choice of learning activities and materials?
- Describe the strategies they used?
- Provide an evaluation of the strategies used?

- Identify their strengths and weaknesses?
- Describe their plans for learning?
- Describe alternative strategies that they could have used?

Area	Aspects
Self	Attitudes Beliefs and expectations Motivation Needs Learning style Preferred learning environment
The English language	Language awareness Systems Varieties Similarities and differences between mother tongue and target language Social appropriacy Cultural appropriacy Pragmatics
Learning process	Activity evaluation Strategy evaluation Self-assessment Goal-setting Monitoring Organising

Table 2.6: The Aspects of Metacognitive Awareness (Sinclair 1999: 102)

Sinclair (ibid 103) proposes that teachers can use these criteria to frame questions when they are discussing work with the learners. For example:

- Why did you do this piece of work?
- Why did you do it in this way?
- What is your plan for next week? Why?
- What, if any, problems did you have?
- Why did you have them?
- What did you do about them?
- What else could you have done?

If this type of question is asked systematically and consistently they can be used for evaluation purposes; Sinclair says “The extent to which the students are able to respond to such questions will provide clues as to their levels of metacognitive awareness” (ibid 103-104). She suggests three levels of metacognitive awareness which are shown by the content and the type of language used by the learners (see Table 2.7), though she stresses that further research is needed.

Sinclair’s approach uses indicators of metacognition, and so of autonomy, to produce a measure of autonomy using a guided or semi-guided interview format. With this method generalisability of results may be an issue as the criteria used would need to be quite detailed in order to begin to establish reliability between different teachers’ judgments. This would be a drawback if it is to be used as a universal measure on its own. Perhaps standardised questions could be developed. In addition, the issue of inter-rater reliability could be addressed, at least in part, by using it in tandem with other measuring techniques.

Level of Awareness	Language characterised by
Level 1 Largely unaware	Description with little or no rationale Formulaic answers Broad statements with little or no support Few or naive questions Little or incorrect use of metalanguage
Level 2 Becoming aware (the transition stage)	Greater use of: Anecdotal evidence Introspection (expression of thoughts and feelings) Use of metaphor, Speak of “epiphanies” Ask questions Use metalanguage
Level 3 Largely aware	Confident and competent use of all the above Can also describe alternative strategies

Table 2.7: Linguistic evidence for metacognitive awareness (Sinclair 1999: 104)

Sinclair's approach is confined to metacognition, or it looks at autonomy from the perspective of metacognition, and it may be that the inclusion of other areas would better reflect the multidimensional nature of autonomy, for example: the learner's beliefs about learning, psychology (e.g. learning style, confidence), motivation, knowledge, and skills. It would also require a skilled and knowledgeable teacher, who may not be available, to carry out the interviews. Being interview-based this method may also be time consuming and consequently be unsuited to larger class sizes. Metacognition does figure prominently in autonomy theory (see Section 2.4.5) and may be a key to the measurement of autonomy.

2.6 Discussion

In this section I will summarise this Chapter on the literature concerning the measurement of autonomy, and underline the conclusions which are most important for shaping the present research. The four questions posed at the start of this chapter were:

1. Is there a need for a measure of autonomy?
2. Should autonomy be measured?
3. Can autonomy be measured?
4. Does the literature provide clues as to how autonomy can be measured?

Questions 1 and 2 were addressed in Sections 2.1.3 and 2.3 and appear to support further research into a measure of autonomy. Clarification of the term *measure* (in Section 2.1.2) to distinguish it from testing and evaluation was useful in highlighting misconceptions about measuring being inappropriate to autonomous learning.

The third question regarding whether autonomy can be measured necessitated considering the conceptual and practical problems (Sections 2.4 and 2.5). The desire to measure autonomy can appear to be misguided in that it oversimplifies or misrepresents concepts of autonomy. Little (1991) presents what he considers to be five commonly-found false assumptions (1991: 3-4) about autonomy, i.e. that it is:

1. synonymous with learning without a teacher
2. the removal of teacher intervention
3. a new method that is taught by teachers
4. a single behaviour manifested in the same way by different learners
5. a steady state which is equally expressed in all areas of learning.

If these were true of autonomy, it would make the task of measuring it more simple than it actually is. This can be shown by imagining them as true of autonomy:

1. If autonomy were synonymous with learning without a teacher it would be easy to establish whether there was a teacher, and a test of learning gain would establish whether learning was taking place.
2. If autonomy were the removal of teacher intervention it would be easy to establish whether teacher intervention had been removed.
3. If autonomy were a new method that was taught by teachers, the description of the method could be checked against the method used by the teacher and the differences quantified.
4. If autonomy were a single behaviour manifested in the same way by different learners, it would be possible to observe learner behaviour and quantify it.

5. If autonomy were a steady state which was equally expressed in all areas of learning, it would be possible to measure one easily quantified area and extrapolate it to all the rest.

All of these appear to make measuring autonomy very simple, but they are very black or white. This simplicity is not matched by the picture of autonomy that has been emerging in this chapter. In order to investigate the measurement of autonomy it will be necessary to extract a picture which is more nuanced and inclusive of as much of the spectrum of autonomy (Figure 2.1) as possible. The simplistic nature of the above statements is partly because Little is making a point. In real situations autonomy is not all or nothing: it may have, for example, some elements which are in a steady state (No. 5), some behaviours which are revealing while others are not (No. 4), and there may be an element of learning without a teacher (No. 1). Measuring autonomy presents challenges in that it has been said to be:

1. context-specific and not transferable
2. highly multidimensional and hard to define
3. socially-based rather than individual
4. stifled by testing
5. not reliably observable
6. a political concept
7. variable

Yet, we also feel that we know that some learners are more autonomous learners than others. This suggests that, rather like Little's statements these challenges are not black or white absolutes. I have argued in this chapter that these points leave some space for manoeuvre.

The initial specifications for a measure of autonomy which motivated this research indicate the collection of quantified data. In addition, a measure should provide data which is most relevant, for example to the learner for self-assessment, and to the teacher to help in areas such as needs analysis. The belief was that data of this nature will be useful for learners, and may assist them in their personal and/or collective struggles for autonomy. For this reason in the present research the political-critical view of autonomy will not be directly included in the search for an autonomy measure. This is not to discount it, but to leave the values to be assigned to the data, i.e. the evaluation, to the learners or teachers themselves. This research does, then, have a practical orientation. It is hoped that it can provide a means of helping teachers by complementing the methods they currently use to gather information about learners' autonomy, which can be either difficult and time-consuming (e.g. Ravindran's CILL) or be essentially informal estimates. I feel the research will have been successful if it can show tangible advantages over either of these.

Question 4 asked if there were indications from the literature as to how to go about practically measuring autonomy. If autonomy has degrees (as indicated in Section 2.4.2) and is in that sense measurable, what are the practical ways in which measurement could be carried out? Sinclair (1999) looks for ways of providing evidence to demonstrate the effectiveness of autonomy, most of which are indirect:

- Learning gains / Proficiency gains
- Feedback from teachers and learners
- Logging learner activity
- Researching the effects of strategy training
- Evaluating the capacity

Morrison (2005) and Dam & Legenhausen (1996) have looked at learning or proficiency gains. This is not a direct measure of autonomy and correlation of learning gains with autonomy levels has not been done. If autonomy can be recognised by “increased motivation and enthusiasm and active involvement in learning” (Sinclair 1999: 97) then perhaps a questionnaire could ask learners to rate their feelings in these areas, a form of feedback, and these might be suggestive of autonomy level. Using logs or diaries to gauge levels of autonomy has the drawback that “Often students record their written comments in as economical a manner as possible” (Sinclair 1999: 98). The degree of strategy knowledge and use may offer insights into learners’ autonomy levels and it may be effective to gather information about these using a questionnaire.

Frequent concepts in descriptions of autonomy	Example of Author
Strategies	Wenden 1991: 15 Ellis 1994: 516 Cotterall 1995: 195 Oxford 1990
Skills	Littlewood 1996: 428 Wenden 1991: 15
Metacognition	Sinclair 1999: 102 Littlewood 1996: 428 Wenden 1991: 15
Confidence	Wenden 1991: 15 Fazey & Fazey 2001: 345-346 Littlewood 1996: 428 Cotterall 1995: 195
Motivation	Littlewood 1996: 428 Ryan & Deci 2000: 58

Table 2.8: Frequent concepts in descriptions of autonomy, with references

Looking at the definitions of autonomy which describe the capacities required of autonomous learners key areas emerge which offer the possibility of ways of quantifying autonomy (which are less indirect than measuring learning gains), such

as metacognition (see Table 2.8 above and Section 2.4.5 above). These areas have been the subject of research and have measures already proposed for them, for example: Motivation, Gardner & MacIntyre (1993) have produced the Attitude/Motivation Test Battery; a measure for metacognition has been considered by Sinclair (1999: 102) and; strategies have been the subject of Oxford's (1990) work to produce the SILL (Strategy Inventory for Language Learning).

Benson (1997: 25) divides views of autonomy into three types; it is seen as:

1. the act of learning on one's own and the technical ability to do so
2. the internal psychological capacity to self-direct one's own learning
3. control over the content and processes of one's own learning

Probing these areas in a way that can render a measure may be possible and if so it would suggest that an autonomy measure may be obtainable. The first classification offers the possibility of probing the learner's technical abilities. The second allows for measurement of the learner's psychological capacities for autonomy, and the third, while it is necessary to consider that there will be elements of the learner's situation which it may not be practical for him or her to control, still suggests the avenue of probing the learner's attitudes to and perceptions of his/her role and the amount of control that he/she wants. The learner's knowledge and beliefs about autonomy may be a fruitful avenue for the measurement of his/her autonomy.

Learners can have mistaken beliefs about learning; Little (1991) and Esch (1996) for example correct some common misconceptions about autonomy. Such misconceptions will affect how autonomous learners can effectively be, therefore some items in my questionnaire should address the learner's beliefs. For example a learner may believe that it is better to learn without a teacher. A learner who believes

this may neglect to obtain help when necessary and may not be aware of the skills which are required to learn autonomously. A learner may have concerns which inhibit his/her autonomy founded on the misapprehension that autonomy means that “intervention or initiative on the part of a teacher is to be banned” (Esch 1996: 37). Learner beliefs such as this may contribute to their level of autonomy and may be practical to measure using a questionnaire.

The areas that may be open to quantification and offer the possibility of more direct autonomy quantification than do learning gains or feedback from teachers are:

- The learner’s skills and knowledge of strategies
- The learner’s degree of metacognition
- The learner’s degree of confidence
- The learner’s degree of motivation
- The learner’s degree of technical abilities for learning on their own
- The learner’s degree of internal psychological capacity to self-direct
- The learner’s degree of control over the content and processes their own learning
- The learner’s knowledge and beliefs about autonomy

These will be considered in the present research.

The practical problems of measuring autonomy which have been presented in this section may be overcome by finding ways to probe the intentions of learners, by distinguishing between *measurement* and *testing*, and by realising that the problems of measuring autonomy are not in fact very different from the problems of measuring other complex concepts such as language ability, which have after much work

become generally acceptable (Benson 2010). In fact, the concept of levels of autonomy is already well accepted in the literature (Section 2.4.2 above).

To sum up, this literature review has considered the need for a measure of autonomy and has evaluated the possible conceptual and practical problems which need to be addressed. Possible routes into the quantification of autonomy have been discussed and it has been shown that there is good reason to believe that autonomy has levels or degrees which may permit measurement. A measure quantifying autonomy would seem to have a valid purpose in monitoring, aiding, and researching language learning.

Little says it is misguided to measure autonomy (2003), but Benson has written “It may simply be the case that the problems that we foresee in the measurement of autonomy appear more acute because we have, to date, largely failed to address them” (2010: 85). In the following Methodology Chapter I shall show how my research has addressed the problems of measuring autonomy. The findings of the literature review will be used to inform my research by guiding the choice and design of the methods to be used in data gathering.

3 METHODOLOGY

3.1 Overview of chapter

In this chapter I aim to explain and justify the rationale which was adopted in the present research in order to provide a clear philosophical underpinning for the research methodology. I will discuss the methodology which is appropriate for researching autonomy, and will provide an overview of the planned research and the actual stages of the research as carried out and show where this differs from the original plan.

In the two chapters following this (Chapter 4 Long List stage, and Chapter 5 Short List stage) I will present the progress of the research chronologically, showing the actual steps taken with explanations of any changes to the original plan which occurred as the research went on.

The three chapters – Methodology, Long List, and Short List – can be seen as basically chronological; they build from the theoretical foundations of the research methodology (Methodology Chapter), to the procedures carried out in the first stages of the research using 256 items (Long List Chapter), and then to the part of the research involving the distillation and use of the Short List of 50 items and its development into a questionnaire to examine the feasibility of a closed-item instrument to measure learner autonomy (Short List Chapter). In this way I hope to show how the research was based on a well thought out rationale and aimed to use appropriate techniques to answer important and specific research questions regarding the possibility of a rapid and user-friendly instrument for the measurement of autonomy.

3.2 Clarifying the research aims

3.2.1 Aims

As described in the Introduction Chapter, I wanted to investigate whether a simple, quick and accessible instrument to measure autonomy and support learners and teachers was possible despite autonomy being seen as multi-dimensional, highly context dependent, and possibly only a Western construct. If I could show that autonomy was, in a useful sense, measurable I thought it would also then help to make autonomy more accessible to “ordinary” learners and teachers.

To investigate whether such an instrument could work it would be necessary to attempt to create such an instrument and test it. It was not a research question which could be answered only by reference to the literature because opinion is divided (as can be seen in the Literature Review) about measuring autonomy.

The questionnaire would be a quantitative data gathering instrument and so it was intended to use quantitative methods to design it and probe it. However, qualitative data would be used to investigate whether learners found the instrument was useable and useful, to find what it meant to respondents, and as Dörnyei (2007) has it, to put “flesh on the bones” of the quantified data. The research can thus be seen as partly focused on the viability and appropriateness of the use of quantitative methods to research and support autonomy. In other words, this research should not be seen as representing a positivist view of autonomy but rather *an inquiry into the viability of such a view in empirical terms*. In this respect, the thesis should not be read as a conventional report on a questionnaire study and its quantitative results, but rather as a critical reflexive narrative of this process of inquiry.

3.2.2 The broad construct of autonomy to be used in the research

Autonomy is not simple and well-defined; it is considered to be multidimensional and there are differences in basic theory about what autonomy is, so this makes it hard to investigate autonomy without aligning oneself to a particular school of thought about autonomy.

I tried to keep an open mind so as to increase the possibility of discovering something new; I sampled from a broad range of literature to find what authors thought autonomy was. I did, however, have to keep in mind the original aim (i.e. exploring the viability of the instrument) but I did not want to prejudge what the contents of the items would be.

However, in order to start at all it is necessary and inevitable that options have to be limited and some general direction established in order not to be pulling in many different directions at the same time. To make it practical to begin I kept in mind the original motivation for the research which indicated the nature of the tool, i.e. it would be simple to use and require little setting up or expertise on the part of the user (whether a teacher or a learner), and this indicated a self-contained closed-item questionnaire. The research would also need to address the question of whether such an instrument would be appropriate for supporting autonomy.

I would limit the selection of items for the questionnaire, but not more than was necessary for the aim of trying to find a useful instrument to measure relevant dimensions of autonomy, i.e. which would be useful for practical formative teaching and learning purposes yet still be recognisably autonomy – at least as seen by some of the authors who had published their theories about its nature. It had to fit in to the

field in a way that made it recognisably part of the field and so it would need to have construct validity.

What were these “relevant dimensions” of autonomy? Since the purposes of the instrument being investigated in this thesis were related to the practical support of teachers and learners in the development of autonomy the relevant dimensions would be those which had immediate and pressing relevance to teachers and learners in the bottom-up support of learning. This would indicate that the relevant dimensions would relate closely to learning English more effectively, in a more self-sustaining way, not dependent on teacher control – on “carrots” and “sticks” education – but having the learners take more control of their learning, and so being able to learn better and move towards their individual potentials.

3.2.3 Construct validity checking

The questionnaire, being made up of items chosen to reflect the field as found in the literature, would embody a view of autonomy; it would in effect be a “theory” of what autonomy was. However, the construct of autonomy intended for any questionnaire by its designers may not necessarily be successfully instantiated in the finished instrument. The statistical technique of factor analysis (see Section 3.12.3 below) can be used to establish whether factor groupings in the data match the groupings which were intended by the designer (i.e. confirmatory factor analysis, see Section 3.12.3.2.1). Alternatively, a questionnaire can adopt a more exploratory approach where the initial construct is broader, and then exploratory factor analysis (see Section 3.12.3.2.2) is carried to discover the underlying groupings within the data collected by the questionnaire. This picture can then be compared with the literature to see if it is a model which is “consistent with theory” (Miller, Maltby,

Fullerton & Acton 2002: 184). This would be a check of the construct validity of the instrument. In the present research this approach was adopted to construct validity. This check would not look at whether the instrument was actually successfully (or indeed usefully) *measuring* learners' autonomy (see Section 3.2.4 below for discussion of the measuring aspect).

Based on a review of the relevant literature I would compile a "Long List" of items. Factor analysis can be used for data reduction (i.e. it could reduce the number of items) and in the present research this was originally intended as a way of removing the less relevant or more poorly performing items.

This item selection procedure would have the advantage of not fixing a precise construct of autonomy in advance of the empirical data. Nevertheless, it would not be possible to have a completely open attitude to the construct for practical purposes of questionnaire length; it would be necessary to limit the number of items in the Long List to narrow the construct. This would be achieved by focusing on the aims of the questionnaire and thus gaining an indication of the more relevant dimensions (the areas are discussed in the Literature Review Section 2.4). I am focusing on "autonomy for language learning" i.e. as a means to the end of learning a language, rather than "language learning for autonomy" i.e. seeing autonomy as the goal of learning rather than the target language (Benson & Voller 1997: 2), and one of the less relevant views of autonomy for this aim is political interpretations of autonomy.

How could I tell if the eventual questionnaire was actually measuring autonomy?

The plethora of autonomy definitions meant that I did not expect the questionnaire to measure some "quintessence" (Benson 2009: 20) of autonomy – it could not be expected to do that. This is beyond the scope of the present research, which as

previously stated is to explore the possibility of a quick and simple autonomy measure to aid teachers and learners rather than an absolute universal standard measure of autonomy. In my experience teachers make and review informal estimates of their learners' autonomy and so my research is relevant to this area of classroom practice where it would appear that a measure could inform or complement teacher estimates.

3.2.4 The measuring aspect

The autonomy measuring aspect of the questionnaire would need to be compared with an acceptable way of measuring autonomy which was independent of the questionnaire in the sense that there would be no cross-contamination of evidence, however indirectly, and one source must not depend on or influence the other. The results of each would be compared to see if they could be said to agree to a significant degree. It would be ideal if the other source was a validated autonomy measuring questionnaire, but there is no such instrument (which is partly what inspired the present research).

The most common way in which autonomy levels are judged is self-assessments by learners of themselves; there are more students than there are teachers, so self-assessment is likely to be more common than teacher assessments. The next most common will be teachers' assessments of their students. My sense, based on my own experiences as a teacher and self-access centre coordinator, is that this is normally carried out in a quite *informal* way (though, as seen in Section 2.5 there have been some more formal attempts).

The absence of a measure meant that (most) teachers were in effect measuring their learners' autonomy based on their exposure to the learners, but not in a principled way using a fixed procedure. If the eventual questionnaire could match or exceed teacher estimates then it would be useful because it would be standardised and repeatable, and would be quicker since no long exposure to the learners would be necessary as is the case with teacher estimates. The teacher will in effect estimate the autonomy of a student based on observations, homework assignments, punctuality, apparent attitude, conversations, beliefs about the students' cultures, etc. (It would be an interesting project to investigate how teachers make their estimates and how accurate they are, but this is beyond the scope of the present research.)

For the questionnaire to be shown to be a worthwhile measure it would need to be shown that it had a net advantage over the currently available methods of informally measuring autonomy, judged in terms of validity, reliability, speed, convenience, timeliness (e.g. available at the start of a course, or at the learner's introduction to a self-access centre), and standardisation of procedure. This would be a major advantage for teachers, and especially for learners (who may be less expert in terms of knowledge of pedagogy and autonomy).

It was thus originally intended that the data produced by the questionnaire would be compared with the measurements produced by teachers as these estimates are the de facto method in most autonomy-aware classrooms.

3.2.5 Qualitative aspect

I felt it would be useful to probe more deeply to see what teachers' and learners' beliefs about autonomy were and how they related to the results of the questionnaire

and therefore interviews were planned (see Sections 5.2.2 and 5.2.3). Even if a relationship would eventually be found between teacher estimates and questionnaire data “the questionnaire data usually reveals [sic] little about the exact nature of the relationship” (Dörnyei 2007: 170) and so interviewing respondents and teachers could potentially add “flesh to the bones” (Dörnyei 2007: 171).

In addition, I would use the instrument with my own preessional group to gain experience of its use in class for myself (see Section 5.4). This would be part of a more qualitative strand in the overall research project (see Section 3.7.2). The project can thus be described as being of the mixed methods type, utilising as it does both quantitative and qualitative methods.

This part of the research, as it transpired, was to be a key area in the eventual outcome of the research as it led me to a new understanding of autonomy and a reappraisal of my aims in this research and a change of emphasis more towards the utility of the instrument for promoting autonomy and less towards its measurement (see Section 8.2.1).

3.3 Research questions

As stated in the Introduction I had the creation of a simple yet viable instrument to measure autonomy as the specific motivation for this research. This would entail both theoretical and practical approaches: on the theoretical side an investigation of the literature of autonomy to find what had been done in this area and to problematize the endeavour from the theoretical side; and practically to design and validate an instrument and, put simply, see if it “worked” by checking its results against some other acceptable standard. To make the research worthwhile it would be

important to show that there were advantages to the use of such an instrument.

Therefore the two research questions are:

1. Can a closed-item questionnaire be used to provide a practical and viable measure of autonomy? (What are the issues involved and can they be overcome? Is there a place for quantitative techniques in the support of autonomy?)
2. What are the uses of a closed-item-questionnaire autonomy-measuring instrument? (What are the advantages and disadvantages? If it works is it just a mass statistical tool or can it be useful for individuals?)

In terms of the quantitative aspect of the research my “null hypothesis” would be “there is no relationship between the simple closed-item questionnaire used in this research and autonomy levels as given in a de facto alternative measure (teacher estimates), and the questionnaire would therefore serve no useful purpose in supporting the development of autonomy for language learning”. This would be falsified if a comparison of the questionnaire’s results with a statistically significant sample of teacher estimates showed a significant correlation. The advantages and disadvantages of the two ways of measuring autonomy, such as speed, necessary training, accessibility (for learners and teachers), support for an individual’s autonomy etc., would also need to be considered in answering the second research question. In addition, the quantitative approach embodied in the questionnaire is itself under investigation in my exploration of the measurement of autonomy by means of both quantitative and qualitative methods including reflection on my own experiences of using the questionnaire with a class; one possible outcome, therefore, is that I may be able to judge for myself the appropriateness of using the instrument

to measure autonomy, and also be in a position to consider the type of use to which it is suited.

3.4 *What would the instrument be like?*

I had envisioned the autonomy-measuring instrument as a simple questionnaire, but it was necessary to establish more precisely what form it would take and how this could be justified in terms of the principles of instrument design. In this section I will discuss these issues and arrive at a more specific description of the desired instrument.

Oppenheim (1992: 10) warns that “A questionnaire is not just a list of questions”; it has to “speak for itself” and cannot so easily incorporate clarifications and probes as can an oral interview. This meant that I had to exercise care with the wording of items (see Section 4.2.1) and as I would be using it with non-native speakers this was doubly important and influenced my decision to use translation (see Sections 4.3.1 and 3.13).

The advantages of using a questionnaire for my autonomy-measuring project include (Gillham 2000: 6): low cost in time and money; it is easy to get information from a lot of people very quickly; and the analysis of answers to (especially) closed questions is straightforward. In addition, Aiken (1997: 46) points out that they “yield a great deal of data on numerous variables” and Dörnyei (2003: 10) adds that they are versatile, being able to cover a variety of topics, people, and situations.

However, questionnaires do have drawbacks which I needed to be aware of and respond to. Firstly, data quality may be low as questionnaires do not provide the opportunity to check with the respondent to correct mistakes, and responses may be

superficial or insincere (Gillham 2000: 8; Aiken 1997: 46; Dörnyei 2003: 10-11). I planned to minimise this by making the items as clear as possible in their wording and by the use of translation. Where the final version of the questionnaire is being used in class the teacher may in fact have the chance to respond to the answers and so check the data quality.

A second issue concerns the psychology of questionnaire response, such as acquiescence bias, i.e. the “Tendency to answer affirmatively (yes or true) to questions or items on questionnaires” (Aiken 1997: 277). This happens particularly when respondents are unsure if they agree or not with an item (Dörnyei 2003: 13). Converse & Presser (1986: 38) however note that acquiescence is more noticeable in the less educated. I limited acquiescence bias by including a “don’t know” option among the possible responses after the first administration of the questionnaire in the United Arab Emirates (UAE).

Thirdly, Gillham (2000: 8) and Dörnyei (2003: 10) both note the problems of motivating respondents. For most people filling in a questionnaire is “an activity which typically they do not enjoy or benefit from in any way” (Dörnyei 2003: 11). I hoped to motivate respondents by administering the questionnaire in class (not impinging on their free time), and offering incentives (see Section 4.3.4), but the voluntary Internet data gathering (see Sections 4.3.2 and 4.3.3) did have low response rates, no doubt partly due to low motivation. When the final version of the instrument is used in class the learners will experience benefits as the teacher will be able to respond to what the learners express in their responses and it is therefore reasonable to hope that motivation will be enhanced.

Fourthly, people often choose the most desirable answer rather than the truthful answer, a feature Dörnyei (2003: 12) calls “Social desirability (or prestige) bias” and which Aiken (1997: 277) describes as:

the tendency on the part of a person to respond in what he or she judges to be a more socially desirable direction, rather than responding in a manner that is truly characteristic or descriptive of him- or herself.

I would need to make clear that I would keep the data private and they would not be shared with others or contribute towards any course grade.

In addition, Dörnyei (2003: 14) warns of the “Fatigue effects” of long questionnaires. To avoid these I would provide a response format which only required ticking a box (while giving respondents the opportunity to make their own comments at the end of the questionnaire if they wished) and I would ensure that the final version of the questionnaire would be relatively short.

3.4.1 Questionnaire type

Oppenheim (1992: 102-103) sorts questionnaires into three varieties using the mode of administration: interview, self-administered, and group-administered. I wanted the instrument to be usable in a self-administered way as well as in class. This indicated a questionnaire that was designed to be self-administered as this could also be used with groups, whereas a group-targeted instrument would not be useable with solitary individuals. In a self-administered situation, Bourque & Fielder feel it is imperative that “the questionnaire be completely self-sufficient, or able to ‘stand alone’” and that the questions must be closed-ended (Bourque & Fielder 1995: 17).

Questionnaires have also been classified by the level of structuredness (e.g. Gillham 2000: 2-3; Cohen et al. 2000: 247) indicated by the balance of open and closed items. According to Nunan (1992: 143) “An open item is one in which the subject can decide what to say and how to say it” whereas closed items, such as multiple-choice questions, have the “range of possible responses [...] determined by the researcher” (Nunan 1992: 143). Questionnaires can be structured (only closed items), unstructured (only open items), or semi-structured (a combination of closed and open items). A major advantage of structured questionnaires is that large quantities of data can be gathered and analysis is rapid. I had to decide which type of questionnaire would be most suitable, and as I wanted the instrument to be able to gather data in SACs and via the Internet as well as in classrooms I chose self-administered because this is the most flexible being useable in or out of the classroom. For ease of data processing and also for ease of use by teachers with large classes I decided on a structured questionnaire with items of the closed type (though I also provided space for respondents to give their feedback for questionnaire development purposes).

I now had to consider the structure and formatting of the questionnaire; length, how items should be grouped, the most effective way to order the questions, the features of the essential constituent parts of a questionnaire, and how to make it look relatively appealing are all issues which are seen as important in the literature. I deal with these in Section 3.10 “Questionnaire development”.

3.5 Methodological stance/paradigm

I am investigating in a critical reflexive way the viability of a positivist method (the questionnaire) in empirical terms in order to probe whether it can viably be used in supporting autonomy, and consequently this thesis is not a standard report on a

questionnaire study and its quantitative results. I use quantitative methods in the development of the instrument and in the larger scale data gathering and analysis, but I also employ more qualitative methods such as interviews and use my own experiences and reflections to explore the use of the questionnaire in a hybrid mixed methods approach. This is the subject of the following section.

3.5.1 Mixed methods research

Methodology is “a set of principles for choosing between procedures” Riley (1996: 253) and Teddlie & Tashakkori (2009: 300) advise that “research questions dictate the research design and procedures”. The aims which underlie my research questions should therefore indicate the appropriate methodology to adopt in the present research. The choice is not limited to one of either qualitative or quantitative. Research using complementary methodologies, or mixed methods research, is being recommended in the language learning field in general, for example by Dörnyei (2007), and more specifically for autonomy, for example by Riley (1996). In fact Riley sees mixed methods research as the most appropriate for autonomy research. Contexts for autonomy, such as self-access centres and classes, have both material and social forms which:

make it pointless to adopt an approach to research which is exclusively ‘qualitative’ or ‘quantitative’, ‘positivist’ or ‘non-positivist’. This will imply methodologically hybrid solutions, which are not easy to find or implement. (Riley 1996: 264)

If mixed methods research is indicated for researching autonomy then in the present research I should aim to combine quantitative and qualitative methods if and when appropriate (Punch 2005: 241).

Mixed methods research can be adapted for different reasons. In the present research for example there are three main considerations. Firstly, mixed methods research is adopted in order to account for both the researcher's and the subjects' perspectives on the autonomy measuring questionnaire. Punch (2005: 242) explains that:

Quantitative research is usually driven by the researcher's concerns, whereas qualitative research takes the subject's perspective as the point of departure. These emphases may be brought together in a single study

Secondly, comparison of questionnaire data and teacher estimates may provide quantitative evidence for the measurement validity of the instrument, but such evidence is often weak when exploring the reasons for relationships and a more qualitative study can be used to complement it (Punch 2005: 242). Thirdly, mixed methods research "may provide a means of bridging the macro-micro gulf" (Punch 2005: 242); purely quantitative research is effective at the large scale, and qualitative research is more suited to the smaller-scale aspects.

Punch (2005: 241) points out that it is necessary to decide whether to combine methods, data, findings, or all three. They should be integrated in a way such that the research design displays "complementary strengths and nonoverlapping weaknesses" (Dörnyei 2007: 63).

Mixed methods are appropriate for researching autonomy for a number of reasons. They can help to make the results of the research acceptable to more readers. They can show that the researcher has awareness of the complexity and/or multidimensionality of questions about autonomy. Mixed methods can help to reduce the chances of a situation akin to Riley's (1996: 251) story of the blind man and the bubble, where the object of the research is not amenable to the means of research.

Certain research questions require a quantitative aspect and a qualitative one. The present research, for example, concerns a quantitative instrument and its viability as a support for learners and teachers. At the macro level, therefore, a quantitative approach is indicated, but at the micro level of understanding individuals' experiences of the instrument and how it applies to them a more qualitative approach is indicated, and each could inform the understanding of the other.

However, mixed methods research is not without problems. Dörnyei (2007: 174) says the reasons why there is not more mixed method research is because many researchers do not feel confident that they can do themselves justice in both quantitative and qualitative areas for various reasons, and also that:

it requires considerable effort to study a phenomenon with two (or more) separate methods, and in the light of all this it is understandable that many (if not most) researchers may prefer to remain on monomethodological grounds.

Mixed methods are indicated for research in autonomy, but mixed methods research is a challenging way of researching, and autonomy is a challenging area to research (especially the measurement of autonomy). In the following section I will consider how other researchers have approached the problem of researching the measurement of autonomy.

3.6 *Other researchers*

In this section I aim to show how my research is informed by the previous autonomy measuring research discussed in Section 2.5. I am investigating a measuring tool, so my research is perhaps closer to Ravindran (2000) and Lai (2001) and less like Dam (2000) as I am not looking at evaluation. I will be probing a quantitative tool, so I

will have to use quantitative techniques to check it in its own terms in order to establish it as a quantitative instrument. Then my research would move to a more qualitative area to see what it “means” and how users feel about it, so gaining feedback on it in qualitative terms. Like Dam (2000) I would probe indicators of levels of “customer satisfaction”.

Adoption of the techniques used in the literature was limited both by my research aims and by the resources available to me. This also served to remind me that the target users of the eventual instrument would also have limited resources. It would not be possible to establish lengthy procedures. Ravindran (2000) had a team which would, over the course of three years, carry out regular training, monitoring, and feedback sessions, and keep detailed records for formal certification. To follow Lai (2001) a teacher would need training and her method would also require much of a teacher’s time if a whole class was to be assessed. Sinclair’s (1999) suggested approach would require a skilled and knowledgeable teacher, who may not be available, to carry out the interviews. Being interview-based this method may also be time consuming and consequently be unsuited to larger class sizes.

Ravindran’s (2000) research concerned formal testing and certification which required rigorous checks to maintain the necessarily very high levels of reliability for the three-year course. This would not be possible in the present research, and in fact the aims of my research required a quick and simple instrument which did not aim at formal certification but at providing a useful support tool. This governs the levels of validity and reliability which will be aimed for in the research.

Dam (2000) does not address reliability and validity. Ravindran (2000) does make great efforts to establish reliability, but rests construct validity on the item selection

which is based on only one author. Lai's (2001) research does not involve comparison of results with another measuring method and so here too validity is not established. It seems that validity is a problematic area in attempts to measure autonomy, and is an area which I will need to address in the present research.

Linked to the question of construct validity is the choice of the construct of autonomy. An intuitive way of approaching measurement is to, in effect, pre-define autonomy and then design an instrument to measure the (measurable) elements of the definition. This is basically the approach adopted by Cotterall (1995), Lai (2001), and Ravindran (2000). I wondered if this could be improved on or modified because I was sensitive to the criticism that Ager (1996: 144) made of performance indicators that they:

measure outputs rather than outcomes, they reduce complex situations to simple numbers, they affect 'performers' so that they perform to indicators, they measure only what can be measured

Or, as Riley (1996: 259) says, "counting what could be counted instead of what counts". I wished to avoid being led by ease of measurability when defining autonomy and also therefore I decided I would follow an item selection procedure which was as inclusive as practicable in the time available, was based on the literature, and would be checked for construct validity (using factor analysis – see Section 3.12.3). I also wanted to avoid providing another definition of autonomy or choosing one extant version – for example, Ravindran (2000) chooses elements to include in the assessment based on Knowles (1975). A large sample of items was indicated, taken from the range of autonomy ideas that were pertinent to the aims of the research.

Cotterall (1995) uses factor analysis and it appears to be a promising technique for interpretation of data. However, the process of interpreting factors appears to involve subjective judgement and is not well defined. In order, therefore, to minimise subjective variations I would seek to establish a procedure which I would follow for all the factors I would be interpreting. Cotterall (1995) identifies factors with as few as two items, which is too few to give a clear indication (see Section 3.12.3.3.5 below). This also points to the need for clear standards to be adopted for factor analysis (see Section 3.12.3.3 below) which are overt and therefore repeatable.

I felt there was an issue in some of the research (Dam 2000; Ravindran 2000) around making assessment of autonomy part of compulsory course work, and Lai (2001) also gave models which the learners had to follow. This appears to reduce the learner choice intrinsic to autonomy and I wanted to avoid this in my research. This was one reason I attempted to make the questionnaire for the present research “free-standing”, i.e. not a part of a course.

3.7 The research design

In designing mixed methods research Punch recommends that three questions should be answered:

Will the two approaches be given equal weight?

Will they be interactive or separate?

How will they be sequenced? (Punch 2005: 241)

In this section I will make clear my answers to these questions.

I am researching the viability of a particular tool which is a quantitative instrument designed to measure autonomy, so I have to use quantitative methods to make and

validate the questionnaire. I can show how it performs quantitatively in one particular i.e. by means of comparison with teacher estimates of their students' autonomy levels.

The qualitative part of the research comes in because to understand what the research means it is necessary to talk to those involved and find what they were thinking, what the questions meant to them, how they decided their answers, what the teacher estimates were based, how strongly they were felt, and how confident the teachers were about them. This is important to know also for understanding what the instrument was measuring, and so could help to feed back into the cycle of development of the instrument. The qualitative part would necessarily have to be smaller scale, and so one or two groups of students and their teachers would be studied, i.e. the questionnaire would be completed and the parties interviewed by me. In addition I would myself use the instrument with a class I was teaching and use this experience to inform my reflections on the research. This qualitative aspect of my research would prove to have a major influence on my interpretation of the findings and on the conclusions of this thesis.

3.7.1 The plan for the research

The research design involved both large scale more quantified data collection and smaller scale more qualitative data gathering.

At the larger scale the responses from the questionnaire items would be coded and quantified for analysis and comparison of questionnaire data with teacher estimates of learner autonomy would be carried out. At the smaller scale I would be investigating whether the quantitative questionnaire was appropriate for supporting

autonomy in class: does it serve a useful purpose validly and reliably; does it have the potential to replace, or improve on, or add to teacher estimates?

3.7.1.1 Larger scale data collection

Larger scale data collection was for three purposes: item selection; construct validity checking; and comparison of the questionnaire data with teacher estimates.

3.7.1.1.1 Item selection

After compiling the initial Long List (I refrain from calling the Long List a questionnaire as it was not the finished questionnaire), the intention was to make it available online and to collect a few hundred responses. Originally, the intention had been to use factor analysis to perform data reduction and thereby form a shorter list of key items for the questionnaire. However, in fact too few responses came in too slowly for this and items were selected using other statistical means (see Section 5.1). Data collection was augmented by paper-based and email-based means to improve the quantity of data available. I also intended to collect feedback from respondents to find any problem items, e.g. those which were not clearly worded.

After item selection the now much shorter list (the “Short List”) of items would contain the items which correlated most strongly into factors. This list of grouped items would be used to form a questionnaire designed to measure autonomy. It could potentially become an autonomy measuring instrument.

3.7.1.1.2 Construct validity checking

The Short List would be used to form a questionnaire and more data would be gathered. The collection of a large amount of data would enable a check on construct

validity to be carried out using factor analysis of the gathered data. The resulting picture of autonomy would be compared with the literature to see whether it was in accord.

3.7.1.1.3 Comparison of teacher estimates and questionnaire data

I originally intended to collect a large amount of teacher estimates to compare them with the corresponding questionnaire data. The purpose of this was to establish the comparative validity and reliability of the questionnaire. The estimates and the questionnaire would be completed at the beginning and end of a course. It would be expected that the teacher estimates would become better with longer exposure to the subject class. This was an assumption since it would appear logical that increased familiarity and knowledge of a class would lead to better estimates. The questionnaire's performance at first and second administrations would not be expected to benefit in any way from the intervening time period in the way that the teachers' performance would. The estimates of the teachers would therefore be expected to move towards the level of the learners' autonomy over time, but the questionnaire results would not be expected to move towards increased accuracy. This is not to say that the questionnaire results will not vary over time – as seen in the literature review it is accepted widely that autonomy varies (see Section 2.4.2). If the results of the questionnaire and the teacher estimates were to move closer over time, i.e. between the first and second administrations, either viewed for individual learners or in terms of class averages, then it would be very suggestive of a change in the teacher's success in estimating autonomy rather than a change in the questionnaire's "ability". Convergence would therefore be a positive result for the questionnaire. If there were divergence it would strongly suggest that the

questionnaire was unreliable over time. If the results moved but remained equally separated with no convergence it would suggest that the questionnaire was matching the estimates but with a bias – this would also be suggestive of a positive result since the movement would indicate a change in the autonomy level of the class which was being picked up by both methods of measurement. A recalibration of the instrument in this situation would theoretically be one possible solution to the disparity.

With larger samples it is easier to establish a significant correlation. This would mean that a number of classes and teachers would be necessary for a definite result to be found, either supporting or not supporting the questionnaire as an alternative to teacher estimates. If the required quantity of classes and teachers was not found any results could not be shown to be significant, and could be no more than suggestive as small samples may be idiosyncratic. Problems did mean that only two sets of teacher estimates could be gathered and therefore statistically significant data would not be gathered.

3.7.2 Smaller scale data collection

This would be more qualitative with concentration on one or two small groups. It would “put flesh on the bones” and would enable a comparison of questionnaire, teacher estimates, and teacher and learner interviews in a more detailed way.

Regarding interviews, Dörnyei says, with obvious relevance to the present research, that conducting interviews with respondents can help to find what they “really meant” and that “This design pattern can also be used for validating test results with a newly developed test” (Dörnyei 2007: 171).

Delays caused by translation, and difficulties with sourcing respondents meant that the large scale comparison of teacher estimates and questionnaire data did not take place and so only the small scale stage of comparison took place. This meant that the production of statistically significant results was not possible. The small scale stage could not substitute for the large scale in terms of statistical significance, but it would still be valuable in two ways. Firstly, it could be a useful rehearsal and trial of the methods intended for the large scale research which may suggest indications of whether there was a correlation. Secondly, it would provide qualitative data as it would shed light and provide examples of what the data may mean in terms of the experiences of individual students and teachers.

I intended to administer the questionnaire to my own preessional class as part of this stage (see Section 5.4) so that I would have first-hand experience of estimating levels and observing students in class which I could relate to their results from the questionnaire. The qualitative experiences of teaching and interviewing learners in conjunction with the questionnaire and making estimates would, in fact, vividly affect my own understanding of the nature of autonomy and my conception of what I was attempting in this research into the measurement of autonomy in a way which I had not foreseen when this stage was initially planned, and so this part of the research was, as it transpired fundamental to the eventual outcome of the thesis (see Sections 7.4.5 and 7.5, and Chapter 8 especially Sections 8.2 and 8.7).

3.7.3 Sampling procedures

La Ganza (2002: 48) stresses that there are many varied contexts for learner autonomy, such as in a workshop with a teacher, in a language laboratory with an advisor, or sitting under a tree with a book. My research was focused on tertiary level

learners of English which indicated class groups of learners as one readily available source, and also the Internet, which would help my sample to be as wide as possible to enhance generalisability, but it would still be relatively small scale due to limits of time and resources. The larger scale and smaller scale data collections would have different sampling procedures.

Punch (2005: 103) gives the three questions which need to be answered in respect of a sampling strategy: How big will the sample be, and why?; How will it be chosen and why?; and What claims will be made for its representativeness? These questions will be addressed here, and where appropriate the large scale and small scale procedures will be dealt with separately.

3.7.3.1 Sample size

Punch's first question is "How big will the sample be, and why?"

3.7.3.1.1 Larger scale

The minimum size of the sample for the larger scale data collection (see Section 3.7.1.1 above) was governed by the intention to use factor analysis for data reduction. (The necessity of carrying out factor analysis for construct validation, and the need for sufficient participants to achieve statistical significance in the comparison between the questionnaire and teacher estimates also indicated a large sample.) In the Long List there would be 256 items and in order to carry out factor analysis it would be necessary to achieve a high ratio of respondents to items. Advice on the size of the ratio varies considerably (see Section 3.12.3.3.1 below) and will depend on the characteristics of the individual data set. However, it was intended to

achieve some hundreds of responses. In reality at the stage of data reduction the sample was not large enough and alternative methods were used (see Section 5.1).

3.7.3.1.2 Smaller scale

Sample size would not be governed by the significance level statistics which are needed for quantitative proof as this would be a qualitative exploration of the response of individuals to the questionnaire, and so a much smaller size sample would be acceptable.

3.7.3.2 How sample chosen

Punch's second question is "How will it be chosen and why?"

3.7.3.2.1 Larger scale

I wanted to investigate the questionnaire with a range of students though always keeping to the specification of language learners in tertiary education. Within this specification there is a wide range of qualities the respondents may have in terms of age, gender, major, L1, and many more.

3.7.3.2.2 Smaller scale

In order to examine the questionnaire in a realistic situation and obtain a sense of how it would perform the sampling would therefore be chosen as being one where I could be the teacher. I would also aim to find another teacher who would use the questionnaire with their class. I would then interview the teacher and their students.

These samples would be chosen on the basis of availability as at Warwick University the available samples would be appropriate to the target of the research as they

would consist of tertiary-level Chinese learners of English (Chinese students form the largest component of the overseas student population in the UK).

3.7.3.3 Sample representativeness

Punch's third question is "What claims will be made for its representativeness?"

Table 3.1 below shows recent statistics for overseas students in UK higher education, and Table 3.2 shows the countries of origin of the respondents in the present research. In both there is a clear preponderance of Chinese students. However, the backgrounds of the respondents do not reflect a deliberate sampling strategy to reflect the overseas student population in the UK, but rather the availability of subjects, which will reflect the idiosyncratic situation of the research and contacts which I had. It is thus a convenience sample and not a random sample of the world's tertiary English language learner population. My department had contacts with China and a number of Chinese students studying in it or in the preessional courses. Pre-essional courses will tend to have more students from countries which do not have significant use of English as a first or official language, and language backgrounds which make English more difficult to learn. This may explain why there are so few in my sample from Europe and the Asian subcontinent. In my work I am exposed to students who have problems with English and this population will be different from the general overseas student body. The items in the Long List and questionnaire were in English or translated into Arabic and Chinese. When translation became an issue (see Section 3.13) it was hoped to provide more languages, but time was not available. The extent of the questionnaire's availability in translation may have had an influence on the composition of the sample. Any claims which will be made for the representativeness of the sample must therefore reflect this.

Region	Total		
	2008/09	Change since 2004/05 %	Proportion of all non-UK countries %
Total (all non-UK countries)	368,970	15.9	100
China	47,035	-10.7	12.7
India	34,065	104.2	9.2
Ireland	15,360	-6	4.2
Nigeria	14,380	76.6	3.9
United States	14,345	-0.3	3.9
Germany	14,130	12.5	3.8
France	13,090	12	3.5
Malaysia	12,695	10.6	3.4
Greece	12,035	-38.9	3.3
Cyprus (European Union)	10,370	82.7	2.8
Pakistan	9,610	46.8	2.6
Hong Kong	9,600	-10.9	2.6
Poland	9,145	318.5	2.5
Italy	6,035	13.5	1.6
Spain	5,690	-5.2	1.5
Canada	5,350	27.7	1.4
Taiwan	5,235	-11	1.4
Saudi Arabia	5,205	113.3	1.4
Thailand	4,675	18.7	1.3
South Korea	4,275	11.2	1.2

Adapted from HESA 2010, table 6a. Cited in UK Higher Education International Unit (2010: 8)

Table 3.1: Top 20 countries of origin for non-UK students at UK higher education institutions, 2008-2009

Country	Number	Proportion of total (%)
China	88	47.57
UAE	54	29.19
Taiwan	25	13.51
Thailand	6	3.24
Turkey	2	1.08
UK	2	1.08
Canada	1	0.54
Colombia	1	0.54
Italy	1	0.54
Japan	1	0.54
KSA	1	0.54
Sweden	1	0.54
Switzerland	1	0.54
Syria	1	0.54
TOTAL	185	100.00

UAE=United Arab Emirates UK=United Kingdom KSA=Kingdom of Saudi Arabia

Table 3.2: Countries of origin of respondents

3.7.4 Subject groups

Table 3.3 below shows the composition of the subject groups in the total combined sample used in the present research. Also indicated in the table are the sections where more description can be found of the data gathering for each group.

3.8 Reliability and validity

The reliability and validity of the eventual instrument will govern what it can justifiably be used for as, for example, the standards for a formal test are much higher than those for a formative class activity. Similarly, if the eventual instrument were to be used as a research instrument the conclusions which could be drawn from its data would have to respect the demonstrated limits of the reliability and validity.

3.8.1 Reliability

Dörnyei says “reliability indicates the extent to which our measurement instruments and procedures produce consistent results in a given population in different circumstances” (2007: 50). He reminds us that reliability is not a property of an instrument but rather “is a property of the scores on a test for a particular population of testtakers” (2007: 50). At the larger scale level in the present research the internal reliability of the scores will be found by means of the Cronbach’s alpha of the component groupings produced by the factor analysis of the data after the final data gathering (see Section 3.12.3.3.3). Dörnyei (2007: 207) recommends reliability figures above .7 and not below .6. Cronbach’s alphas tend to increase with the number of items in the scale (Field 2005: 668) so this will be taken into account with the instrument’s scales.

Sample name	Sample size	Sex M/F	Av. age (years)	Av. time studying English (years)	Countries of origin (N)	Av. time taken (minutes)	Data gathered using the 50 items in:	Section
HCT	54	0/54	NA	NA	UAE (53), UK (1*)	NA	Format A	4.3.1
English Internet	8	0/8	38	15.6	UK (1**), Colombia (1), Sweden (1), KSA (1), UAE (1), China (1), Switzerland (1), Canada (1)	31	Format B	4.3.2
Chinese Internet	6	2/4	22.5	12	China (6)	26	Format C	4.3.3
BNU	48	5/42 U=1	20.3	8	China (48)	28.5	Format D	4.3.4
ELTCS year 3	10	2/8	NA	8.9	China (9), Taiwan (1)	16	Format E	5.3
ELTCS pre-sessional	10	7/3	20	9	China (10)	NA	Format E	5.4
Other pre-sessional	35	15/20	25.5	13.6	China (14), Taiwan (10) Thailand (6), Turkey (2), Italy (1), Japan (1), Syria (1)	9.7	Format F	5.5
Taiwan	14	3/9 U=2	23.5	8.5	Taiwan (14)	27.5	Format E	5.6
Total	185	34/148 U=3	25	10.8	China (88), UAE (54), Taiwan (25), Thailand (6), Turkey (2), UK (2), Canada (1), Colombia (1), Italy (1), Japan (1), KSA (1), Sweden (1), Switzerland (1), Syria (1)	Long List=28.5 Short List=17.7	Format A=54, B=8, C=6, D=48, E=34, F=35	

U=Unspecified sex. NA=Not Available. *studying French. ** Studying Japanese. UAE=United Arab Emirates. UK=United Kingdom. KSA=Kingdom of Saudi Arabia. HCT=Higher Colleges of Technology. BNU=Beijing Normal University. ELTCS=English Language, Translation and Cultural Studies.

Table 3.3: The sample groups involved in the research and the format of items used

At the smaller scale, data from individuals or small groups, such as those obtained from the English Language, Translation and Cultural Studies (ELTCS) preessional group (see Table 3.3 above), can be examined for consistency with the “treatment” which the student would undergo, in this case a five-week preessional course. If there is a correlation between questionnaire and treatment then it would be an encouraging indication. The reliability of the questionnaire is discussed in the light of the data obtained in the Data Analysis Chapter (Section 6.4).

3.8.2 Validity

The routes to validation which were chosen as being practicable within the timescale of the present research were:

- Comparison of questionnaire data with teacher estimates
- Comparison of the construct embodied in the questionnaire with constructs and previous research found in the autonomy literature (construct validity)
- At the smaller scale, comparison would be made of the interpretations of the results of the questionnaire with information gathered from interviews with teachers and learners and observations made with my own class.

The construct validity check would help to ensure that the construct was not representing a view of autonomy chosen only for its ability to be measured. It would have to be a construct which was recognisably autonomy (Miller et al. 2002: 184). Therefore, it was important to carry out a construct validity check on the concept of autonomy which was represented by the items in the questionnaire. In order to achieve this, the statistical technique of factor analysis would be used to reveal the underlying model. This could be different from the expected model, i.e. the areas on

which the item selection was based, but the picture which was revealed by factor analysis will be compared with the literature to establish whether the questionnaire is representing an acceptable construct of autonomy. (Factor analysis is discussed in Section 3.12.3 below.)

3.9 *Item selection*

In this section the rationale for the choice of the Long List of items is explained and justified. The details of this stage of the research are reported in Section 4.2.3.

As it was intended to have learners respond to all the items, practicalities of length meant that judgement would have to be used to decide how many items were acceptable to cover each area (see Section 3.10.1.1 for discussion of the issues around length of questionnaires), with the probability that it would be necessary to have some items covering more than one area to economise on the number of items (Table 10.2 shows the items which were selected and the areas they were intended to cover).

The literature would be surveyed in order to provide the rationale for the choice of the Long List of candidate items for the eventual questionnaire. Items would be chosen to represent areas which the literature suggested were important to autonomous learning. These areas of autonomy have been presented and discussed in the Literature Review.

The items would be chosen with regard to the aims of the research, that is, to investigate the possibility of a quick closed-item questionnaire whose purpose was to measure autonomy (see Introduction Chapter) and therefore the items should represent the relevant views of autonomy and the relevant purpose of autonomy for

language learning, rather than political conceptions of autonomy (see Literature Review Section 2.4.1 and Section 3.2.3). The construct of autonomy which was used to select items for the questionnaire is discussed in the following section.

3.9.1 The construct of autonomy used in the present research

A questionnaire to measure autonomy will embody a construct of autonomy, and it would be necessary to establish an overt, detailed, and firm basis for the construct for use in the selection of items for the Long List. When research into autonomy requires a definition the most common ways to proceed are to:

- Choose a definition, e.g. Ravindran (2000) chose Knowles' (1975) description
- Formulate a definition of autonomy based on some of the literature and one's own experiences, e.g. Cotterall (1995)
- Gather an expert group and agree together on a formulation, e.g. Dam et al. (1991)
- Adopt an institutionally or officially sanctioned definition
- Combine some of the above approaches

Each of these methods will result in a definition which is approved of by some, but will inevitably not be acceptable to all involved in the field. I did not wish to alienate readers by choosing a definition of autonomy at the very outset of the research. I felt it would be preferable to initiate a clear and overt procedure which would reassure the reader that I was open to empirical evidence in the formation of the construct. I would allow the construct to emerge from the research, but with clear parameters permitting a focus on the original aims of the research. I would not have a fully-

formed and final definition of autonomy as an initial assumption. (This attention to the area of making the research acceptable to as wide an audience as possible is also reflected in the choice of mixed methods.)

It would therefore be necessary to have a principled procedure for item selection and for validation of the selection. Item selection would be achieved by reviewing the literature of autonomy for language learning and establishing the main areas then covering the areas with the items. Validation would be with a robust process for checking the construct validity of the construct of autonomy embodied in the items. These two aspects are discussed in the following two sections.

3.9.1.1 The literature review

Choosing items will inevitably involve the researcher in thinking about the subject, but this should not be the only method employed. Gillham (2000: 17) warns that it is dangerous to rely only on your own experience, as it “may lead you into the assumption that you *know* what the issues are because you are familiar with that kind of context”. One way to find ideas for items is to review the literature (Aiken 1997: 33; Dörnyei 2003: 32).

The Literature Review (Section 2.4) has shown the areas which would be used as a basis for the selection of items. As discussed in the Literature Review it is not a comprehensive representation of all possible views of autonomy, because (a) it would be difficult to achieve for practical reasons of length, and (b) because it would not be necessary for the aims of the questionnaire. The aims, as previously stated, were to investigate a quick and simple closed-item questionnaire to measure autonomy for language learning rather than autonomy for life, or language learning

for autonomy (Benson & Voller 1997: 2). These specifications would define the parameters of the construct of autonomy embodied in the research.

Once the main areas have been identified, they can be used in generating an item pool (Dörnyei 2003: 32). Gillham (2000: 41) observes that “A typical weakness of the novice researcher is to try to include [...] too many topics”. In the present research it was considered that at the initial stages of instrument development it was necessary to have a long and inclusive list of topics so as to avoid as much as possible the drawback of prejudging the results.

The multidimensionality of autonomy would be incorporated by including different dimensions. The dimensions chosen are presented below; however, one feature of autonomy – variability – would not be selected as one of the criteria since it is a feature rather than a defining characteristic. The area of transferability is, for the reasons given in the Literature Review (see Section 2.4.2), an important aspect which is seen as making autonomy a worthwhile goal and so the Long List would include items which focused on different areas of language learning skill, for example there would be items which address listening, speaking, reading, writing, grammar, and vocabulary. As a result of the Literature Review the following areas were highlighted (see Section 2.4) as promising categories for the Long List:

- Control
- Skills
- Strategies
- Confidence
- Motivation
- Metacognition
- Social Interaction
- Attitudes to learning
- Actions/Behaviours
- Responsibility

With the areas defined it would be necessary to populate them with items which addressed them (presented in Section 4.2.3).

3.10 Questionnaire development

There is agreement that a questionnaire is not an ad hoc list of items, but has to go through a number of stages of development. For Cohen et al. (2000: 247) it is necessary to plan the development of a questionnaire to ensure that it:

- a) is clear on its purposes
- b) is clear on what needs to be included or covered in the questionnaire in order to meet the purposes
- c) is exhaustive in its coverage of the elements of inclusion
- d) asks the most appropriate *kinds* of questions
- e) elicits the most appropriate *kinds* of data to the research purposes
- f) asks for empirical data

There is a fairly close correspondence between authors on the stages of questionnaire design, though the exact contents and techniques recommended vary (Fowler 1993: 94; Dörnyei 2003: 16-17; Sudman & Bradburn 1982: 281-282). The stages are shown in Table 3.4 below, which also shows the sections which deal with them.

Stage	See Section
Deciding on questionnaire type	3.4.1
Researching the field	3.9, Literature Review Chapter
Writing items	4.2
Gathering feedback on the items	3.11.3, 4.3.3, 4.3.4, 5.1.1
Putting items into a draft questionnaire	3.10.1

Table 3.4: Questionnaire development stages

3.10.1 Putting items into a draft questionnaire

The items have been arranged in six different formats during the course of the research as shown in Table 3.5 below. At each of the first four stages feedback was gathered and changes were made, including removing items, which are detailed in the reports of data gatherings. The different formats were used for data gathering and contributed to the data used for factor analysis of the full data set. Only the 50 items selected for the questionnaire were used in the factor analysis at the end of this research.

Items	Delivery Medium	Language	Notes	Reference	See Section
Long List (i.e. all 256 items)	Paper	English/Arabic in parallel	<i>Did not have "don't know" response option. 5-point Likert scale</i>	<i>Format A</i>	<i>4.3.1</i>
	Internet	English	<i>7-point Likert scale</i>	<i>Format B</i>	<i>4.3.2</i>
	Internet	Mandarin Chinese	<i>7-point Likert scale</i>	<i>Format C</i>	<i>4.3.3</i>
	Emailed Word document.	Mandarin Chinese	<i>7-point Likert scale</i>	<i>Format D</i>	<i>4.3.4</i>
			<i>Items grouped into "face valid" sections</i>		
Short List (i.e. the 50 items selected from the Long List)	Internet	Mandarin Chinese	<i>7-point Likert scale</i>	<i>Format E</i>	<i>5.3</i> <i>5.6</i>
	Internet	English	<i>7-point Likert scale</i>	<i>Format F</i>	<i>5.4</i> <i>5.5</i>

Table 3.5: The formats of the instrument used in the research

3.10.1.1 Questionnaire length

Gillham (2000: 41) notes that it is common for “the novice researcher [...] to try to include [...] too many questions”. Dörnyei (2003: 18) recommends thinking of the

slowest reader in a sample as a guide to limiting the length of a questionnaire, and points out that one cannot include everything. Respondent fatigue is another reason to keep questionnaires short and Gillham feels that this is particularly relevant with rating scale questions (which is the type of response format chosen for this questionnaire). He says “It is extremely boring to answer a series of scaled-response questions; and people stop thinking about what they are doing” (Gillham 2000: 39).

Questionnaire length is usually expressed in terms of how long it takes to complete. The figure of thirty minutes is often mentioned, for example Dörnyei (2003: 18), Aiken (1997: 38), and Fowler (1993: 103). The initial list of items was rather long, but this was necessary in order to deal with all the areas. The latest version of the questionnaire (Formats E and F) has an average completion time of 16 minutes. This is well within the recommended figure.

3.10.1.2 Grouping and sequencing

Sudman & Bradburn (1982: 207) see the order of the questionnaire as requiring the same care as the wording of the items. Cohen et al. (2000: 258) consider that the configuration of the questionnaire indicates “the overall logic and coherence of the questionnaire to the respondents”. The order of the questions and the groups into which they are placed contribute to this impression and lead to better data (Sudman & Bradburn 1982: 207).

Aiken (1997: 38) recommends that the items should follow a “logical conversational sequence” and the questionnaire should not “jump around”. It is important not to produce a questionnaire that “seems to meander at random from question to question” as “Respondents who perceive the questionnaire as being done carefully

are more likely to be careful in the responses they give” (Sudman & Bradburn 1982: 228). Gillham (2000: 25) also recommends that questions should lead logically from one to the next for two reasons: it makes it easier for the respondent to work through, and it prevents “dotting around”, i.e. not answering the questions in sequence. This is important because questions are not “stand alone” says Gillham (2000: 25) and if they are not answered in the intended order the contexts of the questions will be different for each respondent. This context effect is defined by Tourangeau, Rips, & Rasinski (2000: 200) as “the effects of earlier questions on the responses to later ones”. This arises because questions can stimulate memories and with the memories activated, the responses to following questions can be influenced. The answer the respondent gives will be in the context of those activated memories (Converse & Presser 1986: 38). This is an effect which can make such a great difference that individual respondents can be, in effect “answering a different questionnaire” (Gillham 2000: 12). However, Tourangeau et al. (2000: 216) maintain that context effects are greater when the questions are perceived as being related and “When questions shift from one topic to the next without warning, respondents are no longer likely to see earlier questions as carrying implications about the meaning of later questions”. This suggests that “dotting around” may not be so undesirable. There appears to be some disagreement in the literature, about the nature of this effect and I feel that it is so difficult to account for, especially when it is not known whether the effect is present in the questionnaire or not, that it should be looked at only in the event of problems arising.

I put the open ended questions at the end of the questionnaire as they are more difficult to answer, take up more time (Dörnyei 2003: 48, 62), and are perceived as threatening if they occur at the beginning (Sudman & Bradburn 1982: 218, 262).

Demographic questions were also placed at the end in line with advice from Sudman & Bradburn (1982: 218) and Dörnyei (2003: 61). However, Converse & Presser (1986: 39) admit that “It is frequently unclear that one order is better than another” as “each order may reveal a different facet of the issue being studied”. As with question wording, the order of items in a questionnaire is clearly highly important, but it is frequently a matter of judgement of relative merits in a specific situation.

In the Microsoft Word-based version of the questionnaire (Format D) I grouped the questions by theme as recommended by Aiken (1997: 37), which was designed to “break up” the 256 item questionnaire and make it appear shorter. I sequenced these groupings so as to give a “sense of structuredness” (Dörnyei 2003: 21) and these sections were given headings to orientate the respondents to the focus and help to make them feel more involved (Cohen et al. 2000: 258-259).

3.10.1.3 Questionnaire formatting

“The appearance of the questionnaire is vitally important” (Cohen, et al. 2000: 258); the respondents’ perception of the difficulty of the task can be affected by its appearance (Sudman & Bradburn 1982: 243). Consequently “It must look easy, attractive and interesting rather than complicated, unclear, forbidding and boring” (Cohen, et al. 2000: 258). This can be helped by aiming for a clean uncluttered look (Gillham 2000: 39; Dörnyei 2003: 19; Cohen et al. 2000: 258). It is a common mistake to attempt to make a questionnaire look shorter by crowding questions together and using a small type face (Sudman & Bradburn 1982: 244). All this advice was kept in mind for all versions of the questionnaire, and following a suggestion by Dörnyei (2003: 21), different colours were used for the different sections in the Microsoft Word-based version and, rather than numbering the questions from 1-256,

they were numbered within the sections to help make it appear shorter (Cohen et al. 2000: 258-259). Format F (the latest version) has only 50 items and it was not considered necessary to use formatting to make it appear less forbidding.

3.10.2 The parts of a questionnaire

Apart from the items themselves, which make up its bulk, Dörnyei (2003: 25-30) gives the component parts of a questionnaire as: Title; Introduction and general instructions; Specific instructions; Additional information questions; and Final thanks. A questionnaire should speak for itself (Gillham 2000: 38) and this is particularly so in the case of one designed for self-administration such as mine. Respondents are not trained and are unlikely to be motivated (Fowler 1993: 100), so attention has to be given to introducing them to the questionnaire, giving them clear instructions, and making them feel positive about doing the questionnaire.

3.10.2.1 Title

Dörnyei (2003: 25) gives the title's functions as: identifying the domain of investigation, providing the respondent with initial orientation, and activating content schemata.

Aiken (1997: 40) specifically recommends that words such as "questionnaire" should not be included in the title. He does not say why, but Dörnyei (2003: 25) thinks it is because such words are uninformative. I decided not to put the word "autonomy" in the title of the questionnaire because it is a technical term and it was sufficient that respondents understood that it was a questionnaire about how they learned English.

3.10.2.2 Introduction and general instructions

There are some basic areas to include in the introduction and general instructions section at the beginning of a questionnaire. Aiken (1997: 37) recommends saying how long it will take to complete, and what to do with it when finished. Dörnyei (2003: 26) advises naming the organization responsible for conducting the study, promising confidentiality, and saying thank you. It is also desirable to reassure respondents that there are no right or wrong answers in order to encourage honesty. To this should be added the assurance that the results will not go towards grades if the respondents are students. It is generally agreed that a statement should be included about the purpose of the questionnaire (e.g., Aiken 1997: 37; Dörnyei 2003: 26). However, in my questionnaire I did not mention the aim of researching autonomy partly because it is a technical term, and I also felt that respondents might be tempted to show themselves to be more (or less) autonomous, which would introduce an unpredictable bias.

3.10.2.3 Specific instructions

It is important that respondents know exactly how to indicate their answers to the different types of question contained in the questionnaire (Cohen et al. 2000: 258; Aiken 1997: 37) and instructions must be clear (Dörnyei 2003: 27). Oppenheim (1992: 142) warns that “Serious loss of data can result from ambiguous or inadequate answering instructions”. In addition, Dörnyei (2003: 27) stresses that instructions must be distinct from the questions to avoid confusion. The present questionnaire has the advantage of using only one question type for the closed questions, and for this reason it was not necessary to put instructions more than once.

3.10.2.4 Ending the questionnaire

A space at the end inviting optional comments and suggestions on the questionnaire and its content is recommended (Sudman & Bradburn 1982:218; Bourque & Fielder 1995: 104). I did this for all versions of the questionnaire apart from the first where respondents were encouraged to give their feedback verbally to the teacher.

Questions for demographic information should come at the end of the questionnaire (Oppenheim 1992: 132; Aiken 1997: 38; Sudman & Bradburn 1982: 218-219).

Aiken points out that placed at the beginning they would appear to be inconsistent with the questionnaire's title and stated purpose, which will not mention collecting demographic information (1997: 38). Oppenheim's (1992: 132) reasoning is that by the time the respondents have completed the questions they are more likely to be convinced that the inquiry is a genuine one and so be more willing to give personal information. In Format A of the Long List no demographic questions were asked, but these were included in subsequent versions placed at the end of the questionnaire.

3.11 Administering the questionnaire

In this section I will explain my choices regarding how I would administer the instrument for data gathering purposes. I will consider whether to target groups or individuals, the medium to be used (paper, online), and how I would gather feedback on the items and questionnaire.

3.11.1 Groups or individuals

Group administration is where the questionnaire is administered to "groups of respondents assembled together" (Oppenheim 1992: 103). A "surveyor or other supervisory person" (Bourque & Fielder 1995: 4) is present to give instructions and

monitor the respondents, and the precise way that the administration is handled should be decided in advance, and the administrator should have clear instructions (ibid 1995: 5). Ideally the researcher would always be present in person to ensure that procedures and conditions are comparable with other administrations. With a stand-alone questionnaire such as mine (see Section 3.4.1) this problem is reduced as it is less necessary for the researcher to be present and there are fewer of the uncertainties resulting from how the class teacher handles the administration.

As resources were limited and with a large number of responses necessary, the most suitable modes of administration were self- and group-administered. The questionnaire is structured and is therefore well-suited to both self-administration and group-administration modes. A structured questionnaire can be delivered electronically, or administered in paper-based form, and both of these were used.

3.11.2 Medium

There is a choice of media for questionnaires, online or paper-based, and I would aim to have online data gathering as much as possible as I have had experience with both of these modes in the past, and have come to prefer the digital option over the paper-based. The disadvantages of paper-based administration are, firstly, that it takes a disproportionate amount of time to transfer the answers from the sheets to the computer for analysis, and there was also the strong possibility of making transcribing errors. Secondly, the completed questionnaire sheets for overseas administration have to be transported physically to and from the United Kingdom. The price of postage for this can be quite high, and it is a problem to transfer funds to pay for the return postage. Thirdly, respondents sometimes do not answer an item, resulting in incomplete data sets.

The use of computer technology rather than paper helps to avoid mistakes in transcribing, as data can be entered into software such as Excel and SPSS with a few mouse clicks without ever having to physically type the numbers into the computer. In addition electronic formats make the questionnaire more quickly and easily distributable. A web-based questionnaire can also be programmed to check automatically for unanswered questions and prompt respondents to complete them, as was done with the present questionnaire. An Internet-based questionnaire does take time to set up, but this was considered worthwhile because of the advantages already mentioned.

3.11.3 Gathering feedback

I gathered feedback on both the items and the questionnaire to, as Cohen et al. (2000: 260) recommend, check clarity, eliminate ambiguities, see how the questions perform, get feedback on the appearance and layout, see how long it takes, see if respondents remain interested, and to check its appropriateness. I gathered feedback on the items in a number of ways and at different stages of the research. An initial stage was to obtain feedback from my supervisors. At this stage, no decision to translate the questionnaire had been made. Changes were made to the items based on the feedback received, and changes were also made to the format of the questionnaire, for example I decided to use words to describe the Likert options rather than pictures or symbols. For the class-based data gatherings I asked the teachers to gather feedback from their students about the questions. For the online data gatherings I made provision for feedback on the items at the end of the questionnaire.

Another angle of feedback was the translation process which proved to be an excellent way of improving the clarity of the items (see Section 3.13). I found it particularly useful in highlighting how items can be misunderstood by non-native-speaker respondents.

I gathered feedback on the items and the instrument at the following chronological stages:

1. Discussion with supervisors
2. A paper-based administration in the UAE, during which feedback on items was gathered (see Section 4.3.1)
3. Online administration of the English language items (see Section 4.3.2)
4. Translation into Mandarin Chinese (see Section 3.13)
5. Online administration of the Mandarin Chinese items (see Section 4.3.3)
6. Email-based administration of the Mandarin Chinese items in a Microsoft Word form at Beijing Normal University (BNU) (see Section 4.3.4)
7. Web-based administration of 50-item Mandarin Chinese questionnaire with ELTCS (English Language Translation and Cultural Studies) BA students at Warwick University (see Sections 5.3 and 5.4)

3.12 Data analysis

In this section I will cover issues of data analysis such as coding, standardisation of data from different data gatherings, and factor analysis of the data for checking construct validity.

3.12.1 Standardisation

I standardised all data before analysis using SPSS, which is essential if data from different data gatherings are to be combined. For the purposes of factor analysis, large amounts of data are required, specifically the ratio of the number of respondents to the number of items in the questionnaire must be high (see Section 3.12.3 below). Without standardisation it would not be possible to combine the data from different data gatherings. Each data gathering has its own unique context, for example in the present research the format of the questionnaire, the medium, the context of an item within the list of items, the place (home, class etc.), the length of time needed for completion, etc. would all vary. Therefore in order for these data to be compatible and to be pooled for factor analysis an accepted procedure for standardisation would have to be followed. Dörnyei (2007: 205) explains that “The standardization of raw scores involves the conversion of the distribution within a sample in a way that the mean will be 0 and the standard deviation 1”. According to Dörnyei the standard scores express each raw value in terms of its difference from the group mean. The means from the different samples are equalised making the scores comparable without affecting the resulting coefficients in correlation-based analyses such as factor analysis (Dörnyei 2007: 205).

The data were combined for the factor analysis necessary at the construct validation stage. I had intended to use one format of the questionnaire for this, but in the event different formats were combined. I could combine the data because I would not be comparing treatments – sample groups were in effect instances of a universal sample of learners from many backgrounds that would be built up. Most importantly, the 50 items I used were common to all of the data gatherings.

3.12.2 Coding of responses

Responses to the questionnaire items would be coded so that I could process them with statistics software (SPSS or Excel). The coding would be subjective to an extent, as I would need to decide (based on the literature) which items indicated high or low capacity for autonomy. The factor analysis would also tend to counteract this subjectivity since it would group items by their statistical interaction, not by high or low codings.

Each of the 256 items making up the Long List would be a statement. For purposes of data gathering each item would be given a Likert scale for participants to make their responses. Initially these were 5-point (“Format A”), worded “Strongly agree”, “Agree”, “Neither agree nor disagree”, “Disagree”, and “Strongly disagree”. The responses were coded in sequence from 5 “Strongly agree”, to 1 “Strongly disagree”.

For the purposes of data analysis it would be necessary to reverse code those items which were worded negatively (or had a negative connotation as regards autonomy). 73 of the 256 items were reverse coded. The inclusion of negative items is advisable to discourage respondents from marking only one side of the rating scale and also to counter the effects of acquiescence bias (Dörnyei 2007: 205). In this thesis items which are reverse coded are indicated (where appropriate) with an “R”. Normally-coded and reverse-coded items all had “Neither agree nor disagree” = 3.

After the UAE data gathering the 5-point Likert scale was widened for all subsequent formats to increase the definition of the responses. Table 3.6 below shows this and the corresponding changes in the coding of responses. The description of the Formats A to E is given in Table 3.5.

Format A	Coding	Formats B, C, D, E, F	Coding
		Very strongly agree	7
Strongly agree	5	Strongly agree	6
Agree	4	Agree	5
Neither agree nor disagree	3	Neither agree nor disagree	4
Disagree	2	Disagree	3
Strongly disagree	1	Strongly disagree	2
		Very strongly disagree	1
		Don't know	0

Table 3.6: Comparison of Likert item responses

3.12.3 Factor analysis

When all the data gatherings had been completed the data would be standardised (see Section 3.12.1 above) to form the complete data set for the 50 items to be factor analysed. The purpose of this would be to check the construct validity of the questionnaire (as discussed in Section 3.8.2).

Though I would have carefully reviewed the literature to find areas for inclusion in the Long List, no one item could be expected to isolate one area of autonomy, and the area which it was intended to represent may not be how it would be perceived by the respondents and so it may be placed by factor analysis into an unexpected factor. Exploratory factor analysis may put two items together in a factor which come from different areas, but in doing so reveal a hitherto unforeseen area of autonomy, or put an unexpectedly stronger emphasis on one aspect rather than another. The strength of the correlations between factor groupings suggests how areas of the construct may be linked together. This is potentially of great value to learners, teachers, and possibly researchers (see Section 7.3.4).

3.12.3.1 What is factor analysis?

Factor analysis is “a statistical technique based on analysis of correlation coefficients” (Nakatani 2006: 153). It groups items (variables) “into a number of supervariables (‘Factors’) on the basis of their inter-correlation” (Jones 1998: 386).

Since factor analysis groups items together into what are, statistically speaking, similarly behaving sets (factors) of items, it is a procedure which can be used, for example, to show how questionnaire items may be interrelated. As Jones (1998: 386-388) explains the “Relative strength of the Factors is shown by the percentage of the total data-set variance which each one accounts for” which means that it is possible to identify which of the factors found in the analysis account for responses to a questionnaire (Schommer 1990: 499). Factor analysis can be used for data reduction as it can “reduce a large number of variables to a small number of values that will still represent the information found in the original variables” (Nakatani 2006: 153).

Factor analysis, then, produces groupings of items which are behaving similarly in a statistical way. It also shows how the items are more strongly or weakly associated with a particular factor, for example some items will be “cross loading”, i.e. they are strongly associated with more than one factor. The results of a factor analysis are displayed by the statistical analysis software in the form of a matrix of figures formed by two axes, the items (or variables) on one, and the factors on the other. As Jones explains, “The figures in the matrix show how well each raw variable correlates with the Factor as a whole in other words, its relative contribution to the Factor” (Jones 1998: 386-388).

3.12.3.2 Why use factor analysis?

Firstly, factor analysis is a powerful tool for finding the concepts which can be said to underlie a set of items (Regan 1994; Tremblay 2001) such as those found in a questionnaire.

Secondly it can be used for data reduction. The procedure can be used to form multi-item scales, which can then be statistically checked for internal consistency and therefore indicate items which are suspect and can be rejected. This was the original intention in the present research as it would allow the Long List to be shortened while still retaining its necessary coverage of the significant elements of autonomy.

Thirdly, as Green & Oxford (1995) note, it can be used a means of supplying evidence for the construct validity of an instrument. For example, a factor analysis of the data obtained from a questionnaire may provide a “picture” which is “consistent with theory” (Miller et al. 2002: 184) and this means that it is more likely to be a valid instrument than one which provides an incoherent picture.

There are two basic types of factor analysis, confirmatory and exploratory, based on the use to which it is to be put.

3.12.3.2.1 Confirmatory factor analysis

Confirmatory factor analysis (CFA) is where the technique is used to confirm groupings which have already been established by other means, such as expert choice, introspection, observation, or theory. For example, Tseng, Dörnyei & Schmitt (2006: 90) adopted this in their research into self-regulation in vocabulary acquisition as they “assessed the hypothesized model for its fit to the observed data”.

CFA is in effect a way of testing a hypothesis:

Confirmatory factor analysis (CFA) is designed to assess how well a hypothesized factor structure “fits” the observed data. Unlike exploratory factor analysis, the researcher has an explicit prediction concerning both the number of factors that underlie a set of measures and which measures load on the hypothesized factor(s). (Russell 2002: 1638)

In the present research factor analysis would be used not to confirm a prediction or hypothesis as characterised here, but to reveal a construct to assess its validity. It is thus more akin to exploratory factor analysis, discussed next.

3.12.3.2 Exploratory factor analysis

Exploratory factor analysis (EFA), in contrast, is where the technique is used to discover “the patterns that underlie the correlations between a number of variables” (Miller et al. 2002: 174). Researchers have made much use of EFA “to generate hypotheses by identifying characteristics that test items have in common, which do not exist on the surface of the observed data” (Nakatani 2006 153).

EFA produces “a set of empirically-derived categories” (Sakui & Gaies 1999: 475) which are valuable because they provide an alternative source of ideas about what the components of autonomy are and how they may be related (see Section 7.3.4 for discussion and a diagrammatic representation of this for the present research).

Naturally it does not provide a new model for autonomy as it is dependent on many contingent influences, including the coverage of the selection of items used. It is important to cover as much of the relevant field as possible in the questionnaire so that the coverage of the set of items is comprehensive (Dörnyei 2007: 234) and it can thus measure all relevant areas. It is also important for maintaining the construct validity of the picture of autonomy embodied in the items.

EFA can also be used for the purpose of data reduction, for example to eliminate items from a questionnaire which are only weakly correlated with the factors (Russell 2002: 1636). The technique of factor analysis gives a figure for the loading of individual items into factor groups and these figures can be used in selection, so

items which load strongly are candidates for selection and items which load weakly or on to more than one group are candidates for rejection. As previously stated one aim of factor analysis was originally data reduction. This was not possible due to the low ratio of respondents to items. This was because the Long List was too long which meant that there were not enough willing volunteers to complete it.

Consequently an adequate ratio of respondents to items for carrying out factor analysis was not reached. Factor analysis would have to wait until the Short List of 50 items had been selected from the original 256 of the Long List by other statistical means (see Section 5.1). A drawback of this could have been that it reduced the coverage of the field of autonomy before the factor analysis had been carried out. It was therefore important that the Short List would continue to represent the range of areas covered by the Long List and thus maintain its links with the literature on autonomy which had been originally reviewed. This was achieved, though the proportions (i.e. the ratio of items in one area to the total number of items in the Lists) were not identical (see Table 5.3) but they were very similar. The precise ratio was not critical, though it was vital that at least three items were present in each area as this is the minimum for a factor to be recognised (see Section 3.12.3.3.5 below). This and other important considerations to be aware of when carrying out and interpreting factor analysis are discussed in the following section.

3.12.3.3 Key considerations in carrying out factor analysis

EFA is a complex procedure with few absolute guidelines and many options (Costello & Osborne 2005: 1) which could make the process of deciding which to use relatively subjective; and it is tempting to try as many as possible and then,

retrospectively, justify the choice which fits with one's expected outcomes. I avoided this by using a number of FA methods and extracting the commonalities.

3.12.3.3.1 Sample size

To do a statistical analysis such as exploratory factor analysis it is necessary to work with data from a sample of adequate size. Field (2005: 638-640) reviews the literature on sample size for factor analysis, which offers a number of conflicting indications. He reports (2005: 638) that there is a common rule of thumb which suggests a researcher should have at least 10 to 15 participants per variable. Russell (2002: 1632) however, has found that "Minimums of 5 or 10 cases per measure have typically been recommended".

Field's literature review also suggested that 300 cases is a good sample size. He reports (Field: 2005: 640) that Guadagnoli & Velicer (1988) argue that if a factor has four or more loadings greater than 0.6 then it is reliable regardless of sample size. He reports that MacCallum et al. (1999) find that relatively small samples of less than 100 may be perfectly adequate if all the factor loadings are above 0.6. It was therefore decided to aim for 300 participants for the questionnaire research, but additionally, to process the data and inspect the factor loadings to see if it would be possible to stop before reaching 300. Another alternative is to see whether enough data have been gathered using the Kaiser-Meyer-Olkin measure of sampling adequacy (KMO). If SPSS indicates that the value is 0.5 or above this is acceptable, and becomes more acceptable if it is nearer to 1.0. If it is below 0.5 more data should be collected (Field: 2005: 648-650). It is also, states Field, (2005: 650) important to examine the individual KMO statistics for all variables. These should likewise be above 0.5. If any are not the analysis should be run without them and the results

compared. The off-diagonal data in the anti-image matrix should be very small. In addition Field (2005: 652) states that Bartlett's Test of Sphericity should have significance at lower than 0.05 in factor analysis.

Field (648) says that “to do a factor analysis we need to have variables that correlate fairly well, but not perfectly” and, “Any variables that correlate with no others should be eliminated.” The significance values of any variable should be mostly less than 0.05. The correlation coefficients themselves should be below 0.9. If necessary one of the two variables causing the problem should be eliminated. This is a way of potentially reducing the number of items in the questionnaire, as well as improving its accuracy. He also recommends (2005: 648) checking that the determinant is bigger than 0.00001. According to Miller et al. (2002: 184) the criterion of successful factor analysis is loadings which allow an interpretation that is “consistent with theory, or just makes common sense”.

3.12.3.3.2 Number of factor groupings

The first issue faced by any investigator planning an exploratory factor analysis concerns how to extract factors from the data (Russell 2002: 1630) because there is no single best procedure which has been established for factor analysis; for example, ascertaining the correct number of factor groupings to extract from the data is an area which is much debated (Dörnyei 2007; Costello & Osborne 2005). I chose to base my decision on the technique of using indications from the scree plot of eigenvalues. The break of slope in the scree plot indicates the number of factors to extract (Dörnyei 2007; Field 2005; Costello & Osborne 2005).

Dörnyei recommends (2007: 233-236) using either maximum likelihood or principal component analysis as the extraction method, and either oblimin or varimax as the rotation method, giving four combinations for the factor analysis procedure and three alternatives for the number of factors to extract, and this is the procedure which I used. This resulted in 12 separate factor analyses which I then checked for “low-loading” and “cross-loading” items, and for groupings with insufficient items.

3.12.3.3.3 *Criteria for accepting items in factors*

Prior to using the factors to identify components of the construct it is necessary to check that the items are strongly associated with the factor.

Low loading: If an item does not load on to a factor above a certain level it can be ignored. The loading of an item onto a group can be strong or weak, ranging from the maximum 1.000 to 0. The loadings should be below 0.9 (Field 2005: 648) and at the lower level the figure is debatable and depends on the strength of one’s data. For example Nakatani (2006: 154) used the figure of 0.4 or less, while Yang (1999: 520) chose 0.3. I follow Dörnyei (2007: 235) with a figure of 0.3. Items loading below this on all factors were eliminated from further analysis.

Cross loading: Often a single item will load onto two different groups at the significant level or above. This “cross loading” indicates that an item is shared between two or more factors. This would be a criterion for rejecting an item from my analysis as it would add ambivalence to a grouping. Where an item loads on to more than one factor above the 0.3 level it would be ignored. If this cross loading item were retained it would confuse the identification of the factor groupings (see Section 3.12.3.3.4 below).

Cronbach's alpha: Individual items in factors can detract from the overall reliability of the grouping, and the Cronbach's alpha for reliability of scales can be used to find such items. There are two criteria for accepting or rejecting items based on the Cronbach's alpha: firstly, the corrected item-total correlation figure should be above .3 and certainly above .2 (some leeway can be allowed for if an item is otherwise strong); secondly, if deleting an item from the scale improves the Cronbach's alpha for that scale then the item should be deleted (Field 2005: 666-675).

3.12.3.3.4 Identification and interpretation of factor groupings

For a factor grouping to exist it must have three items or more which load on to it at the significant level. In the case of a grouping with only two items it will not be possible to identify reliably the underlying common feature and for this reason such groupings would be removed from the analysis.

Gan et al. (2004: 407) explains how he identified factors, "The name or label for a factor was based on a unifying concept (or unifying concepts) embedded within a pool of items". The factor groupings produced by EFA need to be interpreted by the researcher (this has not been pre-defined as in CFA) and it is beyond the capabilities of software to identify any common underlying feature. There is a danger of misidentification, perhaps caused by unconscious preconceptions. In the present research these groupings were given names such as "Information Literacy" or "Metacognition". Confirmation of this identification should be sought from suitably experienced independent parties. In the present research this would be done in the supervision process.

The process of identification is seen as a subjective process (Field 2005: 666; Dörnyei 2007:236) and detailed procedures for carrying it out seem to be absent from the literature. To find the underlying connection in a principled way a procedure was decided on and followed. The procedure was:

1. Remember not to have a positive or negative judgement of an item
2. Find the key words in each of the items in the scale and highlight
3. Write short phrases synthesising the key words
4. Look at each item and note down its prerequisites or requirements
5. Think of words or short descriptions which represent the most key words, phrases, and prerequisites in a scale to produce a description of the underlying connection for a scale
6. Compare the descriptions for all the scales to ensure they are distinct.

It is clear that this is not an objective method. It does, however, have the advantage of being a procedure, and is therefore to some extent a repeatable process. I am not aware of any other researcher in the autonomy field who has specified an overt procedure for labelling factors. It has the advantage of being a set of guidelines which can easily be referred to. Point 1, for instance, addresses the fact that it is easy to treat an item as reflecting a good or bad feature of the construct under investigation. The procedure was carried out and resulted in themes or areas that seemed to connect the items within a scale.

3.12.3.3.5 Number of items in a factor

Russell (2002: 1632) quotes a figure of four or more items per factor for ensuring identification, but recommends a figure of at least three. In the present research

groupings produced by factor analysis would be treated as forming provisional multi-item scales. The Cronbach's alpha measure of internal consistency reliability for these scales would be calculated and these figures used to check the composition of scales. According to Dörnyei (2003:112) short scales of 3 or 4 items should aim for a reliability of 0.70, and should at least reach 0.60. However, Field (2005: 668) points out that the number of items in the scale is a major influence and longer scales will always tend to have better Cronbach's alphas. He cites a particular case where a figure of 0.57 is "respectable" because there were only three items in the scale.

3.13 Translation

3.13.1 Introduction

Very often interesting things come out of research which were not foreseen at the planning stage. This was the case with my research as issues (e.g. feedback to items 37 and 105) emerged regarding the Arabic translation of the Long List for the UAE administration (see Section 4.3.1). This raised the question of how to translate, and whether to present parallel English and translated items (see Section 3.13.4.6). The importance of all respondents having the same understanding of the items is clear and is consequently discussed at length here. Due to the preponderance of Chinese respondents available for my research Mandarin Chinese was chosen as the first language for the new more considered translation process (further languages were planned should the present research provide encouraging results).

In this section I will first consider the reasons for translating the questionnaire, then look at the fundamental concept of equivalence. Following this I consider the possible ways of conducting my translation and finally I explain and describe the procedure I adopted.

3.13.2 Reasons for translating

Translation is not seen as a central concern in language teaching, and in class it is very uncommon to use questionnaires or tests in translation. However, I think there are good reasons for conducting English language teaching/learning research in the learners' first language. My reasons for translating the questionnaire are:

- respondents will be better able to understand the questionnaire in their own language, and therefore the data obtained are more likely to be reliable
- respondents are more likely to be able to answer all the questions
- they can finish the questionnaire more quickly (particularly important in a long questionnaire)
- lower cognitive load for respondents
- they are more likely to complete the questionnaire
- it is less of an imposition to ask respondents to do a questionnaire in their own language
- respondents' levels of English will differ, introducing unpredictable variations in the data
- a careful translation will take account of the respondents' possible context and not include culturally unfamiliar references.

3.13.3 Translation equivalence

There is a lot of agreement in the instrument adaptation literature that literal translations are not usually appropriate for questionnaires. Harkness & Schoua-Glusberg (1998: 93), for instance, point out that "Languages are not isomorphic and so translation cannot be expected to operate on a one-to-one basis across languages".

Rather, the notion of equivalence is seen as key, an idea which has its roots in the literature of translation studies. Nida (1964: 129), for example, says “equivalence aims at complete naturalness of expression, and tries to relate the receptor to codes of behaviour relevant within the context of his own culture”. The relevance of this to the autonomy measuring instrument is that where items refer to situations which are peculiar to a Western environment they will need to be changed to equivalent situations which are familiar to a Chinese audience. Harkness & Schoua-Glusberg would term this “covert translation”, that is, the translation process “produces a target language text which reads like an original text [...] in the target language and thus does not signal that it is a translation” (1998: 104). Newmark (1981 1991) makes a distinction between what he terms “Semantic” and “Communicative” translation. A semantic translation will be one which aims to remain close to the structure of the original, while a communicative translation will have equivalence in that it “attempts to produce on its readers an effect as close as possible to that obtained on the readers of the original” (Newmark 1981: 39).

My Chinese translation should have communicative equivalence to the original. However, when back translation (see Section 3.13.4.2) was carried out on the initial Chinese version of the questionnaire (see Appendix 10.1), it raised doubts about the communicative nature of the translation. It was not clear whether the problem lay in the initial forward translation or in the back translation. I therefore consulted a Chinese native speaker (a fellow student) who was fluent in English and who confirmed that a semantic style translation had been carried out. It was therefore necessary to do the translation again with much closer liaison with the translators. While communicative equivalence is desirable there is no watertight procedure for

achieving it; as Munday (2001: 43) points out “equivalence inevitably entails subjective judgement”.

3.13.3.1 Sources of bias in translations

In the translation process I would need to include checks for “ethnocentric bias” i.e. assumptions about how a construct is thought of or manifested in the target culture (Van de Vijver & Hambleton 1996: 90), in this case Chinese. Smith (2004) recommends the use of more than one item to probe a point; the same question can be asked using different wordings to make sure that the concept is equivalent between the languages: “Three linguistically distinct measures of the same construct are desirable” maintains Smith, and “if all three agree, one has a clear, robust finding” (2004: 434). Compare this with the discussion (Section 4.2.1.8) on multi-item scales, where more than one item probes the same area.

Staying close to the original text may appear to promote equivalence, for example, by using the same terms in source and target versions. However, this is not advisable when the term is culture-specific such as, for example, the word “Parliament”. It would probably be better to change this to an equivalent term in the target language (Van de Vijver & Hambleton 1996: 91).

These points indicated that I should use procedures involving close liaison between myself and the translators to minimise translation problems; the choice of translation procedure is the subject of the following section.

3.13.4 The choice of translation procedure

There is little hard evidence to guide in the selection of translation procedure (Smith 2004: 446). However, in this section I will give my rationale for the selection of

translation method from among the available options, and describe the process that I chose. The methods I considered, which are the most common methods used in translation, were: One-for-one translation; Back translation; Committee translation; Parallel translation; Advance translation; Bilingual instruments; and Decentering.

3.13.4.1 One-for-one translation

Harkness & Schoua-Glusberg (1998: 100) describe this as “one translator producing one translation in a traditional manner”. This was how the Arabic (see Section 4.3.1) and the first (problematic) Chinese translations were done. Smith (2004: 447) says that this approach is frequently used, being “quick, easy, and inexpensive”, but he adds that it is not usually recommended because it relies too much on one person’s perceptions and skills.

3.13.4.2 Back translation

Harkness & Schoua-Glusberg (1998: 97) define back translation as “the *translation of a translation* back into the source language”. The purpose for doing this is “to compare/contrast the back translation with the source text, usually with a view to assessing the quality of a translation” (1998: 111). The forward and back translations are carried out by separate translators. It is useful for detecting errors in translation, with the added advantage that the researcher does not have to have a command of the target language (Blais & Gidengil 1993: 543).

I used back translation, as previously stated, on the initial Chinese translation of the questionnaire and it was useful in highlighting difficulties with it, showing that the translation would have to be done again with a more reliable procedure.

3.13.4.3 Committee translation

This is where a team of translators and researchers discuss the meaning of the source items and assess possible translations. This method places emphasis on writing good items “not just on translating words” (Smith 2004: 448). It has the advantage of the combined expertise of the committee members, for example some may have knowledge of the target culture and others may be familiar with the construct under investigation (Harkness & Schoua-Glusberg 1998: 101). The committee members should “make independent translations of the same questionnaire” then “compare the translations, reconcile discrepancies, and agree on a final version” (Harkness & Schoua-Glusberg 1998: 101). I thought this a promising technique, but its major drawback was the difficulty of assembling a committee which was able to meet on sufficient occasions and dedicate the necessary time on a voluntary basis.

3.13.4.4 Parallel translation

This is where translators work independently then come together to compare results. When the results differ the translators work with the developers of the questionnaire to find out why (Smith 2004: 448). It did prove possible to have different translators working independently, however it was not possible to bring the translators together to compare their work and discuss it with me. However, parallel translation did contribute to the eventual method which I followed.

3.13.4.5 Advance translation

Harkness & Schoua-Glusberg support the idea that translation should begin during the drafting of the questionnaire, and not be left until after the items have been finalised. Although they do not cite research that provides empirical evidence for

this, they are of the opinion that this is an important procedure which “is often particularly relevant for the languages and cultures furthest removed from the models underlying the source text” (Harkness & Schoua-Glusberg 1998: 105). This makes it particularly relevant for the present Chinese translation, but not as a technique for translation, rather it is the recognition that it should be done at an early opportunity which is important.

3.13.4.6 Bilingual instruments

Making both the source and translated texts available to respondents appears to be a solution to the question of translation equivalence in that the reader, if sufficiently bilingual, can check both versions to work out the correct interpretation. Stansfield reports that “there is a growing belief that it ‘does no harm’” (2003: 201). However, he does not mention the possibility that test takers may be confused by possibly conflicting versions where the translation and the source have different connotations or interpretations. The data gathering in the UAE (see Section 4.3.1) received feedback from one student who was confused because there was a slight difference between the English and the Arabic translation. I feel that if the translation is perfect then it is not necessary to have a bilingual questionnaire, but if it is not perfect and there are differences between the two language versions, it seems to me that some doubt or confusion may result. In fact when discussing this point with a Chinese translator she recommended that bilingual text be avoided because it would be confusing. On this basis I decided to aim for a good monolingual Chinese translation.

3.13.4.7 Decentering

Decentering is a form of advance translation (3.13.4.5). Decentred translation, says Johnson (1998: 18-19), “may involve multiple iterations of translation and back-

translation, with each language version being continually refined to bring them into closer concordance of meaning”. There are various procedures for doing this. In one a draft questionnaire in the source language is used to produce two final questionnaires, one in the source and one in the target language. It is seen as a way of producing a text that is symmetrical, as opposed to being “centred” or “anchored” in a specific culture and language (Harkness & Schoua-Glusberg 1998: 98). Van de Vijver & Leung (1997: 39) describe decentering as “the removal of words and concepts in a source language that are difficult to translate or specific to a culture” and in this way it seeks to remove item bias and increase the equivalence.

I chose decentering for my translation procedure. It involved substantial discussion about individual items and if necessary the changing to alternative phrasing of the original text so that they worked in both languages as equivalently as possible. One translator would make a translation and then another translator would check it and add an alternative Chinese translation if they thought it necessary. I then worked with a translator and discussed which the best translation was, asking how it translated back into English, and if it would be easier to translate into natural sounding Chinese if the English text was changed. The resulting translation was then checked by independent Chinese L1 speakers who also made suggestions which I sent to other members of the translating team for their comments.

There were some items where I went against the advice of the translators because I felt my instincts were possibly right and I needed to await feedback from the administrations to see how the items performed. For example regarding item 251 “I use real English texts (i.e. not made for students) in my learning”, different translators thought I should use the term “authentic English texts”. However, I was

not convinced that this would be a term in everyday speech. The translators I was using were all in the English language teaching field which meant that this term was familiar to them, but I did not think it would be familiar to the Chinese respondents, who would be learners of English but not teachers.

3.13.5 Cultural differences in questionnaire response

In addition to translation, Smith (2004: 432) notes that other features of questionnaires may be subject to cultural differences. In this section I will discuss the techniques I considered to minimise such differences.

3.13.5.1 Nonverbal scales

In order to avoid using words and so reduce problems of translation and cross-cultural equivalence, the item response options can be expressed as numerical scales (e.g. 1-10, or percentages), ladders, thermometers, or smiley faces, etc. However, Smith (2004: 437-438) identifies some problems with this approach:

- Some scales are complicated
- Numerical scales are not invariant in meaning
- No research establishes whether numerical scales are used consistently across nations
- Different cultures have different lucky and unlucky numbers
- Numerical scales do not actually eliminate words, only reduce them

I therefore chose to use a conventionally-worded Likert scale with the present instrument (see Section 3.12.2).

3.13.5.2 Response styles

Acquiescence bias is the tendency for respondents to want to answer positively to questionnaire items; this makes the use of reverse coding a useful technique. This is where two questions cover the same construct but one is worded positively and the other has the statement worded negatively and reverse-coded (Tourangeau et al. 2000). I have included reverse-coded items in the questionnaire.

Some people will tend to favour the middle responses and others the extremes, regardless of their true attitude to an item (Smith 2004: 440). This is often given as a difference between East Asians and Westerners. However a study by Chen, Lee & Stevenson (1995) did not find evidence for East Asians preferring the midpoint, a result which is of importance to the present research. There is an issue as to whether to include the middle options in Likert scales. Smith says “Research from several countries finds that providing ambivalent respondents with a clear response option produces more reliable results” (Smith 2004: 441), but he does not specify which countries. However, Dörnyei (2001: 207) says “this [...] appears to be a relatively unimportant question that is not expected to modify the results significantly”. In the light of this, I chose to include a middle response to avoid forced choices (see Section 3.12.2)

After being translated and checked the administration of the Long List of 256 items continued, though much time had been lost due to the unforeseen need for a lengthy process to convert the list for use with Chinese respondents. However, I feel my research benefitted through working with the native Chinese speaking translators which helped me later to understand the ELTCS students involved in my research. I

will return to highlight the issues and implications of translation in language learning research in Sections 7.7 and 8.3.3.

3.14 Overview of eventual research stages

Though there was an original guiding plan it was soon modified. Table 3.7 provides the overview of the actual stages of research as carried out.

	Stages	See
1	Researching the field, literature review.	3.2.2
2	Writing items for the Long List of 256 items.	4.2
3	Data gathering in UAE. Parallel English/Arabic translation. 5-point Likert Scale. 256 items. (Format A)	4.3.1
4	Online data gathering of the English items. 7-point Likert Scale introduced (Format B)	4.3.2
5	Translation into Mandarin Chinese	3.13
6	Online data gathering with the Mandarin items (Format C)	4.3.3
7	Questionnaire designing	
8	Email data gathering with the Mandarin items in a Microsoft Word form at Beijing Normal University (BNU). (Format D)	4.3.4
9	Item selection process reducing number of items to 50	5.1
10	Web-based data gathering with 50-item Mandarin questionnaire (Format E) with 3 rd year ELTCS BA students at Warwick University (“T1”)	5.2, 5.3
11	Interview with one 3 rd year ELTCS student	5.2
12	Interview with 3 rd year ELTCS teacher. Teacher’s estimates of students’ autonomy levels are made	5.2
13	Preessional course with ELTCS group starts at Warwick University. I am their teacher. I make initial estimates of their autonomy. (“P1”)	5.2, 5.4
14	First Web-based data gathering with 50-item Mandarin questionnaire with my preessional ELTCS BA students. (Format E)	5.2
15	I teach preessional course with ELTCS students. I observe and make notes regarding students’ autonomy.	5.2
16	I make post-course estimates of ELTCS students’ autonomy.	5.2
17	Second Web-based data gathering with my preessional ELTCS group. (“P2”) (Format E)	5.2
18	Data was obtained from students in other preessional classes at Warwick University, Internet based, in English. (Format F)	5.5
19	Data obtained from 14 students in Taiwan. (Format E)	5.6
20	Standardisation of all data. Factor analysis of combined standardised data. Analysis of results of factor analysis and comparison with literature. Comparison of quantitative and qualitative data.	3.12

Format (A-F) refers to the presentation format of the items – see Table 3.5 for details

Table 3.7: Overview of the actual stages of research as carried out

Table 3.8 (page 153) shows the original outline plan for the research as envisaged at the outset and next to this are notes of problems and issues which emerged which caused changes to the plan.

3.15 Summary

In this chapter I have presented and explained the aims and rationale for the research and presented the plan which was based on those considerations. I have also, for reasons of clarity, given an overview of how plans were changed.

In the following two chapters, each covering a distinct period in the research, I describe in chronological order how the research actually progressed. In the first of them I describe the part of the research based around the Long List of items, which includes showing how the items were designed to cover a broad range of areas of autonomy, and how care was taken with the wording of the items. This Chapter also deals with the unforeseen issue of translation (which had far-reaching consequences on the timescale of the research). In the subsequent chapter I deal with the Short List, including how it was selected from the Long List, and I look at the issues involved in presenting it as a questionnaire.

Step	Description of original step	Notes	Section
1	Survey the literature of autonomy to find the main areas.	The range of areas to be covered was very large and this meant that selection had to be more focused on those which appeared to relate to the practical aims of the eventual questionnaire. This meant deciding which areas were peripheral to these aims and eliminating them. This meant that the “objectivity” was reduced.	3.2.2, 3.2.3, 3.7.1.1.1, 3.9.1.1
2	Compose items which cover the areas of autonomy. Cover each area with multiple items worded differently so that a very large number of items, perhaps as many as 200, would result.	Scales to cover each area and sub-area were found to be impractical for the present research due to the great number of items which would be necessary. Respondent fatigue would be a major limitation. Items had therefore to be chosen more subjectively with the aims of the eventual questionnaire in mind.	4.2
3	Compose a Likert scale; upload the list of items to the Internet as a form in a webpage with each response coded with a “score” to gather responses to the items. Respondents would give their feedback on the items to highlight any which they found unclear.	Access to the Internet was an issue in China. In other places students did not have access to computers in class. This meant that paper versions had to be printed and distributed and the papers returned by post and the data digitised. This severely slowed down the research. The issue of translation emerged as a major issue. The research was intended to be international, but items were not being understood as intended, and consequently it was decided to translate the items into a number of different languages. Quality of translation became a major issue. This became time-consuming as the items were translated into Mandarin Chinese and checked. Most of the respondents were Chinese, but more translations would have been done if time had been available.	4.3.3 4.3.4 3.13
4	Having amassed at least 200 responses from a wide range of respondents, items which had caused confusion would be removed. The data would then be factor analysed. The results of this would be used to eliminate items.	The number of respondents fell short of the numbers hoped for, which slowed down the research further. It meant that the numbers of responses required for factor analysis would be reached much later than envisaged. In order to proceed with the research it was necessary to reduce the number of items to make the ratio of respondents to items acceptable for factor analysis. This meant reducing the number of items to 50. This item reduction was carried out by statistical means	5.1
5	The remaining items would be used to form a well-designed questionnaire.		
6	The questionnaire would then be administered over the Internet to as broad a sample of respondents as possible, numbering at least 200.		

Step	Description of original step	Problems which emerged during the research	Section
7	Items which had caused confusion for respondents would be eliminated and the data would be factor analysed to eliminate items which did not contribute to the factors found.		5.1
8	The questionnaire would again be formatted in accordance with guidelines in the questionnaires literature.		
9	Data would be gathered from a large number of respondents. Teacher estimates of autonomy would be gathered longitudinally, ideally at beginning and end of a course.	At the large scale the stage of data gathering was not carried out. Only one independent group (i.e. in addition to my own class) was found. This was partly due to the delays caused by translating the questionnaire, but mainly due to lack of responses to my requests for volunteers.	5.2
10	Small scale data gathering in tandem with the questionnaire. Learners would be observed, interviewed, and feedback sought from their teacher to gather data on the respondents. I would teach one class and find another class and teacher at Warwick so that I could interview students who were not from my own class.	Qualitative data gathering in tandem with the questionnaire was carried out on a preessional course and on a third year undergraduate course. The preessional was my own class and the undergraduate class was that of one of the lecturers at Warwick University. Only one student who had agreed to be interviewed actually came for interview. This was caused by the earlier delays in the research caused by translation and by the difficulty of finding sufficient respondents to the Long List. The delay meant that interviews clashed with end of year assessments.	5.2
11	Comparison of the questionnaire and the qualitative data to gain insights on its viability and validity.	Comparison with the literature: the data so far gathered were standardised so that they could be combined to carry out factor analysis. The results of the factor analysis would be used for construct validity. Comparison was made of the small scale qualitative and quantitative data which had been gathered to see whether the questionnaire was functioning well for individual learners and classes to provide a qualitative indication of validity.	3.12.1

Table 3.8: Original outline plan for the research and subsequent changes

4 THE LONG LIST STAGE

4.1 Overview

In this chapter I describe the compilation of the Long List and give details of the data gathering which was carried out using it. The purpose of this stage was to gather data in order to carry out item reduction to select the best-performing items. As previously stated, factor analysis had originally been envisaged for this, but insufficient returns were obtained (the actual procedure used is given in Section 5.1).

4.2 Writing items

The rationale for item selection is given in Section 3.9 above. The Long List was a list of items from which it was intended to select items for the autonomy-measuring questionnaire. It was therefore necessary to consider item-wording guidelines for questionnaires when composing the items.

The wording of items is important, and this is particularly so if the questionnaire is self-administered. As Cohen et al. (2000: 251) point out, words are inherently ambiguous and great care needs to be taken to minimise this. In order for the present research to maximise the clarity of the items and to avoid item quality issues I decided to aim for the highest standards possible and consequently I will discuss item design in depth in this section. Guidelines for item writing are plentiful (e.g. Tourangeau et al. 2000; Oppenheim 1992; Dörnyei 2003; Cohen et al. 2000; Converse & Presser 1986).

4.2.1 Item wording guidelines

In Table 4.1 a number of authors' guidelines on item wording are summarised. This reveals the main areas of concern, which are vocabulary, length of items, simplicity

of grammar, leading or loaded questions, ambiguity, vagueness, double-barrelled items, the middle option in multiple choice answers, and multi-item scales. These will be discussed here as they relate to the item-authoring for the Long List.

4.2.1.1 Vocabulary

It was important not to make assumptions about which words respondents knew or do not know. Oppenheim (1992: 129) maintains it is necessary to find out by gathering feedback on the questions. However, rather than relying entirely on this, which may not reveal all difficulties with all populations, it was prudent to choose vocabulary which is not overly technical. Aiken (1997: 37) recommends that researchers “always choose the simplest way to say something”. Converse & Presser (1986: 15) agree that common concepts rather than abstract ones from the academic field should be used. They add (1986: 11) that chatty, over familiar language or “some subculture’s slang” should also be avoided. I carefully adhered to this.

4.2.1.2 Shortness

Shorter items are more likely to be understood (Converse & Presser 1986: 11-12; Dörnyei 2003: 52-53; Gillham 2000: 25). Oppenheim (1992: 128), Dörnyei (2003: 52), and Aiken (1997: 40) recommend no more than 20 words per question. The average length of the items in the Long List is 9.3 words, a figure well within this recommendation. However, two items have 21 words each:

166. I know which sense is best for me to use when learning (i.e. sight, or hearing, or touch, or physical movement)

167. I select learning techniques (i.e. taking notes, or drawing diagrams, or by listening, etc.) that suit my best way of learning.

Author	Tourangeau et al. (2000: 61)	Oppenheim (1992: 128-130)	Dörnyei (2003: 52-56)	Ellard & Rogers (1993: 17, from Dörnyei 2003)	Cohen et al. (2000: 248-249)	Converse & Presser (1986)
Vocabulary		Use simple words, avoid acronyms, abbreviations, jargon and technical terms		Thou shalt match the vocabulary used in items to the vocabulary of those who will respond to them.	Avoid highbrow questions	Use standard English, Use common concepts, not abstract concepts from the academic field.
Grammar	Avoid complicated syntax		Use simple and natural language	Thou shalt not use complex grammatical forms		
Shortness		Questions should not be too long	Aim for short [...] items			Use short questions when possible
Simplicity	Keep questions simple		Aim for [...] simple items		Avoid complex questions	Better to ask two or three simple questions rather than one complex question
Negatives		Avoid double negatives	Avoid negative constructions	Thou shalt not use “no” and “not” or words beginning with “un”	Avoid questions that use negatives or double negatives	Use direct rather than hypothetical questions Do not use double negatives. Beware of words which are implicitly negative.
Leading/Loading		-Beware ‘leading’ questions. -Beware loaded words Avoid proverbs and other popular sayings	Avoid [...] loaded words and sentences	Thou shalt not permit any loaded questions to appear in your questionnaire	Avoid leading questions	
Ambiguity	Define ambiguous or unfamiliar terms	-Beware the dangers of alternative usage. -Some words are notorious for their ambiguity and are best avoided or else defined	Avoid ambiguous [...] words and sentences			

Author	Tourangeau et al. (2000: 61)	Oppenheim (1992: 128-130)	Dörnyei (2003: 52-56)	Ellard & Rogers (1993: 17, from Dörnyei 2003)	Cohen et al. (2000: 248-249)	Converse & Presser (1986)
Vague/Specific	Avoid vague concepts, and provide examples when such concepts must be mentioned. Replace vague quantifiers with ranges that specify exact probabilities, frequencies, and so on					Wording that is specific and concrete is more likely to communicate uniform meaning”
Double-barrelled items		Avoid double-barrelled questions	Avoid double-barrelled questions	Thou shalt not use double-barreled (sic) items		Do not use double-barrelled questions
Don't Know/Middle option		Don't Know and Not Applicable categories are too often left out		Thou shalt not permit a non-committal response		
Memory	Decompose questions that cover multiple possibilities into simpler questions that cover a single possibility apiece	Don't over-tax the respondents' memories				
Questionnaire design		Pay due attention to detail such as layout...probes...	Include both positively and negatively worded items Avoid items that are likely to be answered the same way by everybody	Thou shalt have 40% to 60% true- or agree-keyed items Thou shalt not mix response formats within a set of questions		
Pre-testing		All closed questions should start their lives as open ones		Thou shalt pretest questions before collecting data		
Other					Avoid irritating questions	use redundancy but do not overdo.

Table 4.1: Comparison of item wording guidelines

Removing one word would bring these into line, for example deleting “i.e.”, but this would not necessarily make them any clearer. The length is due to the provision of clarification (within brackets) which fits with other guidelines recommending less vague or ambiguous language. In fact, Bourque & Fielder (1995: 46) make the point that sometimes a longer question is more specific and precise.

4.2.1.3 Simplicity and grammar

Gillham (2000: 25) recommends avoiding complex constructions, and Converse & Presser (1986: 16) recommend asking two or three simple questions instead. I used the first person as much as possible to make the items simple and immediate.

Converse & Presser (1986: 23) and Aiken (1997: 37) recommend asking about actual experiences rather than using hypothetical items. I aimed to use the present simple tense and avoid conditional or passive structures where appropriate.

It is frequently stated (Dörnyei 2003: 54; Oppenheim 1992: 128; Converse & Presser 1986: 13) that negative constructions should be avoided as they can confuse the respondent and are difficult for the researcher to interpret. However, I have retained some for reasons of clarity, such as item 147. “I worry if I don’t understand all the words in a text” which is difficult to alter. I inspected the respondents’ written feedback from the administrations to see whether negative wording did emerge as an issue, but no problems were reported.

4.2.1.4 Leading/Loading

Leading or loaded questions and vocabulary mean that the item “indicates the ‘good’ or ‘correct’ or ‘socially desirable’ answer” (Gillham 2000: 26) or “the questioner’s

own point of view” (Oppenheim 1992: 137). Dörnyei (2003: 54) makes the point that apparently innocuous words such as “merely”, “modern”, and “natural” can also influence a respondent. There were seven items in the questionnaire which may conceivably contain loaded words, for example:

64. Science books contain **only** facts

173. I approach a topic in a **careful**, step by step manner

174. I consider facts and come to **objective** conclusions

175. I look at causes and effects **logically**

One cause of loading is (unconscious) assumptions that the researcher presupposes are shared with the respondent (Oppenheim 1992: 128; Tourangeau et al. 2000: 42). According to Tourangeau et al. (2000: 43) unconscious assumptions are “inescapable in natural language questions”. It becomes a problem when the assumptions are not shared. The situation here is similar to that with negatives, discussed above, and I took the same measures to ensure that they did not have a significant disadvantageous effect.

4.2.1.5 Ambiguity

Definitions are not always shared (Converse & Presser 1986: 18) and in addition many common words (such as “week”, “dinner” etc.) have alternative meanings (Oppenheim 1992: 129). In fact Cohen et al. (2000: 249) feel that ambiguity cannot be avoided, rather we can only attempt to minimise it. This is another indication of the importance of collecting feedback about the items in the Long List. One item (221) was changed: “I want to learn in a more Western way” was adapted to “I want

to learn in a more learner-centred way”, though this was due to feedback from tutors and my fellow students rather than any negative feedback from respondents.

4.2.1.6 Vague/Specific

Converse & Presser (1986: 31) say “The more general the question, the wider the range of interpretations it may be given”. Tourangeau et al. (2000: 61) advocate the avoidance of vague concepts and the provision of examples when such concepts must be mentioned. Dörnyei (2003: 54) says words to be avoided are non-specific frequencies and values (such as “good” and “often”), and universals (such as “all”, or “never”). In the Long List there are some items with such “vague” concepts, for example item 11. “I am *good* at planning my learning”. In this case it would be unnatural and unduly pedantic, I feel, to specify exactly what “good” means; what matters is that the respondents are clear on the meaning, and that they are able to respond with their “true” perceptions and feelings.

4.2.1.7 Double-barrelled items

Double-barrelled items ask two questions at the same time so that respondents may not be able to agree or disagree to both parts at the same time. This can result in a “don’t know” or non-response, or if they do put an answer the researcher will not know if it applies to one “barrel” or both (Oppenheim 1992: 128; Dörnyei 2003; Aiken 1997: 37). There are some double-barrelled items in the questionnaire:

44. I know my strong points and weak points (How to respond if I know my weak points but not my strong points?)

49. I have changed the way I learn after thinking about it (What if I thought about it but did not change?)

55. I know why I did well or did badly (What if I know why I did well, but not why I did badly?)

If they are unanswerable or too open to randomness in the responses chosen by the respondents it might be safer to eliminate all, but it is a judgment of mine to retain them since the situations which may lead to difficulty in responding (given in brackets) appear to be quite forced and unlikely. Respondent feedback should indicate if there is a problem with them.

4.2.1.8 Multi-item scales

Dörnyei (2003: 33) recommends four or more items to cover a single point, forming a multi-item scale, but Oppenheim (1992: 143) argues that “it is simply not possible to develop multiple-item scales for everything”. This is the case with the Long List stage of the questionnaire. However, (and to conclude this section on item wording guidelines), Converse & Presser (1986: 10) distinguish four key concepts in item writing: simple language, common concepts, manageable tasks, and widespread information. All of these are reflected in the advice which has been discussed in this section, and so the emphasis of the items in the Long List is on the immediate and tangible, rather than the vague and hypothetical. There is a degree of judgement and subjectivity involved at this stage in the development of the questionnaire (and any questionnaire), but I am well aware of the possible pitfalls, and the use of respondent feedback on the items (Section 3.11.3) will be useful in identifying problems.

4.2.2 The choice of closed items

In order for the autonomy measuring instrument to be able to gather data from (potentially) large numbers of respondents in a form which could be analysed

quickly and simply a closed-response questionnaire was indicated. There is a choice to be made as there are basically four types of closed questions: rank ordering, multiple choice, dichotomous, and rating scale. I will look at each of these options in turn and explain how my decision was arrived at.

Rank ordering questions ask respondents to indicate priorities in a selection of responses (Cohen et al. 2000: 252). This type of question shows the respondents' ideas of the relative values of different items (Gillham 2000: 31). There are two major drawbacks for the present research: it is more difficult to answer than individual items and respondents may not be able to make the distinctions necessary (Cohen et al. 2000: 252; Dörnyei 2003: 44; Aiken 1997: 46); and also the results are not easy to process statistically (Dörnyei 2003: 45). For this reason I decided against using rank ordering questions.

Multiple choice questions are a familiar question type which has a statement followed by possible responses designed to represent fully the probable range of answers. The individual responses do not overlap (Cohen et al. 2000: 251) and they have the advantage of producing easy to process data (Cohen et al. 2000: 251).

Dichotomous questions are items limited to two possible responses, such as yes/no or agree/disagree, and do not offer middle options (see Section 4.7.1.1.8). As I hoped to obtain gradations of response beyond the black or white this method appears unsuited for a putative autonomy-measuring instrument, and for this reason I decided not to use dichotomous questions.

The remaining choice was rating scales. The most familiar manifestation of rating scales is the Likert scale, which is widely used and popular for questionnaires

because it is “simple, versatile, and reliable” (Dörnyei 2003: 36). It is a very common response format, and therefore it should be familiar to the respondents meaning they are more likely to answer in a reliable way. It also allows for more subtle responses than dichotomous questions (Cohen et al. 2000: 253). Another advantage is that the data produced are easy to process. For these reasons I chose the Likert scale for this research. However there are a number of points which need to be considered.

Firstly, as with all closed response formats, a disadvantage is that there is no way of telling if respondents might have wished to add something to their response (Cohen et al. 2000: 254), and the reasons why a certain option was chosen are not recorded (Gillham 2000: 32). However, during my research I planned to interview a number of respondents to find out more about why they answered in the ways they did, and a teacher using the questionnaire with a class would also be able to do this.

Secondly, there is the question of how many points to include in the scale, and whether to include middle and “don't know” options. Tourangeau et al. (2000: 248-249) talk of scale range effects, where the available range of response options can change beliefs about the question. The presence of an option, such as neutral, can actually be perceived as making that option an acceptable one. While it is generally true that the more options an item contains the more accurate it will be (Dörnyei 2003: 42) it is also the case that respondents often avoid the extremes of the scale (Tourangeau et al. 2000: 248-249; Gillham 2000: 32; Cohen et al. 2000: 254).

Gillham feels that this makes seven-point scales redundant, while Cohen et al. actually see larger scales as a way of reducing the effect of avoidance of extremes. For the present purposes I considered that a wide range of responses should be offered in order to find gradations which would make it easier to distinguish between

responses. At first, a five-point Likert scale was chosen. This was used in the United Arab Emirates, but then for the subsequent versions I used a seven-point scale to permit a greater distinction between responses.

Having an equal number of options produces a scale with no middle. Converse & Presser (1986: 36) ask whether offering the middle alternative encourages a non-committal response which is the easiest for the respondent or provides an additional gradation of opinion in the data. Cohen et al. (2000: 254) consider that having an odd number of items is better, as in not forcing a choice one does not hide the respondent's true feeling of ambivalence. I therefore decided to offer the middle option.

There is much of agreement in the literature that the "don't know" option should be included in rating scales (Muijs 2004:48; Aiken 1997: 45; Converse & Presser 1986: 35; Tourangeau et al. 2000: 43; Oppenheim 1992: 129). Authors point out that in any item there may be presuppositions. These are "inescapable in natural language questions" (Tourangeau et al. 2000: 43). Presuppositions can take the form of "assumptions not only about the nature of what is to be measured, but also about its very existence" (Converse & Presser 1986: 35). Consequently, it is perfectly possible that respondents may not be familiar with an element of the question, and so may really be unable to give a definite answer. It would be tempting to omit the "don't know" choice to constrain a decisive response, but as Oppenheim (1992: 129) asks "do we really want to obtain 'forced' responses which are virtually meaningless?". In the light of this, I decided to offer a "don't know" option, though this decision was taken after the first data gathering in the UAE.

In this section a variety of question/answer types have been discussed. Gillham (2000: 34) thinks that to maintain respondent interest question types should be varied. Cohen et al. (2000: 258) recommend including “questions that are likely to be of general interest” and by mixing attitude and behaviour questions. This was done in the Long List of items, and feedback from respondents indicated that they found it interesting (see for example Section 4.3.2).

4.2.3 Items in the Long List

Table 10.2 in the Appendix shows all the 256 items in the Long List and gives the areas relevant to autonomy which they were intended to cover. These are the areas established in the Literature Review as described in Section 3.9.

It is not clear that it is possible to write one item which covers one discrete area, and it is not clear that the intention of the author can precisely match the understanding of the reader (see Section 4.2.1). It is also apparent that it is not possible to have four or more items for each possible dimension of autonomy (as discussed in Section 4.2.1.8). With the help of my supervisors I decided to use a grid (which developed into Table 10.2) in which I could place items and control the coverage of all the areas – though it was not possible to show how the items would objectively (i.e. without my decisions) interrelate until the factor analysis stage.

I populated the grid with items to form the Long List. To help me with ideas for questions and examples of phrasings I surveyed existing questionnaires, e.g. Oxford’s (1990) Strategy Inventory for Language Learning, Sinclair (1999) on metacognition, motivation questionnaires from Gardner & MacIntyre (1993) and Dörnyei (2001), Cotterall’s (1995) learner beliefs questionnaire, and many others. I

also looked at learner training materials, and discussed ideas for items with supervisors. I then wrote, adapted or borrowed items making sure they followed the guidelines I have described in Section 4.2.1. In overview, the complete questionnaire development process from start to finish went from more subjective to more objective. Compiling the Long List was the most subjective stage, and I therefore took care to ensure (using the grid) that the items represented the wide range of specified areas which had been identified in the Literature Review as central to autonomy for learning. In this way I intended that the subsequent more objective steps (Short List and then final factor analysis) would be selecting from the widest possible range of items. Table 4.2 shows the main sections of the thesis where each of the areas is discussed.

Area	Section
Social Interaction	2.4.8
Motivation	2.4.7
Responsibility	2.4.6
Actions/Behaviours	2.4.4
Metacognition	2.4.5, 2.5.1.5
Control (capacity for)	2.4.3
Attitudes to learning (beliefs)	2.5.1.3
Confidence	2.5.1.3
Strategies (knowledge and use of strategies)	2.5.1.4
Skills (Areas of Critical thinking, Reading, Writing, Listening, Speaking, Information, Grammar, Vocabulary etc.)	2.5.1.2

Table 4.2: The general areas for questionnaire items with thesis sections

4.3 Subject groups

At the stage of the Long List there were two data gatherings, one in the United Arab Emirates (UAE), and one via the Internet which was open to all volunteers.

4.3.1 The UAE data gathering

The respondents were 53 Emirati women in their second year at the Higher Colleges of Technology (HCT) at two campuses, Fujairah Women's College and Abu Dhabi

Men’s College (though female students) in the United Arab Emirates (UAE). HCT is an English-medium tertiary institution. All the students had attended state high schools before passing the national Common Educational Proficiency Assessment (CEPA test) with a score of 150 or more, which is the level required for university or Higher Diploma courses at the HCT (UAE Ministry of Higher Education and Scientific Research 2006). It was assumed that regarding culture and education the participants were all very similar.

Sample size		54
Country of origin	UAE	53
	UK	1
Average age	(Not gathered)	? years
Sex	Male	0
	Female	54
Study Level	Final year high school	1
	2 nd year Higher Diploma	53

Table 4.3: Participants in HCT UAE data gathering

The respondents were presented with the full 256 items of the Long List (“Format A”). Each was a statement with 5-point Likert scale response, worded “Strongly agree”, “Agree”, “Neither agree nor disagree”, “Disagree”, and “Strongly disagree”.

The responses were coded in sequence from 5 “Strongly agree”, to 1 “Strongly disagree”. 73 of the 256 items were reverse coded, i.e. where items were thought to bear a negative relation to autonomy the coding of the responses would be reversed. Normally-scored and reverse-scored items all had “Neither agree nor disagree” = 3. Use of symbols instead of words was rejected as it was not certain that they would be equally well understood in different cultures.

The questionnaire was administered in paper format in three classes by the class teachers during 55 minute English lessons. It was not possible to be present in the classes, but the teachers informed the class that the questionnaire was a survey and not a test, and that the results were not going to be part of their course, and that it was completely anonymous. This was also written at the top of the questionnaire. The items in the questionnaire were written in both English and Arabic. Most participants completed the questionnaire in less than 40 minutes. Respondents were allowed to complete the questionnaire outside class if they wished.

I will now examine all the feedback which I received from respondents. The length of the questionnaire caused problems: *“Can we take it home?”*, *“It’s boring”*, *“We need breaks”*, *“It’s too long”*, and queries arose regarding some items:

- *“Some questions are confusing, specifically 37 vs. 38”*

37 is “If I find a word that I don’t know, I always ask a teacher first” and 38 is “If I find a word that I don’t know, I look it up in a dictionary first”. This indicates either a problem with the translation of these items, or that respondents are not be processing the word “first”.

- 16

This item was queried: it reads “When I learn something new I feel good because I can stop learning it”. The problem with this was not specified, but it may be a translation problem.

- 26 *“lucky”?*

Item 26 is “I feel lucky when I get good marks”. I judge this to be a translation problem produced by the unexpected appearance of a word such as “lucky” in the questionnaire.

- 62 *“What’s the point of this question? Is it linked to the previous one?”*

Item 61 is “I find it difficult to tell facts from opinions when reading” and item 62 is “Sara eats a lot of sweets so she must be fat”. These are both items aimed at probing the skill area of critical thinking and are not connected in any other way.

-88 *“Did Mr David write the questions himself? I think there should be a comma here.”*

Item 88 is “If I am not sure about something it bothers me”. Only one respondent made this point.

- 90

Item 90 is “I learn exclusively about college subjects”. This was a candidate for rewording.

- 105” *Arabic ‘might be’ is English ‘is’”.*

Item 105 reads “When I read in English I think about what the source of the text is”. This is probably a subjunctive construction in Arabic. The meaning does not appear to be significantly different. This again appears to be a translation problem.

- 184 and 185 – *“What’s the difference?”*

Items 184 “I like negotiating with other students in class” and 185 “I like class discussions”. The distinction is pair or small group work versus whole class discussions.

The two lowest-scoring items, 26 and 88, are also ones which received negative feedback. Five of the lowest scoring items were near the end of the instrument where fatigue effects are to be more expected.

4.3.2 English language Internet data gathering

This version of the questionnaire (Format B) is a web page form. It is essentially the same as the UAE version, though the previous 5-point Likert scale response format was changed to a 7-point scale to obtain more discrimination between responses. A “don’t know” box was also added. Spaces were provided at the end for item-specific and general feedback, and a section also asked for background information on age, nationality, gender, place/mode of study, study level (undergraduate or postgraduate), years learning English, and time taken to complete the questionnaire. The participant details are shown in Table 4.7 below.

Sample size		8
Country of origin	UK	1*
	Colombia	1
	Sweden	1
	KSA	1
	UAE	1
	China	1
	Switzerland	1
	Canada	1
Average age		38 years
Sex	Male	0
	Female	8
Study Level	Undergraduate	3
	Postgraduate	4
	Non-student	1
Average time studying English		15.6 years
Average completion time		31 minutes

*This respondent was an L1 English answering the questions for L2 Japanese.

Table 4.4: Participants in English Language Internet group

Problems were reported in feedback from respondents with items:

- 198. I hate to study with less than my best effort. “It is unclear.”
- 17. If I must finish a job at a certain time I finish early. “The word ‘job’”.
- 62. Sara eats a lot of sweets so she must be fat. “A trick question?”.
- 68. I trust the Internet. “What aspect of the Internet?”.
- 76. I rely on the teacher. “Rely on the teacher - For what?”.
- 183. I am happy to use different worksheets from the rest of the class.
“What worksheets? - maybe, ‘I am happy to work with a different task from the rest of the class’???”.

The Chinese respondent said “I think there are some repetitions in the questionnaires. That is, some questions reflect the same aspect of learning English”. The Swedish respondent said:

A general comment: Maybe a comment in the beginning about keeping an eye open for the tricky questions. It took time to scroll through again to find the ones I hesitated on. Maybe the questions are too many???

The first point was not repeated by other respondents, but the second shows problems with the length of the questionnaire.

One of the respondents (from the UK, learning Japanese) gave more feedback via email:

just did your survey! Really made me think as I am just wondering whether to start serious study of Japanese again. I tend to be plunged into language using situations so far above my ability level (at work) that it feels like whatever I do wouldn't show any results for a long time, so I need a lot of energy and consistency but lack it. The survey made me realise I am illiterate as regards libraries and reference works in Japanese. It was a stage in taking myself seriously as a J-learner again, so thank you!

This was interesting feedback as it concurs with the idea that an autonomy measuring instrument has a function as a formative tool to encourage reflection.

4.3.3 Mandarin Chinese Internet data gathering

Prior to this data gathering the items had been translated into Chinese due to issues which had emerged with the UAE data gathering (see Section 3.13). This was the Long List with 7-point Likert scale for responses delivered online (Format C). It can be seen in parallel translation in the Appendices Section 10.3. Details of the participants are given in Table 4.5.

Sample size		6
Country of origin	China	6
Average age		22.5 years
Sex	Male	2
	Female	4
Study Level	Undergraduate	5
	Teacher	1
Average time studying English		12 years
Average completion time		26 minutes

Table 4.5: Participants in Mandarin language Internet group

In the feedback section there were comments about perceived repetitions: “*something has repeated for several times!!*” and “*There are some repetitions of these questions*”. One of the respondents contacted me via email and specified items 200, 202, and 219 as having “*similar connotations with the previous questions*”. He also commented on how long the questionnaire was, and reported a mistake with the Chinese characters in item 97. He queried item 251 “I use real English texts (i.e. not

made for students) in my learning” suggesting it be reworded to “*authentic texts*”.

This was not changed as non-technical language was preferred.

There were a small number of returns for this data gathering. This can be ascribed to its length (also reported as a problem in feedback from the other Long List data gatherings) or to one of the following reasons:

- Compulsory item completion. When the “submit” button was pressed the page was automatically checked for unanswered items and if there were any it showed a message asking the user to complete the questions marked with a red asterisk (this is a design option built in to the Warwick SiteBuilder web page authoring software). This message was not very prominent and it is possible that the message was not being noticed and users thought they had successfully finished and left the page without it actually being submitted.
- In China servers external to the country are only accessible if Internet subscribers pay extra. Despite prior enquiries regarding accessibility, this fact had not been revealed to me.
- The questionnaire took the form of one long page, and this may have caused problems for potential respondents with slow Internet connections.

4.3.4 The BNU Mandarin Chinese data gathering

All respondents were English majors at Beijing Normal University (BNU) and details are shown in Table 4.6 below.

The Long List of items translated into Chinese was again used. However, following the problems with the Internet delivery of the questionnaire in China (see Section

4.3.3) the questionnaire was reformatted as a form in a Microsoft Word document. Another variation on the previous data gatherings was that the items were grouped with the intention of “breaking up” the list to make the 256-items seem less daunting to respondents. There were two feedback sections, the first to report individual items which had caused problems, and the second for general reflections.

The Long List was distributed by email via two local administrators, one of whom was briefed by me and the other by the first. An incentive was offered (free umbrellas) to reward the respondents for their time and effort. The administrators informed the respondents that the questionnaire was a survey and not a test, that the results were not going to be part of their course, and that it was completely anonymous.

There was much feedback from this group. 11 respondents said it was repetitious, for example one said “I think some of the questions are just repeating one another. There are some overlapping”. Three respondents referred to its length, for example:

I think the questionnaire is a little bit long. Anyone who is facing this questionnaire should have plenty of patience. Yet, it is quite convenient for us to finish, especially using a computer, because we can click the buttons.

Four said it had made them reflect on their learning, for example:

I think the questionnaire is very useful and helpful for me to rethink my learning methods, study attitude, and some of my believes

Two respondents commented on the difficulty of selecting a response:

I find some questions are too absolute, which makes me hard to decide, such as the use of “always” and so on.

and

some of the questions are not so clear and the choices are very general, so it is a little bit difficult to distinguish the difference among the choices

It is interesting that more respondents did not make these points as the questions were written or selected with an aim of not assuming a specific context. It appears that most respondents did not have great difficulties answering the items.

Sample size		48
Country of origin	China	48
Average age		20.27 years
Sex	Male	5
	Female	42
	Not given	1
Study Level	1 st year undergraduates	21
	3 rd year undergraduates	27
Average time studying English		8 years
Average completion time		28.5 minutes

Table 4.6: Participants in BNU group

4.4 Chapter summary

In this chapter I have examined in some depth the item wording guidelines from the literature which were followed in order to obtain well-formed items for the Long List. I have given details of the data gathering which was carried out using the Long List. In the next chapter I will describe how the data were processed to reduce the length of the Long List and so produce the Short List.

5 THE SHORT LIST STAGE

In this chapter I will describe the stage of the research involving the selection of the Short List of items and its subsequent use for data gathering, combined with the collection of teacher estimates of autonomy and interview data.

The Long List stage had not proceeded entirely according to plan. There had been problems and delays due to difficulties obtaining enough responses to requests for participants. There had also been at the Long List stage the unforeseen and lengthy translation process. Time constraints meant that it would not be possible to wait for more participants. The original timetable and the plan for the research had to be reconsidered to make something that would still be worthwhile and achievable, but would be modified. As a result the Short List stage described here is not how it was originally expected to be. The main changes are:

1. Item selection by factor analysis was not possible and an alternative method was found (see Section 5.1).
2. Construct validity checks using factor analysis (see Section 3.8.2) would still be carried out but would be delayed until the ratio of respondents to items had been improved by reducing the number of items, and also by amassing more data.
3. Large scale comparison of questionnaire data and teacher estimates and observations would not be possible due to the shortage of volunteers to participate. This would now only take place at the small scale (see Section 5.2).

5.1 Item selection process for Short List

The item selection process was based on multiple parallel techniques rather than relying on one procedure. It was hoped that this would help to preserve the construct and reduce the chances of an idiosyncratic result. The techniques used were:

- Respondents' feedback
- Range of response
- Standard deviation
- Polarisation of response
- Discrimination index

As previously stated it had been intended to use factor analysis and Cronbach's alpha reliability coefficients, but this resulted in a selection of only 20 items as there were not enough respondents at this point in the data gathering for these procedures to function. In meetings with my supervisor it was decided that 20 items were probably not enough to cover a multidimensional construct such as autonomy. I therefore decided to find a different selection procedure which would produce a longer list of still robust items, but not so long that it would produce respondent fatigue and so I aimed at a number of items which could be completed in between 15 and 30 minutes (for the slower respondents). A figure of 50 items was thought to be a good balance between construct coverage and respondent fatigue.

Item analysis using the Rasch technique (Hughes 2003: 228-233) was also considered as a way of reducing the number of items, but was rejected as it involved deciding on an order of "difficulty" of questions which had at this stage not been empirically established. I left the question of whether any subjective choice of items

would be necessary until after the selection procedure had given its results, and then would only be carried out for clearly stated reasons.

This selection process did not look at whether items were good indicators of autonomy or not. The purpose of this process was to look at the qualities of the instrument's items in terms of their statistical performance only. Maintaining the coverage of autonomy by the 50 items was important and this was checked, as shown in Table 5.2, and found to be satisfactory. In Sections 5.1.1 to 5.1.7 which follow I present the process used in the item selection.

5.1.1 Respondents' feedback

In all, respondents were unclear or gave negative feedback on 19 items, and these are shown in Table 5.1. More details of the feedback can be found in the individual reports of the data gatherings in Sections 4.3.1, 4.3.2, 4.3.3, and 4.3.4. Numbers or question marks in the group columns indicate how many individuals queried the item. In the HCT column question marks indicate that the class teachers did not record how many respondents queried the item. The BNU column is blank because, although two respondents did say that two pairs of items seemed too similar, it was not possible to identify clearly to which items they were referring. Most respondents in all data gatherings did not mention any items as problematic. There were two items, 62 and 198, which received negative feedback from respondents in different groups which makes them the stronger candidates for rejection by the respondents' feedback criterion.

No.	Item	Respondent Groups			
		HCT	English Internet	Chinese Internet	BNU
16	When I learn something new I feel good because I can stop learning it?	?			
17	If I must finish a job at a certain time I finish early?		1		
26	I feel lucky when I get good marks.	?			
37	If I find an English word that I don't know, I always ask a teacher first.	?			
38	If I find an English word that I don't know, I look it up in a dictionary first.	?			
62	Sara eats a lot of sweets so she must be fat.	?	1		
68	I always trust the information I find on the Internet.		1		
76	I rely on the teacher when learning.		1		
88	If I am not sure about something it bothers me.	?			
90	I learn exclusively about college subjects, and nothing else.	?			
105	When I read in English I think about what the source of the text is.	?			
183	I am happy to use different worksheets from the rest of the class.		1		
184	I like negotiating with other students in class.	?			
185	I like class discussions.	?			
198	I hate to study with less than my best effort.		1	1	
200	I try to find out how to learn better.			1	
202	I look for opportunities to practice English.			1	
219	I reflect on my learning.			1	
251	I use real English texts (i.e. not made for students) in my learning.			1	

HCT=Higher Colleges of Technology. BNU=Beijing Normal University

Table 5.1: Summary of all items queried by respondents

5.1.2 Range of response

Consideration of the range of responses is useful as it indicates whether the item elicits differing or predominantly the same responses. A wider range of responses is desirable for the instrument as the measurement of autonomy requires that an item can detect differences among respondents. The range of items was expressed as the

number of different options on the Likert scale which had been selected by more than x per cent of the respondents, where x is a figure that enables a ranking order of items to be produced. Initially it was set at $x=2.1\%$, but this did not produce a useful ranking and it was raised to 10%. The results were added to an Excel spread sheet of all the items, the table was sorted by this criterion, and then the top 50 items were indicated with shaded boxes. Appendix 10.4 shows a table covering the complete selection process. The Range of Response criterion figures can be seen in the column labelled A.

5.1.3 Standard deviation

Questionnaire items with a larger standard deviation will be more useful for the autonomy measuring instrument since standard deviation indicates the spread of responses for an item. A wider spread of responses will permit more discrimination between the individuals in the group answering the questionnaire. The standard deviations for each item were added to the selection table (Appendix 10.4) in the column labelled C, and the table was then sorted by this criterion and the top 50 items were indicated by a shaded box.

5.1.4 Polarisation of response

Polarisation is here intended to mean the extent to which an item divided respondents. The normal bell curve distribution has most respondents in the middle of the range, whereas a polarised distribution has most respondents occupying the extremes of the range with few in the middle. This is desirable for the instrument as its items need be able to detect differences among respondents in order to measure autonomy. Column B in Appendix 10.4 shows the polarities in the selection table. The polarity of the items was calculated by finding items where the middle Likert

option had fewer responses than the surrounding options. This difference was quantified as the difference between the middle option and the closest of the neighbouring options. The table was sorted in Excel by this criterion and the top 50 items were indicated with a shaded box. The selection table shows that the respondents had not favoured the extremes of the Likert scale. This indicates that they had not treated it as a test but rather, as advised, they had treated it as a survey.

5.1.5 Selection table

The sum of the figures for each item in columns A, B, and C was added together and the total added to the table in column D. The table was sorted by this criterion and again the top 50 were indicated with a shaded box. Finally, the number of shaded boxes for each item was put in column E and the table sorted by this figure and, secondarily, by standard deviation as many items had the same number of shaded boxes. The resulting table had 75 items with at least one shaded box. This ranking will be referred to as List A. A separate procedure was then employed to produce a second list (List B) using discrimination indexes, as described in the following section.

5.1.6 Discrimination indexes

A discrimination index (DI) indicates how consistently an item distinguishes between higher and lower performing respondents in tests (Hughes 2003: 226-228). If an item does not correlate with the overall result indicated by the instrument it is not an effective question.

While the autonomy measuring instrument was not envisaged as a test this technique can still be used to indicate whether an item is an effective question in an instrument

since the DI looks at the relation of a single item to all the items. The procedure for compiling the DI was as follows.

All respondents were given a figure by calculating the total coded score for all their answers. The respondents were then ranked high to low. This set was divided into two sets, the respondents with higher overall scores and those with lower overall scores. Next individual items were addressed: the total for each item was summed within the higher group and separately within the lower group. Item totals within the higher group should be greater than the same items totalled in the lower group. The larger the difference the more effectively the item has discriminated between the groups. If an item has a similar score in both groups it has not performed well in indicating differences in degree of autonomy. This is either because the question addresses a point which does not vary with levels of autonomy, or because it was in an ineffective question (for example with confusing wording, or with an obvious “best” answer). Items which had a large difference between the scores in the higher and lower groups were selected for the new questionnaire. Items which had not discriminated would be less effective items and would be candidates for removal. I now had two lists of items ranked in order of decreasing effectiveness by their respective criteria, i.e.:

List A. items ranked by their performance in standard deviation and polarity of response.

List B. items ranked by their discrimination index.

In the following section I will show how these were combined and explain my rationale.

5.1.7 Combined selection procedure

The two lists were now combined, taking the highest ranking items from each list so as to obtain a final selection of 50 items. The rationale for this was that the two lists could have compared in three possible ways:

1. All of the top 50 items are identical between the lists (the selection would be straightforward)
2. None of the top 50 items are shared between the lists (it would be simply a matter of taking the top 25 items from each list)
3. The two lists share some of their top 50 items (the actual situation)

Case 3 was the actual situation as many of the items appeared in both lists (which was encouraging because it confirmed their selection), but there were also many items which did not co-occur.

One option which was considered was to take only the common items and to do this look further down the rankings beyond the top 50, perhaps to 75 or 100 or beyond. This would ensure that all items had been selected by both criteria. However, the problem which was found with this approach was that since all the items are common in the full lists (since they are both composed of the same items though in different orders) it meant that an item high in list A may be very low in list B and therefore an undesirable match. Likewise, if starting in list B with a highly ranked item it may match at a much lower level in list A. It would therefore be necessary to find a way to balance the two lists in a manner which, as much as possible, favoured the selection of items which were highly ranked in both lists. For this reason I decided to favour the highly ranked items rather than relying solely on the matching

items. This meant choosing the top x items from each list, then combining these two lists. The items which occurred in both lists therefore had two instances in the combined list, one from each of the initial two lists. For each item which had two instances one instance was removed, and the total number of remaining items in the combined list was summed. If the total was less than 50 the process was repeated with the top $x+1$ items from both lists (i.e. moving one step down the ranking). If the total was more than 50 the process was repeated with the top $x-1$ items from both lists (i.e. moving one step up the ranking). The process was repeated until a value of x was found which produced the desired 50 items. The diagram below (Figure 5.1) illustrates the process.

The 50 items which resulted from this process were checked to see whether they were items which had undergone any rewording as a result of respondent feedback which may have changed their response qualities due to changes of meaning or connotation. If there had been such changes it would have made it problematic to combine the data from the item responses in all the data gatherings. All the 50 items were found to be identical except for item 76 which had originally read “I rely on the teacher” but had been changed to “I rely on the teacher **when learning**”. Item 76 had been changed as a result of feedback from one respondent in the English language Internet data gathering (see Section 4.3.2), but this change was not regarded as significant in the context of a questionnaire about learning and therefore no action was taken.

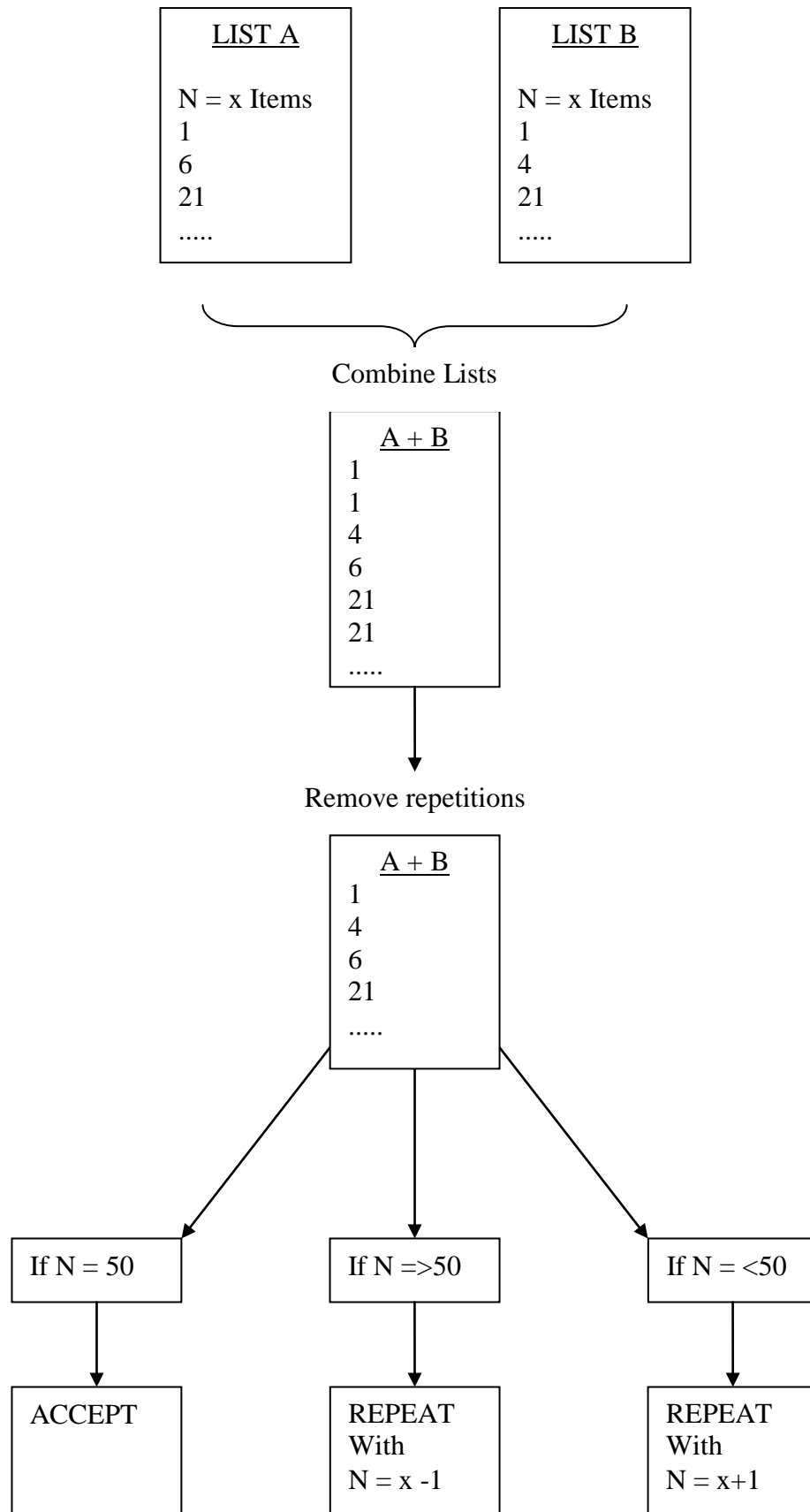


Figure 5.1: Process for combining lists

5.1.8 Items in the Short List

Table 5.2 (overleaf) shows the items selected for the Short List with the areas of autonomy which they cover. In order to show (in terms of representative coverage of the autonomy areas) that the Long List and the Short List were broadly similar, Table 5.3 shows the comparison of the Long and Short Lists.

This new version of the questionnaire, Format E (in Chinese) and Format F (in English) is Internet-based and can be found at:

<http://www2.warwick.ac.uk/fac/soc/al/research/groups/llp/circal/omall>

The 50 items are arranged into face valid groupings to “break up” the questionnaire and to give a sense of logical organisation (see Appendices Section 10.3). The rationale for breaking up and grouping items in a questionnaire has been discussed previously in Section 3.10.1.2.

5.2 *Small scale research*

It had been intended (as described in Section 3.7.1.1.3) to conduct large scale research at this point, involving longitudinal class-based administrations of the Short List and observations and estimates of autonomy made by the class teachers, followed by interviews. The rationale for this had been the quantitative validation of the questionnaire using sufficient quantities of data to allow statistical significance to be investigated. In the event, the large scale research was not possible because of translation and slow data gathering at the Long List stage.

Item	Reading	Writing	Listening	Speaking	Grammar	Vocabulary	Attitudes to learning	Social Interaction	Motivation	Confidence	Responsibility	Actions/Behaviours	Strategies	Metacognition	Control	Skills
008. I am good at studying on my own							x		x		x					
017. If I must finish a job at a certain time I finish early							x		x		x	x				
021. All lessons are equally valuable							x		x				x	x		
023. Students should always do what the teacher says							x		x		x				x	
026. I feel lucky when I get good marks							x		x	x	x			x	x	
030. The teacher's job is to give me all the information							x		x		x				x	
046. I can describe the learning Strategies I use										x			x	x		
049. I have changed the way I learn after thinking about it.							x		x	x	x	x	x	x	x	x
055. I know why I did well or did badly							x						x	x		x
076. I rely on the teacher							x			x	x				x	
086. When I read an English text I need to understand every word in it	x					x	x			x			x	x	x	
095. To read you must proceed word by word	x					x	x			x			x			x

Item	Reading	Writing	Listening	Speaking	Grammar	Vocabulary	Attitudes to learning	Social Interaction	Motivation	Confidence	Responsibility	Actions/Behaviours	Strategies	Metacognition	Control	Skills
100. Last time I read an English text I predicted the content of it	x						x		x	x		x	x		x	x
109. I predict the content before I listen			x				x			x		x	x		x	x
110. Every word is important for understanding a listening text			x			x	x			x			x		x	
111. The last time I listened to English I tried to predict the content			x				x					x	x	x	x	x
112. I worry if I don't understand everything when I listen			x				x			x			x		x	x
125. I change the way I write according to who will read it		x						x		x	x	x	x	x	x	x
130. My writing is better now than it was a year ago		x							x	x				x		
138. I know techniques to help me remember vocabulary						x				x	x		x	x	x	x
140. To remember vocabulary you need to be talented						x	x								x	
142. I fix my problems in vocabulary						x			x	x	x		x	x	x	x
147. I worry if I don't understand all the words in a text	x					x	x			x			x	x	x	

Item	Reading	Writing	Listening	Speaking	Grammar	Vocabulary	Attitudes to learning	Social Interaction	Motivation	Confidence	Responsibility	Actions/Behaviours	Strategies	Metacognition	Control	Skills
150. I worry if I don't understand all the grammar in a text					x		x			x				x	x	
153. Learning grammar is a talent some people have, but not all					x		x								x	
175. I look at causes and effects logically							x						x			
187. I think learning English is more difficult for me than for the average learner								x	x	x				x		
189. I learn English because I have to							x		x		x			x	x	
193. The other students know English better than me								x	x	x				x	x	
194. The other students are more confident than me at speaking English				x				x		x				x		
195. I worry that other students will laugh at me when I speak English				x				x	x	x				x	x	
196. I am confident I can learn English well							x		x	x				x	x	
203. I organise my time for studying							x		x	x	x	x	x		x	
205. I notice how other people use English								x					x			
212. I talk to others about how I feel about learning English				x			x	x		x		x	x	x		
220. I am ready to learn in unfamiliar ways							x		x	x		x	x	x		

Item	Reading	Writing	Listening	Speaking	Grammar	Vocabulary	Attitudes to Learning	Social Interaction	Motivation	Confidence	Responsibility	Actions/Behaviours	Strategies	Metacognition	Control	Skills
229. I can choose the method of learning that suits me best							x			x	x		x	x	x	
230. My way of learning will never change							x			x			x		x	
231. I can study independently							x	x	x	x		x	x	x	x	
234. It is my job to check my work for mistakes		x			x	x	x		x	x	x	x				
236. I am good at making choices									x	x				x	x	
237. I am an active dynamic person										x		x		x		
238. I choose the exercises I work on							x		x	x	x				x	
243. I decide what I need to read	x						x		x	x	x		x	x	x	
246. Memorizing answers is the best way to learn							x			x			x	x		
251. I use real English texts (i.e. not made for students) in my learning	x						x			x		x	x			
252. I know how to find information in a library										x		x			x	x
253. I know how to use English language reference books (encyclopedias, dictionaries, etc.)	x					x				x		x			x	x
254. I know the parts of a book (index, glossary, contents, chapters)	x															x
256. I know how to find the information I need on the Internet	x						x			x		x	x			x
Total	9	3	4	3	3	9	35	8	22	37	16	16	27	27	31	14
Percentage	18	6	8	6	6	18	70	16	44	74	32	32	54	54	62	28

Table 5.2: Short List's coverage of areas of autonomy

		Reading	Writing	Listening	Speaking	Grammar	Vocabulary	Attitudes to learning	Social Interaction	Motivation	Confidence	Responsibility	Actions/Behaviours	Strategies	Metacognition	Control	Skills
Long List	N	24	14	6	15	22	35	194	47	127	157	114	82	142	144	152	90
(*256)	%	9%	5%	2%	6%	9%	14%	76%	18%	50%	61%	45%	32%	55%	56%	59%	35%
Short List	N	9	3	4	3	3	9	35	8	22	37	16	16	27	27	31	14
(*50)	%	18%	6%	8%	6%	6%	18%	70%	16%	44%	74%	32%	32%	54%	54%	62%	28%

*= Number of items in list. N= the number of items which refer to the area. % = N expressed as a percentage

Table 5.3: Comparison of coverage of the areas of autonomy in the Long List and Short List

The smaller scale research took place with two classes. Both classes were for the ELTCS (English Language Translation and Cultural Studies) at Warwick University, one was at third year undergraduate level (see Section 5.3 for a description of the class) and one was a preessional course (see Section 5.4 for a description of the class).

The procedures for gathering the data in the two classes were as follows. For the ELTCS year-3 group:

1. Students complete the questionnaire on-line
2. Teacher observation of students, leading to
3. Estimates of autonomy
4. I interview volunteer students
5. I interview the teacher.

Ideally for the research there would have been stages prior to this where the students answered the questionnaire and the teacher, with only first impressions of the students, made estimates of their autonomy. This was not possible as the class came to my research only after the course had already started. At step 4 volunteers were requested but (despite incentives) only one volunteer attended for interview.

For the ELTCS preessional group the data gathering procedure was:

1. The teacher (myself) estimates the levels of autonomy of the students based on first impressions. On the same day ...
2. Students complete the questionnaire on-line (but the data are stored only and not analysed)
3. Observation of students during the course, leading to ...

4. Second estimates of autonomy at the end of the course. On the same day ...
5. Students complete the questionnaire a second time.

There would be no interviews with my own students for reasons of objectivity.

5.2.1 Estimating autonomy levels

A positive result in the comparison of questionnaire data and estimates gained from large samples would have persuasively indicated that an autonomy measuring instrument such as the one used in the present research was viable and could potentially serve a useful function in similar contexts. In the event the necessary level of response did not materialise and the data gathering was limited to the small scale. At this small scale, if the questionnaire did match the estimates it would be suggestive only, because significant correlations are hard to demonstrate with small samples. The reasons why comparison was nevertheless carried out at this small scale were to gather qualitative data, and for the purpose of gathering quantitative data to demonstrate the principle and techniques which had been envisaged for the large scale research. A good result for this part of the research would be a strong suggestion that the instrument (or the principle of such an instrument) merited further research.

I decided to use teacher estimates of autonomy for comparison with the questionnaire data because estimates are what I have always used to gauge my students' levels of autonomy. Estimates are the de facto way of making low stakes decisions because they do not require the commitment of large amounts of time and resources. They are what the autonomy measuring instrument seeks to match or improve upon and are

therefore a source of data which is appropriate for the purposes of the present research.

I wanted to have a situation which would be as close as possible to that found in a normal course. In fact both classes were normal courses and not especially set up for research purposes; this was so that data gathering would stay close to the aim of the research i.e. to reflect how estimates are actually made. For this reason I did not specify in advance what should be observed – it was clear to me that this would have defeated the idea of using estimates; it would have been, in effect, another untried and untested procedure for measuring autonomy and would therefore have been unsuitable for comparison with the autonomy measuring questionnaire.

Different teachers will base their estimates on differing evidence. An experienced practitioner in the field of autonomy, a naïve practitioner, or one somewhere between in terms of experience, will probably make their estimates based on their own understandings of autonomy. The two teachers (myself and the other teacher) involved in the present research are knowledgeable and experienced in autonomy. I would not be investigating *how* teachers make estimates (though this would be an interesting area for separate research). It was not necessary to know this to find whether the questionnaire produced results which correlated well with teacher estimates. However, for qualitative putting of flesh on bones I intended to ask the ELTCS year-3 teacher during interview how she decided on the estimates with the question “Are there any features which identify the more and less autonomous students?”.

5.2.2 Student interview

There were two aims for carrying out student interviews in the small-scale data collection stage of the research. Firstly, I hoped to be able to probe more deeply the respondents' questionnaire answers and so gain insights beyond the limitations of the Likert scale. I hoped that interviewees would be able to give more reasons for their answers and so illustrate to me whether and how the questionnaire was functioning.

Secondly, I wanted to obtain further information to improve the design of the questionnaire and the wording of the items. I was interested in how they had interpreted items, especially to reveal whether the decontextualized nature of the items had confused them and whether they had answered in a general way or with a specific personal context in mind. Additionally, I wanted to find more about how the Likert scale had functioned, for example why they had respond "strongly agree" rather than "very strongly agree" or "agree".

The above aims indicated using the questionnaire itself as the basis of the interview, i.e. a semi-structured format based on working through the autonomy measuring instrument, item by item, with the interviewee and probing more deeply where interesting points arose. The questionnaire is designed for the general context of tertiary level learners of English and is not aimed specifically at a single specific group. This means that it would be useful to obtain responses from the interviewees which provided more detailed, particular, and concrete information than the abstract and decontextualized data from the autonomy measuring instrument (Richards 2003: 53).

There were time constraints as potential respondents were in the final term of the year with exams due shortly. I therefore proposed a maximum of 1 hour for each interview.

5.2.3 Teacher interview

The interview guide (below) I drew up was intended to allow for opportunities for comparisons between the questionnaire and the interview. A key element of the interview was the elicitation of the teacher's estimates of individual students' autonomy levels on a scale of 1-10. The questions I intended to ask were:

- In general how would you characterise the level of autonomy of the ELTCS students?
- Is this the same as when they began the course in Warwick?
- Would you say that some students are more autonomous than others?
- Are there any features which identify the more and less autonomous students?
- 10 students completed the autonomy measuring questionnaire: could you estimate their levels of autonomy so that I can see if there is any correlation with their questionnaire results?
- In general do you think they are socially confident? (For example, whether students worry what other students will think about their accent, etc.)
- In general do you think the students approach language tasks in a top-down or bottom-up way?
- In general are they able to find the information they need either in books and reference works or on the Internet?
- Do they tolerate ambiguity?

- Are they good at making their own choices?
- Do they actively reflect on their own learning practices?
- Do you think they analyse input, i.e. do they notice the way people speak or the way language is used in texts?
- Do you think they have self-confidence in their language use?
- In general are they able to learn independently?
- Do you think they engage in self-evaluation?
- Do you think they have a sense of agency in their learning?

5.3 ELTCS year 3 data gathering

5.3.1 Description

The questionnaire was the 50-item web page version in Mandarin Chinese with 7-point Likert scale and feedback section.

Sample size		10
Country of origin	China	9
	Taiwan	1
Sex	Male	2
	Female	8
Study Level	3 rd year undergraduate	10
Average time studying English		8.9 years
Average completion time		16 minutes

Table 5.4: Participants in ELTCS third year group

This respondent group was involved in the BA in English Language, Translation and Cultural Studies course (ELTCS). This is a four-year course, the first two of which are at Renmin University in China where the students study English, translation, and

Chinese and British culture. Then the following two years are spent at Warwick University where they do cultural studies, translation, and English language and linguistics. The participants were 12 third year students on the ELTCS BA nearing the end of their first year at Warwick University. Their English language level was the equivalent of the University of Warwick English Language Test grade BBB.

Respondents were given the incentive of a one-to-one grammar consultation on language problems in their essays. The introduction to the questionnaire informed the students that the results were not part of their course and that it would be confidential, and the class teacher informed the respondents that the questionnaire was a survey and not a test.

Respondents completed the questionnaire outside class. It was hoped to interview some of the respondents and all were invited though only one student (known here as “T1-St-b”) actually attended the interview.

5.4 Preessional ELTCS group data gathering

This group, as the previous one, was involved in the ELTCS BA course. The students were following a preessional course at Warwick University in preparation for their first year in the UK. I was their teacher for the 5-week preessional course. I made a more detailed data gathering with this group. On day 1 I initially estimated the students' levels of autonomy and noted them down. They then completed the 50-item questionnaire (Format E) in Chinese. I informed the respondents that the questionnaire was a survey and not a test, that the results were not going to be part of their course, and that it was completely anonymous. Details of the sample are given in Table 5.5.

Sample size		10
Country of origin	China	10
Average age		20 years
Sex	Male	7
	Female	3
Study Level	2 nd year undergraduate	10
Average time studying English		9 years

Table 5.5: Participants in ELTCS preessional group

For the next five weeks I taught the students and was also observing and making notes on them for their end of course reports, including notes on their apparent autonomy. In week 5 I estimated the students' levels of autonomy for a second time and noted them down, had the students complete the questionnaire a second time, and wrote reports of the students' autonomy based on my observations.

5.5 Other Warwick preessional students data gathering

I took the opportunity while teaching on the ELTCS preessional to gather data from other students in different classes totalling 35 respondents. These data were standardised and added to the total data set to improve the factor analysis process. As can be seen from Table 5.6 most were East Asian learners in tertiary education. Their teachers informed the respondents that the questionnaire was a survey and not a test, that the results were not going to be part of their course, and that it was completely anonymous.

Sample size		35
Country of origin	China	14
	Taiwan	10
	Thailand	6
	Turkey	2
	Italy	1
	Japan	1
	Syria	1
Average age		25.5
Sex	Male	15
	Female	20
Study Level	Tertiary	34
	Teacher	1
Average time studying English		13.6 years
Average completion time		9.7 minutes

Table 5.6: Participants from other preessional groups

5.6 Taiwan group data gathering

The 14 respondents in this data gathering were all studying English at Fu-Jen Catholic University, Taiwan. As can be seen from Table 5.7 all were of Taiwanese background. The sample consisted of three males, nine females and two were not specified. Respondents completed the questionnaire online in class and the average time for completion was 27.5 minutes. The questionnaire was the 50-item web-based Mandarin Chinese version with 7-point Likert scale and feedback section (i.e. Format E). The introduction to the questionnaire informed the students that the results were not part of their course and that it would be confidential, and the class teacher informed the respondents that the questionnaire was a survey and not a test. It was not possible to interview any of the respondents. The data were standardised and added to the total data set to improve the factor analysis process.

Sample size		14
Country of origin	Taiwan	14
Average age		23.5 years
Sex	Male	3
	Female	9
	Unspecified	2
Study Level	Tertiary	14
Average time studying English		8.5 years
Average completion time		27.5 minutes

Table 5.7: Participants in Taiwan group

5.7 Ethics

Punch (2005: 277-278) gives 11 relevant questions to ask oneself, based on Miles & Huberman (1994: 290-297), before, during, and after research, which can be summarised as addressing the issues of “harm, consent, deception, privacy and confidentiality of data” Punch (2005: 277). Warwick University has published ethics guidelines for its researchers available at:

www2.warwick.ac.uk/fac/soc/pais/staff/page/administration/phdguidelinesweb1.doc.

The issues of ethics which were dealt with in my research were addressed in four areas: harm, consent, honesty, and confidentiality.

The Warwick University ethics guidelines say “protect participants (such as interviewees) from any harm as a result of their research”. No harm was expected as a result of responding to the questionnaire, and indeed there were indications of actual benefits to the respondents in the form of increased awareness and reflection on learning.

Regarding consent, the Warwick University ethics guidelines say “Researchers should carry out investigations, interviews etc. with the informed, and prior, consent of participants” and Punch (2005: 277) recommends asking oneself the question “Do the people I am studying have full information about what the study will involve?”. The consent should also be freely given Richards (2003: 140).

In order to avoid confusing or influencing the respondents before or during the research I introduced the questionnaire in general terms as an investigation of how learners learn and of learner attitudes to learning and did not discuss concepts of autonomy, or use the word autonomy within the items. I was aware that, as Richards (2003: 140) says, there is a “fine dividing line between limited description and deliberate deception” but I felt it was important to elicit data from respondents who had not been influenced by me so that they would remain in that respect more representative of the general population of language learners. Where educational institutions were involved informed consent was sought from the representative of the institution, in this case the data gatherings at the Higher Colleges of Technology, The Beijing Normal University, and the University of Warwick.

Regarding confidentiality, the Warwick University ethics guidelines say “Confidentiality of participants and their data must be respected. Details that would allow individuals to be identified must not be published or made available to anyone not involved in the research project unless explicit consent is given”. All data were anonymised and the respondents were informed that their contributions would be confidential. In addition when the respondents were participants in a course they were also informed that the questionnaire and its data were not part of their course.

5.8 Limitations

In this section I will summarise the constraints which I have identified in the course of the three methodology-related chapters (see especially Table 3.8). There were three unforeseen sources of delay which forced the original plan for the research to be modified. Firstly, there was the necessity of translating the items, and then the discovery that the casual translation method which had been adopted was not adequate and would require much time and organisation to put right. Secondly, the response rate to the questionnaire was low. This was partly due to the number of items to be responded to (as indicated by feedback such as that in Section 4.3.1), and partly due to Internet access problems in China (see Sections 4.3.3 and 4.3.4). The delays meant that:

1. It was necessary to change the method of data reduction as there was an insufficient ratio of respondents to items for factor analysis (see Section 5.1).
2. It was not possible to organise large scale data gathering, and so the statistically significant comparison of teacher estimates with questionnaire results was not possible, though small scale comparison went ahead (see Section 5.2).
3. Delayed interviews now coincided with end of year assignments so only one student attended (see Section 5.2.2).
4. It was necessary to standardise and combine data from different formats of the questionnaire to permit factor analysis for construct validity checking (see Section 3.12.1).

Limitations are discussed in Section 8.3.3.

5.9 Conclusion

At the conclusion of this stage in the research a total of 185 useable responses to the 50 items had been gathered. Observations and autonomy estimates had been carried out for the two ELTCS groups, and one student and one teacher had been interviewed. In the following chapter I present the findings of the analysis of the data and provide a discussion of the issues and questions which arise such as how significant the data are, and the implications of the restricted data gathering.

6 DATA ANALYSIS

6.1 Introduction

In this chapter I will present the different data analyses and results. I will distinguish large scale and small scale data. I will use the larger scale data to look at the validity of the instrument using factor analysis to find the autonomy-related model which is embodied in the questionnaire items and which it is necessary to examine to see how much validity and reliability it has. At this larger scale (standardised accumulated data gathered from 185 respondents to the 50 items) it is possible to produce statistically significant data which makes this an important step. In order to do this I will investigate how the questionnaire performs using three methods:

- Checking Cronbach's alpha for internal reliability of scales
- Comparing the model with a previous model in the autonomy measuring literature produced by Cotterall (1995)
- Checking to establish whether the model reflects the areas found in the Short List or is novel.

At the smaller scale I will continue to examine the questionnaire using quantitative and qualitative data. I will investigate at this level by:

- Seeing whether there are correlations between the questionnaire results and the two teachers' estimates (i.e. mine and the ELTCS year-3 teacher's) in order to shed light on the viability of the instrument
- Seeing whether there are correlations between the questionnaire and the interview with the ELTCS year-3 teacher

- Comparing two students' questionnaire results with more qualitative data to see whether the questionnaire corresponds with the qualitative data, and to understand more about their experiences in order to see whether the questionnaire has produced useful information about them

The data gathered and analysed here will be useful for exploring the questionnaire's viability and credibility at the level of individual learners and teachers. At this level the quantified data will not be treated as statistically significant but have a purpose in prompting questions for further discussion.

In the light of the above small and larger scale investigations, I will consider the validity and reliability of the instrument's data and consider how much it is appropriate to conclude from them.

6.2 *Larger scale data*

6.2.1 Introduction

In this section I will present the analysis of the larger scale data gathered using 185 pooled and standardised responses to the questionnaire. I will describe the process of factor analysis which was adopted and the groupings which emerged. This picture of the questionnaire's representation of autonomous learning will be investigated by comparing it with the literature.

6.2.2 Factor analysis procedure

After all data gatherings were complete the questionnaire data for the 50 items were standardised and combined to be factor analysed in SPSS. In Section 3.12.3 in the Methodology Chapter I described the process to be used to extract factors. The procedure adopted was cautious with much redundancy built in to it so as to avoid

idiosyncratic results. The procedure I adopted called for 12 separate factor analyses as I would use the two recommended extraction methods (maximum likelihood and principal component analysis) and the two recommended rotation methods (oblimin and varimax) with the three possible alternatives for the number of factors to extract which had been found by examining the scree plot of eigenvalues. The break of slope in the scree plot occurred at 8, 9, or 10 which indicated the number of factors to extract. I would run the analysis for each of these. Table 6.1 below shows these 12 procedures and gives each a reference number. These were then checked for low-loading and cross-loading items, and for factors with insufficient items for identification to be possible (as described in Section 3.12.3.3).

Extraction method	Rotation method	No. of factors forced	Reference No. of factor analysis run
Maximum likelihood analysis	Oblimin	8	1
		9	2
		10	3
	Varimax	8	4
		9	5
		10	6
Principal component analysis	Oblimin	8	7
		9	8
		10	9
	Varimax	8	10
		9	11
		10	12

Table 6.1: The 12 runs of factor analysis

An example of one of these 12 runs (number eight in Table 6.1) is included here to illustrate the stages in the process. The factor matrix with loadings is given in Table 6.2.

Scale (*CA)	Items	1	2	3	4	5	6	7	8	9
1. (.773)	046. I can describe the learning strategies I use.	.624								
	109. I predict the content before I listen.	.608								
	049. I have changed the way I learn after thinking about it.	.535								
	138. I know techniques to help me remember vocabulary.	.456								-.325
	237. I am an active dynamic person.	.455				-.311				
	130. My writing is better now than it was a year ago.	.424			.398					
	142. I fix my problems in vocabulary.	.419								-.388
	212. I talk to others about how I feel about learning English.	.403								
	234. It is my job to check my work for mistakes.	.403								
055. I know why I did well or did badly.										
2. (.702)	147. I worry if I don't understand all the words in a text. (R)		.768							
	112. I worry if I don't understand everything when I listen. (R)		.653							-.305
	150. I worry if I don't understand all the grammar in a text. (R)		.612							
	086. When I read an English text I need to understand every word in it. (R)		.603							
	110. Every word is important for understanding a listening text. (R)		.375							
	026. I feel lucky when I get good marks. (R)		.349			-.348				
3. (.731)	256. I know how to find the information I need on the Internet.			.710						
	196. I am confident I can learn English well.			.674						
	254. I know the parts of a book (index, glossary, contents, chapters).			.628						
	253. I know how to use English language reference books (encyclopedias, dictionaries, etc.).	-.388		.450				.444		
	175. I look at causes and effects logically.			.428	-.349					
	252. I know how to find information in a library.			.411			.367			
	125. I change the way I write according to who will read it.			.405						
4. (.676)	231. I can study independently.				-.630					
	220. I am ready to learn in unfamiliar ways.				-.609					
	008. I am good at studying on my own.				-.593					
	229. I can choose the method of learning that suits me best.	.370			-.459					
	236. I am good at making choices.				-.404					
	205. I notice how other people use English.				-.336					

Scale (*CA)	Items	1	2	3	4	5	6	7	8	9
5. (.100)	017. If I must finish a job at a certain time I finish early					-.581				
	111. The last time I listened to English I tried to predict the content.					.560	.312			
	203. I organise my time for studying.				-.358	-.449				
	100. Last time I read an English text I predicted the content of it.					.403				
	023. Students should always do what their teacher says. (R)					.325				
	076. I rely on the teacher when learning. (R)									
6. (.447)	243. I decide what I need to read.						.543			
	238. I choose the exercises I work on.						.530			
	251. I use real English texts (i.e. not made for students) in my learning.			.316			.493			
7. (.718)	194. The other students are more confident than me at speaking English. (R)							-.764		
	193. The other students know English better than me. (R)							-.733		
	187. I think learning English is more difficult for me than for the average learner. (R)					.328		-.466		
	195. I worry that other students will laugh at me when I speak English. (R)				-.342	.377		-.443		
8. (.099)	153. Learning grammar is a talent some people have, but not all. (R)								.588	
	021. All lessons are equally valuable (R)								.458	
9. (.634)	230. My way of learning will never change. (R)									-.657
	246. Memorizing answers is the best way to learn. (R)						-.426			-.563
	140. To remember vocabulary you need to be talented. (R)									-.557
	189. I learn English because I have to. (R)									-.516
	095. To read you must proceed word by word. (R)									-.450
	030. The teacher's job is to give me all the information. (R)								.329	-.356

*CA= Cronbach's alpha



Item groupings



Items to be rejected

Table 6.2: Loadings from factor analysis run eight

Summarised in Table 6.3 below is the procedure followed for accepting/rejecting items (showing the reasons) both by Cronbach's alpha and from the indications shown in the factor matrix.

Scale	Cronbach's alpha		Matrix indications	Conclusion
	Item-total correlation	CA if item deleted		
1	-	-	Item 55 is low-loading	Item 55 rejected
2	Item 26 is below .3	Item 26 would improve scale if deleted	-	Item 26 rejected
3	-	-	Item 253 is cross-loading	Item 253 rejected
4	-	-	-	Keep all items
5	All items are below .3	Items 111 and 100 would improve scale if deleted	Item 76 is low-loading	Reject whole scale
6	All items are below .3, but above .2	-	-	Keep all items
7	-	Item 195 would improve scale if deleted	-	Item 195 rejected
8	Both items are well below .3	-	Only two items loading on this factor	Reject whole scale
9	Items 30 and 189 are both below .3 but are otherwise strong	-	Item 246 is cross-loading	Keep items 30 and 189 Item 246 rejected

Table 6.3: The rejected items and scales with reasons for exclusion

The groupings which remained are shown in Table 6.4, below with suggested factor names. Similar processing was carried out for the other 11 runs. Since all the 12 runs did not agree precisely on the composition of the final factors they were combined. This was achieved by selecting the most frequently occurring items in each factor grouping. The process is illustrated in Table 6.5 below.

Scale	Items	Loading	Factors
1	046. I can describe the learning strategies I use.	.624	Metacognition
	109. I predict the content before I listen.	.608	
	049. I have changed the way I learn after thinking about it.	.535	
	138. I know techniques to help me remember vocab'.	.456	
	237. I am an active dynamic person.	.455	
	130. My writing is better now than it was a year ago.	.424	
	142. I fix my problems in vocabulary.	.419	
	212. I talk to others about how I feel about learning English	.403	
234. It is my job to check my work for mistakes.	.403		
2	147. I worry if I don't und' all the words in a text. (R)	.768	Linguistic Confidence
	112. I worry if I don't und' everything when I listen. (R)	.653	
	150. I worry if I don't understand all the grammar in a text. (R)	.612	
	086. When I read an English text I need to understand every word in it. (R)	.603	
	110. Every word is important for understanding a listening text. (R)	.375	
	026. I feel lucky when I get good marks. (R)	.349	
	256. I know how to find the information I need on the Internet.	.710	
196. I am confident I can learn English well.	.674		
254. I know the parts of a book (index, glossary, contents, chapters).	.628		
253. I know how to use English language reference books (encyclopedias, dictionaries, etc.).	.450		
175. I look at causes and effects logically.	.428		
252. I know how to find information in a library.	.411		
125. I change the way I write according to who will read it.	.405		
4	231. I can study independently.	-.630	Self-Reliance
	220. I am ready to learn in unfamiliar ways.	-.609	
	008. I am good at studying on my own.	-.593	
	229. I can choose the method of learning that suits me best.	-.459	
	236. I am good at making choices.	-.404	
	205. I notice how other people use English.	-.336	
5	243. I decide what I need to read.	.543	Making Choices
	238. I choose the exercises I work on.	.530	
	251. I use real English texts (i.e. not made for students) in my learning.	.493	
6	194. The other students are more confident than me at speaking English. (R)	-.764	Social Comparison
	193. The other students know English better than me. (R)	-.733	
	187. I think learning English is more difficult for me than for the average learner. (R)	-.466	
7	230. My way of learning will never change. (R)	-.657	Locus of Control
	140. To remember vocabulary you need to be talented. (R)	-.557	
	189. I learn English because I have to. (R)	-.516	
	095. To read you must proceed word by word. (R)	-.450	
	030. The teacher's job is to give me all the information. (R)	-.356	

Table 6.4: Final groups for factor analysis run eight with suggested identifications

Group	Items	1	2	3	4	5	6	7	8	9	10	11	12	Selection (N)	
1	150. I worry if I don't understand all the grammar in a text. (R)	x	x	x	x	x	x		x	x	x	x	x	11	
	147. I worry if I don't understand all the words in a text. (R)	x	x	x	x	x	x		x	x	x	x	x	11	
	112. I worry if I don't understand everything when I listen. (R)	x	x	x	x	x	x		x	x	x	x	x	11	
	086. When I read an English text I need to understand every word in it. (R)	x	x		x	x	x		x				x		7
	110. Every word is important for understanding a listening text. (R)		x			x			x				x		4
	153. Learning grammar is a talent some people have, but not all. (R)		x				x								2
2	254. I know the parts of a book (index, glossary, contents, chapters).	x	x	x	x	x	x	x	x	x	x			10	
	253. I know how to use English language reference books [..]	x	x	x			x	x	x		x	x	x	9	
	256. I know how to find the information I need on the Internet.				x	x	x	x	x	x	x			8	
	252. I know how to find information in a library.	x		x			x	x	x		x	x	x	8	
	251. I use real English texts (i.e. not made for students) in my learning.	x	x	x			x	x			x	x	x	8	
	175. I look at causes and effects logically.				x	x			x	x	x	x		6	
	196. I am confident I can learn English well.				x	x			x	x	x			5	
	125. I change the way I write according to who will read it.				x	x			x	x				4	
	130. My writing is better now than it was a year ago.				x	x								2	
	243. I decide what I need to read.												x	1	
	238. I choose the exercises I work on.												x	1	
3	194. The other students are more confident than me at speaking English. (R)	x	x	x			x	x	x		x	x	x	9	
	193. The other students know English better than me. (R)	x	x	x			x	x	x		x	x	x	9	
	187. I think learning English is more difficult for me than for the average learner. (R)	x	x	x			x	x	x		x	x	x	9	
	195. I worry that other students will laugh at me when I speak English. (R)								x		x			2	
	189. I learn English because I have to. (R)							x					x	2	
4	230. My way of learning will never change. (R)	x	x	x	x	x	x		x	x		x		9	
	140. To remember vocabulary you need to be talented. (R)	x	x	x	x	x	x		x	x		x		9	
	246. Memorizing answers is the best way to learn. (R)	x	x	x	x	x	x			x		x		8	
	189. I learn English because I have to. (R)				x	x	x	x		x		x		6	
	095. To read you must proceed word by word. (R)				x	x			x			x	x	5	
	030. The teacher's job is to give me all the information. (R)				x	x							x	3	
	110. Every word is important for understanding a listening text. (R)				x								x	2	
	187. I think learning English is more difficult for me than for the average learner. (R)				x									1	
021. All lessons are equally valuable (R)												x	1		

Group	Items	1	2	3	4	5	6	7	8	9	10	11	12	Selection (N)	
5	237. I am an active dynamic person.				x	x	x	x	x	x				6	
	234. It is my job to check my work for mistakes.				x	x	x	x	x	x				6	
	212. I talk to others about how I feel about learning English.				x	x	x	x	x	x				6	
	138. I know techniques to help me remember vocabulary.				x	x	x	x	x	x				6	
	109. I predict the content before I listen.				x	x	x	x	x	x				6	
	049. I have changed the way I learn after thinking about it.				x	x	x	x	x	x				6	
	046. I can describe the learning strategies I use.				x	x	x	x	x	x				6	
	142. I fix my problems in vocabulary.				x			x	x	x				4	
	130. My writing is better now than it was a year ago.							x		x			x	4	
	238. I choose the exercises I work on.				x	x	x							3	
	236. I am good at making choices.				x	x			x					3	
	196. I am confident I can learn English well.							x					x	x	3
	175. I look at causes and effects logically.							x					x	x	3
	125. I change the way I write according to who will read it.							x					x	x	3
	251. I use real English texts (i.e. not made for students) in my learning.				x	x									2
	231. I can study independently.				x										1
	229. I can choose the method of learning that suits me best.				x	x			x						3
	203. I organise my time for studying.				x	x									2
055. I know why I did well or did badly.							x	x						2	
6	231. I can study independently.	x	x	x					x			x	x	6	
	236. I am good at making choices.	x	x				x		x			x		5	
	229. I can choose the method of learning that suits me best.	x	x	x			x		x					5	
	220. I am ready to learn in unfamiliar ways.						x		x			x	x	4	
	237. I am an active dynamic person.	x	x	x										3	
	234. It is my job to check my work for mistakes.	x	x	x										3	
	205. I notice how other people use English.								x			x	x	3	
	203. I organise my time for studying.	x	x				x							3	
	008. I am good at studying on my own.								x			x	x	3	
	212. I talk to others about how I feel about learning English.	x												2	
	046. I can describe the learning strategies I use.	x	x											2	
	238. I choose the exercises I work on.	x												1	
	195. I worry that other students will laugh at me when I speak English. (R)												x		1

Group	Items	1	2	3	4	5	6	7	8	9	10	11	12	Selection (N)
6 cont.	142. I fix my problems in vocabulary.	x												1
	138. I know techniques to help me remember vocabulary.	x												1
	109. I predict the content before I listen.	x												1
	095. To read you must proceed word by word. (R)	x												1
	049. I have changed the way I learn after thinking about it.	x												1
7	237. I am an active dynamic person.										x	x	x	3
	234. It is my job to check my work for mistakes.										x	x	x	3
	212. I talk to others about how I feel about learning English.										x	x	x	3
	138. I know techniques to help me remember vocabulary.										x	x	x	3
	109. I predict the content before I listen.										x	x	x	3
	049. I have changed the way I learn after thinking about it.										x	x	x	3
	046. I can describe the learning strategies I use.										x	x	x	3
	142. I fix my problems in vocabulary.										x	x		2
	055. I know why I did well or did badly.										x		x	2
	229. I can choose the method of learning that suits me best.										x			1
130. My writing is better now than it was a year ago.										x			1	
8	196. I am confident I can learn English well.	x	x	x										3
	130. My writing is better now than it was a year ago.	x	x	x										3
	125. I change the way I write according to who will read it.	x	x	x										3
	256. I know how to find the information I need on the Internet.	x	x											2
	212. I talk to others about how I feel about learning English.				x									1
	095. To read you must proceed word by word. (R)		x											1
9	030. The teacher's job is to give me all the information. (R)									x	x	x		3
	110. Every word is important for understanding a listening text. (R)									x	x			2
	095. To read you must proceed word by word. (R)									x	x			2
	076. I rely on the teacher when learning. (R)									x	x			2

Group	Items	1	2	3	4	5	6	7	8	9	10	11	12	Selection (N)
9	021. All lessons are equally valuable (R)									x		x		2
	153. Learning grammar is a talent some people have, but not all. (R)										x			1
	140. To remember vocabulary you need to be talented. (R)										x			1
10	236. I am good at making choices.									x	x			2
	220. I am ready to learn in unfamiliar ways.									x	x			2
	008. I am good at studying on my own.									x	x			2
	231. I can study independently.										x			1
	229. I can choose the method of learning that suits me best.									x				1
	205. I notice how other people use English.									x				1
11	203. I organise my time for studying.										x			1
	251. I use real English texts (i.e. not made for students) in my learning.								x	x				2
	243. I decide what I need to read.								x	x				2
12	238. I choose the exercises I work on.								x	x				2
	195. I worry that other students will laugh at me when I speak English. (R)		x											1
	111. The last time I listened to English I tried to predict the content.		x											1
13	017. If I must finish a job at a certain time I finish early		x											1
	246. Memorizing answers is the best way to learn. (R)											x		1
	230. My way of learning will never change. (R)											x		1
	140. To remember vocabulary you need to be talented. (R)											x		1

R = Reverse-coded

 Acceptable

 Borderline


 Not acceptable

Table 6.5: Combining the 12 factor analyses

In the 12 analyses some groupings occurred repeatedly, composed of substantially the same items. The groupings in the table (Table 6.5) are arranged in order of their frequency of occurrence, for example the first group is identifiable in 11 of the 12 analyses. Within the groupings the individual items are arranged in descending order of frequency. This arrangement made it possible to find the most prevalent groupings with the most robust compositions. In total it can be seen that there were 13 groupings, but below group 6 the groupings are composed of items which had already occurred more frequently in preceding groups (items indicated in Table 6.5 with the “Not acceptable” hatching). This analysis therefore produced six groups which were selected. The reliability of these scales was checked by Cronbach’s alpha. Two items were removed: item 153 (to improve the Cronbach’s alpha of group 1, and because its corrected item total correlation was below .3), and item 195 (to improve the Cronbach’s alpha of group 3).

In Table 6.6 below I summarise the process of identifying the underlying connection between the items in each group. The steps in the process have already been presented in Section 3.12.3.3.4. Firstly, as shown in the column “Key Words”, I took or interpolated the key words or concepts contained in the items. Then, as seen in the column “Interpretation”, I attempted to put the key concepts together to make an interpretation which would fit with the key words and the sense of the items. Finally, as seen in the column “Factor”, I attempted to choose one idea which would encapsulate the underlying connection between the items in the grouping.

Group (*CA)	Items	Key Words	Interpretations	Factor
1 (.712)	150. I worry if I don't understand all the grammar in a text. (R) 147. I worry if I don't understand all the words in a text. (R) 112. I worry if I don't understand everything when I listen. (R) 086. When I read an English text I need to understand every word in it. (R) 110. Every word is important for understanding a listening text. (R)	worry (confidence) text understand every word, everything, all grammar receptive skills	Confidence in approach to texts Tolerance of ambiguity	Linguistic Confidence
2 (.749)	254. I know the parts of a book (index, glossary, contents, chapters). 253. I know how to use English language reference books [..] 256. I know how to find the information I need on the Internet. 252. I know how to find information in a library. 251. I use real English texts (i.e. not made for students) in my learning. 175. I look at causes and effects logically. 196. I am confident I can learn English well. 125. I change the way I write according to who will read it.	know book parts find information internet library authentic text	Familiar with how information is stored and able to retrieve information	Information Literacy
3 (.718)	194. The other students are more confident than me at speaking English. (R) 193. The other students know English better than me. (R) 187. I think learning English is more difficult for me than for the average learner. (R)	other students English confident the average learner	Learners' comparing themselves with perceptions of other students'	Social Comparison
4 (.642)	230. My way of learning will never change. (R) 140. To remember vocabulary you need to be talented. (R) 246. Memorizing answers is the best way to learn. (R) 189. I learn English because I have to. (R) 095. To read you must proceed word by word. (R)	never change need to talented memorizing have to must word by word	Being in control; having power , or being powerless	Locus of Control

Group (*CA)	Items	Key Words	Interpretations	Factor
5 (.764)	237. I am an active dynamic person. 234. It is my job to check my work for mistakes. 212. I talk to others about how I feel about learning English. 138. I know techniques to help me remember vocabulary. 109. I predict the content before I listen. 049. I have changed the way I learn after thinking about it. 046. I can describe the learning strategies I use. 142. I fix my problems in vocabulary. 130. My writing is better now than it was a year ago. 238. I choose the exercises I work on.	Active dynamic check work talk to others feel techniques predict thinking about [learning] learning strategies fix problems better now choose exercises	Actively and consciously using techniques and strategies to aid learning,	Metacognition
6 (.709)	231. I can study independently. 236. I am good at making choices. 229. I can choose the method of learning that suits me best. 220. I am ready to learn in unfamiliar ways. 205. I notice how other people use English. 203. I organise my time for studying. 008. I am good at studying on my own.	Independently making choices choose unfamiliar ways notice organise my time studying on my own	Ability to self-direct	Self-Reliance

*CA = Cronbach's alpha. (R) = Reverse Coded

Table 6.6: Identification of factors

6.2.3 Groupings emerging from the factor analysis process

I shall look at the individual factor groupings in the following paragraphs, and explain my analysis of each one. These groupings will be discussed further in the Discussion Chapter, Sections 7.3.2 and 7.3.3.

6.2.3.1 Group 1 Linguistic Confidence

All the items involve the idea of approaching the “understanding” of a text from a building-block or atomistic level. The first three items all use the word “worry”, which implies a feeling of insecurity or a lacking of confidence in the approach to a target language text. It suggests an anxious learner who is not tolerating ambiguity and is getting caught up in details at the expense of a broader top-down approach. However, it must be remembered that the Likert scale registers from agreement to disagreement. The poles are security-insecurity, and this suggests that the underlying connection is (level of) confidence with regard to target language text, hence the label Linguistic Confidence. It is the strongest category to emerge from the selection process (see Table 6.5) in the sense that it is the most consistently represented across the 12 factor analyses.

6.2.3.2 Group 2 Information Literacy

This is the second strongest category to emerge from the selection process, and it relates to information literacy skills, i.e. the ability to retrieve information from online and printed sources. The items when arranged in order of frequency (see Table 6.5) show that the first four items all deal directly with using sources to find information. The fifth-ranked item (item 251) deals with using authentic texts, a subject which has a superordinate connection with the idea of finding information in

English language sources. The items lower in the category have less specific relevance to the central idea, but could still have a link in that the logic of cause and effect (item 175), and confidence with regard to learning (item 196) are both, if not absolutely necessary, at least useful supports for information literacy. The final item (125) shows an awareness of different types of text which, again, is a useful skill to have for the effective interpretation of text.

6.2.3.3 Group 3 Social Comparison

It was immediately clear that this group involved a relation or comparison to other people, i.e. “other students” (194, 193) and “the average learner” (187). One of the items, 194, specifically uses the word “confident”; it is the highest loading item and so suggests that this factor involves self-belief based on comparisons with others.

6.2.3.4 Group 4 Locus of Control

I see this category as being linked by ideas of power and feeling in control. The ability to learn is seen, by the more highly-scoring respondents, not as a talent bestowed on a lucky few (item 140), but perhaps as a potential for success which can be harnessed by all perhaps by diligence or by knowledge of techniques, or by feeling at liberty to try. Item 246 does not specifically follow in the theme of power, but there is a sense in the idea of memorizing which is suggestive of conservatism and inflexibility in learning, and therefore of the opposing ideas of trying (and being able to try) new ways of learning, of being able to be flexible, and of being in control rather than a passive subject. Item 95 is akin to 246 in the sense that the respondent who agrees with the item will see the text as being in control; the text has to be followed in an externally preordained linear way rather than in the way one chooses to approach it for oneself. The expressions “need to”, “have to”, and “must” in items

140, 189, and 95 all indicate a control which, though not necessarily external, is limiting to the individuals' perceptions of the scope of their will.

This category seems to be showing an element which deals with learners being empowered and having a consequent broader range and outlook on what is possible, and so gives them the sense of learning being within their own control. For this reason I have labelled this category Locus of Control. It corresponds well with the idea of autonomous learners being in control.

6.2.3.5 Group 5 Metacognition

The items in this group (see Table 6.6) appear to be united by learners making active and conscious use of techniques and strategies to aid their learning. Learning strategies and techniques are evident in items 138, 46, 212, and 109. Reflecting on learning is shown in items 49, 46 and 130. Awareness of reflection can be seen in items 49 and 46. The active engagement with learning combined with reflection which is clear in this grouping suggests that this category is associated with learner metacognition.

6.2.3.6 Group 6 Self-Reliance

Items 231 and 8 both deal with independent study, while the use of "I can" (items 231 and 229) indicates ability, and the expressions "I am good at" (items 8 and 236) combined with item 220 "I am ready to learn in unfamiliar ways" indicate a positive assessment of the learners own abilities regarding learning. Items 236 and 229 both deal with making choices, and possibly also item 203 suggests decision making. Item 220 underlines readiness to learn in different ways, and similarly 205 suggests active gathering of L2 usage in order to advance learning. This group seems to be

characterised by qualities related to the individual’s sense of ability and self-reliance in learning, so I have labelled the group “Self-Reliance” to reflect this proactive independence.

6.2.4 Comparison with literature

Here I attempt to establish the construct validity of the questionnaire by comparing its factors with the groupings found in Cotterall’s (1995) work, hers being the most readily comparable approach available. In the Discussion Chapter (Section 7.4) I will discuss the questionnaire model with the wider literature of autonomy for language learning.

Factor Groupings	
Cotterall 1995	Present Research
A 1. Role of the teacher	4. Locus of Control
B 2. Role of feedback	4. Locus of Control?
C 3. Learner independence	6. Self-Reliance
D 4. Learner confidence in study ability	1. Linguistic Confidence
E 5. Experience of language learning	5. Metacognition
F (6. Approach to studying)	-
G -	2. Information Literacy
H -	3. Social Comparison

Table 6.7: Possible matches of questionnaire factor groupings with Cotterall (1995)

Cotterall’s (1995) study (see Section 2.5.1.3) produced six factor groupings.

Comparison of these groupings with those produced by the present research reveals a reasonably close match (see Table 6.7 above). Common areas are control (rows A and B), independence (row C), and learning confidence (row D). Cotterall’s “Experience of language learning” is, she maintains, similar to metacognition, which suggests a match with my Metacognition component. There is no match with “Approach to studying”, and in fact in a later paper (Cotterall 1999) she discards this grouping because it “was not considered to relate specifically to language learning” (Cotterall 1999: 498). This leaves the questionnaire’s categories 2 and 3 without

matches. I will look at each of Cotterall's five factors and compare them with the questionnaire's equivalents.

Category 4 Locus of Control is similar to Cotterall's (1995, 1999) "Role of the teacher" factor. Her items are specifically couched in terms of the role of the teacher rather than a wider spectrum of possible sources of perceived control. The items in the questionnaire's Locus of Control category do not address teacher role but are a more varied collection of control-related items. This difference may reflect differing approaches to item selection between those used for the present questionnaire and for Cotterall's. Cotterall selected all the 34 items herself (she does not give details of any selection process). The autonomy measuring questionnaire had the advantage of using a large initial selection of items (Cotterall 1995: 256) which were then reduced in number statistically using data from trials and not by using judgements of the researcher or experts. This process has, I feel, produced a more interesting "data-driven" collection of items for the autonomy measuring questionnaire which is less likely to reflect preconceptions. In Category 4 there are items which (like Cotterall's Role of the Teacher) accommodate possibly external loci (e.g. 189. I learn English because I have to) as well as items which can potentially reflect more internal attributions (e.g. 230. My way of learning will never change). This broadens the coverage of the control area from the narrow focus on the role of the teacher.

Factor 2 Role of feedback. Her respondents like to get feedback. The items seem to indicate dependence on external approval, a reliance on external evaluation rather than self-evaluation. This seems to suggest that her grouping is not finding out about autonomy in general but about her specific students. Perhaps this grouping could be labelled "Reflection on own performance", and then her respondents' answers can be

interpreted as evidence of a lack of this. Reflection would imply that this is a form of metacognition, making Cotterall's factor grouping closest to my Category 5 Metacognition.

The items in Cotterall's Factor 3 Learner independence, all show "characteristics of an active learner" (1995: 199), which is very much my interpretation of the autonomy measuring instrument's Category 6, Self-Reliance.

Cotterall's Factor 4 Learner confidence in study ability is only composed of two items which makes it impossible to produce a reliable identification of the underlying connection. However, both the items are worded such that confidence in one's ability is a possible interpretation: "I know how to study languages well" and "I know how to study other subjects well". Self-belief, and perhaps therefore confidence, is suggested by the final "well".

Factor 5 Experience of language learning. Again, this group is composed of only two items ("I have been successful in language learning in the past", and "I have my own ways of testing how much I have learned"), but Cotterall believes that they can be interpreted as showing that learners have "awareness about themselves, about language learning and about strategies" in other words metacognitive knowledge (1995: 201). The autonomy measuring instrument has ten items to cover this area, and is much more specific and more clearly identifiable with metacognition.

Social comparison and information literacy are not reported in Cotterall's research which raises questions about the selection process used for her items – only 34 were used compared with 256 for the present questionnaire's development. The comparison with Cotterall's groupings adds to the evidence for the autonomy

measuring questionnaire having construct validity in that it broadly agrees with Cotterall's findings.

In Section 6.3.2 I will use the factors which have been preliminarily established here to investigate which areas teachers use in estimates, and whether this is a balanced range of the components of autonomy. I will also use these components to evaluate the questionnaire's performance, and compare it with a detailed investigation of two learners from the ELTCS groups.

6.2.5 Comparison with Short List categories

This section looks at whether the questionnaire's factor analysis groupings were new or aligned closely with the criteria originally used for the item selection. In Table 6.8 the areas represented in the 50-item Short List are presented for comparison with the six factors emerging from the factor analysis of the combined standardised data for the 50 items.

Short List areas	*%	Factor analysis	**Ranking
Confidence	76	Linguistic Confidence	1
Attitudes to learning	70	Information Literacy	2
Control	62	Social Comparison	3
Strategies	54	Locus of Control	4
Metacognition	54	Metacognition	5
Motivation	44	Self-Reliance	6
Actions/Behaviours	32		
Responsibility	30		
Skills	28		
Social Interaction	18		

*proportion of total items which involve the area as percentage. **the strongest to weakest factors from 1-6

Table 6.8: Comparison of input and output groupings

Both are arranged in order of strength, i.e. for the Short List areas by the proportion of items covering that area, and for the factors by the strength of the factors from high to low (factor number). There are parallels between the two grouping schemes,

but there has been a re-sorting of items. As can be seen in Table 6.9 below, the items have not been organised by the factor analysis into the same groupings as before.

The Long List categories and how the items were judged to cover them (shown in Table 10.2) were based on the areas identified in the Literature Review (see Section 2.4) and on personal judgement and discussion with supervisors. Following this the Short List was based on a number of statistical analyses aimed at selecting those items with a stronger ability to produce a wider spread of responses. With the factor structure produced by factor analysis the remaining items have been grouped statistically without the necessity of the subjective judgement found in the first stages of item selection. The final groupings, as can be seen in Table 6.9 do not exactly match the initial ones. This indicates to me that it was correct not to rely on judgement alone in the item selection process, but to go through the stages from more subjective but also broad, to less subjective and more focused. Correlations between the factor analysis results and the literature are more suggestive for purposes of construct validity checking when the narrower final selection of items has been arrived at more objectively

Factor	Item	Attitudes to learning	Social Interaction	Motivation	Confidence	Responsibility	Actions/Behaviours	Strategies	Metacognition	Control	Skills
1	086. When I read an English text I need to understand every word in it	x			x			x	x	x	
	110. Every word is important for understanding a listening text	x			x			x		x	
	112. I worry if I don't understand everything when I listen	x			x			x		x	x
	147. I worry if I don't understand all the words in a text	x			x			x	x	x	
	150. I worry if I don't understand all the grammar in a text	x			x				x	x	
2	125. I change the way I write according to who will read it		x		x	x	x	x	x	x	x
	175. I look at causes and effects logically	x						x			
	196. I am confident I can learn English well	x		x	x				x	x	
	251. I use real English texts (i.e. not made for students) in my learning	x			x		x	x			
	252. I know how to find information in a library				x		x			x	x
	253. I know how to use English language reference books...				x		x			x	x
	254. I know the parts of a book (index, glossary, contents, chapters)										x
	256. I know how to find the information I need on the Internet	x			x		x	x			x
3	187. I think learning English is more difficult for me than for the average learner		x	x	x				x		
	193. The other students know English better than me		x	x	x				x	x	
	194. The other students are more confident than me at speaking English		x		x				x		
4	095. To read you must proceed word by word	x			x			x			x
	140. To remember vocabulary you need to be talented	x								x	
	189. I learn English because I have to	x		x		x			x	x	
	230. My way of learning will never change	x			x			x		x	

Factor	Item	Attitudes to learning	Social Interaction	Motivation	Confidence	Responsibility	Actions/Behaviours	Strategies	Metacognition	Control	Skills
5	046. I can describe the learning Strategies I use				x			x	x		
	049. I have changed the way I learn after thinking about it.	x		x	x	x	x	x	x	x	x
	109. I predict the content before I listen	x			x		x	x		x	x
	130. My writing is better now than it was a year ago			x	x				x		
	138. I know techniques to help me remember vocabulary				x	x		x	x	x	x
	142. I fix my problems in vocabulary			x	x	x		x	x	x	x
	212. I talk to others about how I feel about learning English	x	x		x		x	x	x		
	234. It is my job to check my work for mistakes	x		x	x	x	x				
	237. I am an active dynamic person				x		x		x		
	238. I choose the exercises I work on	x		x	x	x				x	
6	008. I am good at studying on my own	x	x	x	x						
	203. I organise my time for studying	x		x	x	x	x	x		x	
	205. I notice how other people use English		x					x			
	220. I am ready to learn in unfamiliar ways	x		x	x		x	x	x		
	229. I can choose the method of learning that suits me best	x			x	x		x	x	x	
	231. I can study independently	x	x	x	x		x	x	x	x	
	236. I am good at making choices			x	x				x	x	

1 Linguistic Confidence. 2 Information Literacy. 3 Social Comparison. 4 Locus of Control. 5 Metacognition. 6 Self-Reliance

Table 6.9: Comparison of items in Short List and factor analysis groupings

6.2.6 Instrument performance indications

In this section I will consider how the items and factor groupings in the questionnaire contributed to its overall result. This overall result will be expressed as a “Categories Average” which is a summary score calculated for each individual questionnaire respondent. It is found by summing the respondent’s scores for each scale and then dividing by the number of scales (six) to produce an average. As stated in Section 3.12.2 items were coded either negatively or positively. Examination of the final list

of items in the factor structure reveals that all items have a clear positive (autonomous) or negative (un-autonomous) value and that any ambiguous items have been removed in the course of the item selection procedure.

Table 6.10 below shows the significant correlations of the factor groupings (or “Categories”) with the Categories Average (using combined and standardised data from all respondent groups). This shows how the categories were predictive of the aggregate result. The Social Comparison category has the highest correlation to the Categories Average, and Linguistic Confidence the lowest.

Categories	Categories Average		N
	Pearson correlation	Sig. (1-tailed)	
Social Comparison	.751(**)	0	161
Locus of Control	.709(**)	0	161
Metacognition	.665(**)	0	161
Information Literacy	.661(**)	0	161
Self-Reliance	.569(**)	0	161
Linguistic Confidence	.376(**)	0	162

**Correlation is significant at the 0.01 level

Table 6.10: Significant correlations of categories to Categories Average in descending order (standardised pooled data)

This is an expression of the autonomy-related model embodied in the questionnaire as expressed by the sample. It shows that Social Comparison was the strongest indicator of overall score. If the instrument were measuring autonomous-learning level then this category would be the most important for a teacher or learner to consider. It may also be that this category has an underlying influence on all the other categories, i.e. it supports aspects of the autonomy construct present in the other categories, and hence appears in the statistics as that most significant to the overall autonomy performance. This suggests that the construct measured by this instrument tends to be related with positive social comparisons and less with the individual’s judgement of language skills. However, these results should be treated

G	Items	Categories Average		
		Pearson r	Sig. (2- tailed)	N
3	193. The other students know English better than me. (R)	.452**	0	161
3	187. I think learning English is more difficult for me than for the average learner. (R)	.400**	0	161
-	055. I know why I did well or did badly.	.381**	0	161
6	231. I can study independently.	.379**	0	161
6	229. I can choose the method of learning that suits me best.	.378**	0	161
-	195. I worry that other students will laugh at me when I speak English. (R)	.370**	0	161
4	189. I learn English because I have to. (R)	.363**	0	161
2	175. I look at causes and effects logically.	.356**	0	161
2	251. I use real English texts [...] in my learning.	.355**	0	161
2	196. I am confident I can learn English well.	.348**	0	161
3	194. The other students are more confident than me at speaking English. (R)	.338**	0	161
2	254. I know the parts of a book [...]	.308**	0	161
6	236. I am good at making choices.	.301**	0	161
5	234. It is my job to check my work for mistakes.	.293**	0	161
2	253. I know how to use English language reference books [...]	.293**	0	160
5	237. I am an active dynamic person.	.287**	0	161
5	046. I can describe the learning strategies I use.	.272**	0	162
5	142. I fix my problems in vocabulary.	.269**	0.001	162
4	246. Memorizing answers is the best way to learn. (R)	.268**	0.001	160
6	220. I am ready to learn in unfamiliar ways.	.267**	0.001	161
5	138. I know techniques to help me remember vocabulary.	.266**	0.001	162
6	205. I notice how other people use English.	.260**	0.001	161
4	095. To read you must proceed word by word. (R)	.257**	0.001	162
1	147. I worry if I don't understand all the words in a text. (R)	.239**	0.002	161
2	125. I change the way I write according to who will read it.	.236**	0.003	160
6	008. I am good at studying on my own.	.230**	0.003	162
1	112. I worry if I don't understand everything when I listen. (R)	.226**	0.004	162
6	203. I organise my time for studying.	.225**	0.004	161
4	230. My way of learning will never change. (R)	.224**	0.004	161
5	212. I talk to others about how I feel about learning English.	.215**	0.006	161
4	140. To remember vocabulary you need to be talented. (R)	.211**	0.007	162
5	049. I have changed the way I learn after thinking about it.	.199*	0.011	162
-	026. I feel lucky when I get good marks. (R)	.193*	0.014	161
1	110. Every word is important for understanding a listening text. (R)	.189*	0.016	162
2	256. I know how to find the information I need on the Internet.	.187*	0.018	160
5	130. My writing is better now than it was a year ago.	.178*	0.024	162
1	086. When I read an English text I need to understand every word in it. (R)	.166*	0.035	162
-	030. The teacher's job is to give me all the information. (R)	.165*	0.036	162
2	252. I know how to find information in a library.	.165*	0.036	161

*Correlation is significant at the 0.05 level. **Correlation is significant at the 0.01 level. G = Group
1. Linguistic Confidence. 2. Information Literacy. 3. Social Comparison. 5. Metacognition.
6. Self-Reliance

Table 6.11: Significant item Correlations with Categories Average in all data combined in descending order of Pearson's correlation

with caution as this result would also be found when the top categories are the highest-scoring due to having items which are too generally true of most of the

respondents in the sample, or which most students prefer to answer in a positive way (see Section 3.13.5.2).

Table 6.11 (above) shows all the individual items which correlated significantly with the Categories Average. These items can be seen as the most important ones in the questionnaire and as the most indicative of an overall “autonomy score” as indicated by the Categories Average figure.

The two highest items (193 and 187) are from the Social Comparison category.

When ranked in order of the size of correlation, in the top third of the items five are related to confidence (items: 193, 187, 195, 196, and 194), compared to four from Information Literacy, three from Self-Reliance, and one from Locus of Control.

Confidence is also highlighted by teacher estimates (see Section 6.3.2 for discussion of this).

6.2.7 Summary results for the full sample

As can be seen in Table 6.12 Linguistic Confidence has the second lowest average item score (57.72%) of all the categories, just behind Social Comparison (56.77%).

This suggests that confidence is a weakness for these respondents. The highest scoring category is Information Literacy with an average score of 74.19 per item.

Categories 1 (Linguistic Confidence) and 3 (Social Comparison) have the lowest average scores. Data of this type indicate how the results of the questionnaire could potentially be used as an indicator of strong or weak areas in a class or possibly in individual learners and so serve as a tool to support teachers in the development of autonomous learning. I would suggest that the aggregate Categories Average figure will be a very general indicator only as it will not indicate identifiable

Category and Items	Min	Max	Mean	Ave
1. Linguistic Confidence				57.72
086. When I read an English text I need to understand every word in it. (R)	0	100	61.20	
110. Every word is important for understanding a listening text. (R)	0	100	61.64	
112. I worry if I don't understand everything when I listen. (R)	0	100	54.56	
147. I worry if I don't understand all the words in a text. (R)	14.29	100	56.29	
150. I worry if I don't understand all the grammar in a text. (R)	14.29	100	54.92	
2. Information Literacy				74.19
125. I change the way I write according to who will read it.	0	100	70.10	
175. I look at causes and effects logically.	0	100	70.14	
196. I am confident I can learn English well.	0	100	78.23	
251. I use real English texts (i.e. not made for students) in my learning.	0	100	74.04	
252. I know how to find information in a library.	0	100	72.45	
253. I know how to use English language reference books [..]	20	100	74.59	
254. I know the parts of a book (index, glossary, contents, chapters).	0	100	75.94	
256. I know how to find the information I need on the Internet.	0	100	78.04	
3. Social Comparison				56.77
187. I think learning English is more difficult for me than for the average learner. (R)	0	100	64.24	
193. The other students know English better than me. (R)	0	100	52.80	
194. The other students are more confident than me at speaking English. (R)	0	100	53.26	
4. Locus of Control				65.51
095. To read you must proceed word by word. (R)	0	100	61.97	
140. To remember vocabulary you need to be talented. (R)	0	100	57.06	
189. I learn English because I have to. (R)	14.29	100	63.70	
230. My way of learning will never change. (R)	0	100	72.71	
246. Memorizing answers is the best way to learn. (R)	0	100	72.12	
5. Metacognition				70.17
046. I can describe the learning strategies I use.	0	100	66.70	
049. I have changed the way I learn after thinking about it.	0	100	69.70	
109. I predict the content before I listen.	0	100	68.97	
130. My writing is better now than it was a year ago.	0	100	76.00	
138. I know techniques to help me remember vocabulary.	0	100	68.60	
142. I fix my problems in vocabulary.	0	100	66.90	
212. I talk to others about how I feel about learning English.	0	100	72.66	
234. It is my job to check my work for mistakes.	0	100	72.25	
237. I am an active dynamic person.	0	100	69.56	
238. I choose the exercises I work on.	0	100	70.36	
6. Self-Reliance				72.73
008. I am good at studying on my own.	14.29	100	74.40	
203. I organise my time for studying.	0	100	70.08	
205. I notice how other people use English.	40	100	80.77	
220. I am ready to learn in unfamiliar ways.	0	100	74.09	
229. I can choose the method of learning that suits me best.	0	100	71.83	
231. I can study independently.	20	100	74.76	
236. I am good at making choices.	0	100	63.16	

R = Reverse coded

Table 6.12: Item scores (%) in each category (combined standardised data)

areas of strength or weakness, but can be a quick summary to be noted and then developed by considering the other categories and by considering whether it is an expected result or an indicator of something which the teacher is not aware of. An

example of its use is given in Section 6.3.5.2. The categories average is also a way of examining how the instrument performs, that is, examining the weightings of the different categories within the instrument's underlying construct. In order to understand how the questionnaire's results really relate to individual learners, and therefore to investigate whether the questionnaire can be used at that level, it is necessary to look at the smaller scale data.

6.3 *Smaller scale data*

6.3.1 Introduction

In this section I will continue to examine the questionnaire, now looking at how its data and construct correspond with smaller scale or more qualitative sources of data. The data used were gathered from the two ELTCS classes (20 students), an interview with one of the ELTCS teachers, and an interview with one of the students, therefore in this section the evidence will be of a different nature to the larger scale data analysed in the previous section. Conclusions will not go beyond the limits of the data but interesting questions can be prompted by the qualitative data here reviewed.

6.3.2 Correlating teacher estimates with the questionnaire

It had originally been the intention to compare large numbers of teacher estimates with the questionnaire data and so establish whether the instrument was giving similar results to those obtained from teachers. If correlations had been found then it would have suggested that the questionnaire could be used as a tool to substitute for teachers' estimates of autonomy.

At the smaller scale now being considered, correlating teacher estimates with the questionnaire data will not provide statistically significant generalisable results.

However, it is worthwhile and interesting to emulate the procedure at a micro-level because it was an opportunity to gather evidence of the types of issue which could be encountered when using the questionnaire in an authentic teaching environment.

As well as looking at the correlation between the total questionnaire score and the teacher estimates, it was also decided to look at the individual categories and see whether these correlated and, if so, which gave the closest match. This was extended to looking for correlations between the individual items and the teacher estimates.

The purpose of this was to continue the quality control of items in order to find indications of issues which would need to be addressed. Information from the ELTCS year-3 teacher's interview will also be examined to find if her verbal descriptions of her students shed light on her estimates.

The two groups of students have been described in Sections 5.2, 5.3, and 5.4. In both the ELTCS year-3 group and the ELTCS preessional group there were 10 students. I myself taught the latter class. This group did the questionnaire twice, once at the beginning and once at the end of the five-week preessional, though only 6 of the students were present for the second administration.

Both teachers (i.e. the ELTCS year-3 teacher and myself) provided estimates of their students' autonomy levels on a scale of 1 to 10. In the case of the ELTCS year-3 group the teacher gave her estimates at the same time as she was being interviewed by me after the questionnaire had been administered. In my case I made a set of estimates on the first day of the preessional course and also administered the questionnaire (this stage will be referred to as P1). The students completed the questionnaire a second time at the end of the course, and I made my second set of

estimates of their autonomy levels, now based on five weeks' experience of teaching them (this stage will be referred to as P2).

It was hoped to shed light on which areas teachers actually use in everyday situations to evaluate their students' autonomy, and therefore these three sets of estimates were made without guidelines or suggestions as to how to carry them out.

6.3.2.1 Categories level

Looking first at the categories, the significant correlations with teacher estimates are shown in Table 6.13. (As the data from the questionnaire are from a 7-point Likert scale but the data for the teacher estimates are from a 10-point ordinal scale it was necessary to use the non-parametric Spearman's r rather than Pearson's r .)

	Teacher Estimate		
	Spearman's Correlation	Sig. (1-tailed)	N
1. Linguistic Confidence	.523(*)	0.019	16
4. Locus of Control	.519(*)	0.02	16
Categories Average	.473(*)	0.037	15

*Correlation is significant at the 0.05 level (1-tailed). **Correlation is significant at the 0.01 level (1-tailed).

Table 6.13: Significant correlations between Teacher Estimate and the categories and Categories Average (ELTCS year-3 and ELTCS preessional P2)

Linguistic Confidence has the strongest correlation. It should be remembered that the teacher estimates are based on observation of the learners and the teacher's own judgement, which are both subjective and will almost inevitably favour some manifestations of autonomy more than others. It would not be surprising if those students with better English are be judged as more autonomous. This kind of overemphasis on one area can be expected to show in the correlations figures such as those in Table 6.13.

The “Categories Average” is an overall score given to each individual questionnaire respondent. It is calculated by adding the score for each category and then dividing by the number of categories (six) to produce an average. This produces a score which gives an equal weight to each category. When individual categories to Categories Average correlations are made within the questionnaire (Table 6.14) it is possible to compare the balance of components used by teachers with that predominating in the questionnaire.

Group	Categories	Categories Average	
		Pearson Correlation	Sig. (1-tailed)
3	Social Comparison	.829(**)	0
1	Linguistic Confidence	.804(**)	0
6	Self-Reliance	.596(**)	0.01
5	Metacognition	.583(*)	0.011

N=15. Correlation is significant at the 0.05 level (1-tailed). Correlation is significant at the 0.01 level (1-tailed).

Table 6.14: Significant correlations of categories to Categories Average arranged in descending order (groups ELTCS year-3 and ELTCS presessional P2)

There is a broader range of categories, including Linguistic Confidence as with the Teacher Estimate, but also having Social Comparison, Self-Reliance, and Metacognition. This is a similar pattern to that found with item correlations. In Table 6.16 below, only two items correlated significantly with the teacher estimates (110 and 150). When item to Categories Average correlations are calculated for the questionnaire (see Table 6.15 below) there are 23 items which significantly correlate, two of which are not in the six categories, but including the two items found to correlate significantly with Teacher Estimates. Information Literacy does not appear, and (as with the significant correlations with the items) Locus of Control is not present.

Group	Item	Categories Average	
		Pearson Correlation	Sig. (1-tailed)
3	193. The other students know English better than me. (R)	.803(**)	0
3	194. The other students are more confident than me at speaking English. (R)	.788(**)	0
1	112. I worry if I don't understand everything when I listen. (R)	.770(**)	0
6	231. I can study independently.	.713(**)	0.001
6	229. I can choose the method of learning that suits me best.	.690(**)	0.002
1	150. I worry if I don't understand all the grammar in a text. (R)	.678(**)	0.003
3	187. I think learning English is more difficult for me than for the average learner. (R)	.655(**)	0.004
6	008. I am good at studying on my own.	.642(**)	0.005
-	026. I feel lucky when I get good marks. (R)	.642(**)	0.005
2	256. I know how to find the information I need on the Internet.	.633(**)	0.006
5	130. My writing is better now than it was a year ago.	.595(**)	0.01
1	110. Every word is important for understanding a listening text. (R)	.572(*)	0.013
5	138. I know techniques to help me remember vocabulary.	.543(*)	0.018
1	086. When I read an English text I need to understand every word in it. (R)	.539(*)	0.019
6	220. I am ready to learn in unfamiliar ways	.523(*)	0.023
5	142. I fix my problems in vocabulary	.511(*)	0.026
-	017. If I must finish a job at a certain time I finish early	.510(*)	0.026
1	147. I worry if I don't understand all the words in a text. (R)	.507(*)	0.027
6	236. I am good at making choices.	-.484(*)	0.034
5	234. It is my job to check my work for mistakes	.481(*)	0.035
2	253. I know how to use English language reference books (encyclopedias, dictionaries, etc.).	.475(*)	0.037
6	205. I notice how other people use English	.468(*)	0.039
2	196. I am confident I can learn English well.	.467(*)	0.04

N=15. ** Correlation is significant at the 0.01 level (1-tailed). * Correlation is significant at the 0.05 level (1-tailed).

Table 6.15: Significant correlations of individual items to Categories Average arranged in descending order (groups ELTCS year-3 and ELTCS preessional P2)

6.3.2.2 Items level

I will now look at correlations at the items level. The data from the two groups ELTCS year-3 and ELTCS preessional P2 were combined to provide a larger set of data. It was expected that these groups would provide better correlations than ELTCS preessional P1 since the two teachers (myself and the year-3 teacher) had had longer to familiarise themselves with the students. Table 6.16 shows the significant correlations which were found between the average scores for items and our estimates, and also shows to which categories the items belonged.

Group	Items	Teacher Estimate	
		Spearman's Correlation	Sig. (2-tailed)
1	110. Every word is important for understanding a listening text. (R)	.658(**)	0.006
1	150. I worry if I don't understand all the grammar in a text. (R)	.601(*)	0.014

** Correlation is significant at the 0.01 level (1-tailed). * Correlation is significant at the 0.05 level (1-tailed). (R) = Reverse coded. N=16. Group 1 = Linguistic Confidence

Table 6.16: Significant correlations of items to Teacher Estimate using Spearman's r

Both of the items belong to the Linguistic Confidence group, which suggests that students' linguistic confidence has a part to play in the process which the year-3 teacher and myself went through to estimate their students' autonomy. This was also the indication of the analysis at the Categories level (see Table 6.14). It is possible that the learners with the best English are those who we score more highly because we have assumed that these learners will have the greater autonomy. If this is the case, then it suggests that we were overly influenced by a limited section of the autonomy spectrum. It therefore suggests that the questionnaire could have a place in enabling a more balanced measure of learner autonomy. However, there is also a correlation at the .05 level between Teacher Estimate and Categories Average, which suggests that we two teachers and the questionnaire are, for the two groups ELTCS year-3 and ELTCS preessional P2 (i.e. second estimates), producing similar overview judgments of the autonomy levels of the learners. This on a larger scale could amount to a correlation of the questionnaire data with teacher estimates, though at this smaller scale it is suggestive rather than a transferable conclusion.

In order to investigate further how the ELTCS year-3 teacher made her estimates I will look at the interview I carried out with her.

6.3.2.3 Interview

In the interview she often stated that she had some difficulties in assessing the students. For example (Lines 68-74):

now, [T1-St-c], I've got to try and make sure I've got the right name [unclear] oh yes [T1-St-c] yes she is, she's very kind of difficult to know because she's always smiling and pleasant ... she probably, I mean I don't really know because I don't really know her well enough

Of the same student she says "she's so inscrutable" (line 85) and "I can't really make up my mind about her (line 88). Speaking of another student she says:

I wouldn't be surprised if there were quite a few homeworks missing from him [T1-St-g] and things like that, although you know I would have to check that (Line 195-196)

About one particular student she says: "anyway, [T1-St-d] I should know quite well she's my tutee" (line 90), which suggests that she is not as familiar as this with the other students.

The teacher makes a point which is particularly important in larger classes "[T1-St-j] now difficult to analyse or assess her because she is so quiet" (line 299), and, in an email communication she writes "when there are 33 you can't know each one very well - so I may have got them wrong".

In view of these difficulties, the fully validated questionnaire could have a role to play in teacher training by raising their awareness of how they make their assessments and what other considerations they should be looking for. The issue also highlights that teachers can teach students for a substantial time and yet not be able to become familiar with them. This suggests that a valid and reliable instrument of the type being analysed here could have a use in helping teachers to know their students better than they would otherwise.

6.3.3 ELTCS year-3 teacher interview and confidence-related categories

In analysing the interview with the ELTCS year-3 teacher I thought I discerned two distinct uses of confidence-related concepts. The teacher hints at a distinction when she says “confident members of the group and confident language users in the group...” (lines 364-365), which implies that she sees a distinction between two types of confidence, one perhaps more general and the other more specifically relating to using language.

In order to investigate whether the teacher’s division between the confident and the confident language users is reflected in her interview comments, I have presented in Table 6.18 the two uses of confidence which occur in the interview. One type is confidence concerning relating to the others in the group which I have labelled “people confidence”; the second type is where the teacher talks about “self-belief” and “self-esteem” and is referring to the learner’s abilities in English, and this I have labelled “learning confidence”. I have marked whether her descriptions are predominantly positive or negative. These can be compared to her estimates, and the questionnaire results for Linguistic Confidence and Social Comparison (which involved confidence in its highest-loading item), and Categories Average, which are shown in Table 6.17 below.

The ELTCS year-3 teacher says in her interview (Lines 366-367) that:

... looking at the ones I put down low [i.e. in her estimates of autonomy level] ... I would say yes all of these I would say lack confidence, you know they are socially probably shy

Table 6.17 shows that she in fact gave good autonomy marks to some students who in the interview she said were lacking confidence (see transcript excerpts in Table

6.18). For example students T1-St-a and T1-St-d have the highest autonomy estimates (see Table 6.17), but do have negative reports about their people confidence from the teacher who describes student T1-St-d as “very shy”, but estimates her autonomy at 7.5 out of ten. However, they also have positive-sounding reports of their learning confidence; she for example describes student T1-St-d as having “self-belief in her abilities”, which would compensate for her shyness. The two students with the lowest autonomy estimates both have negative reports of their confidence: T1-St-e for learning confidence and T1-St-j for her people confidence, and neither of these are balanced by positive reports, which corresponds with the teacher giving them lower autonomy estimates. In Sections 6.3.2.1 and 6.3.2.2 it also emerged that the year 3 teacher’s estimates and mine correlated most closely with the Linguistic Confidence scale, both when viewing items and categories. This still seems to be the case when considering individual students, as can be seen in Table 6.17. The suggestion is that the teacher will base estimates of autonomy on a readily observable quality such as Linguistic Confidence. This is the kind of effect which the questionnaire with its autonomy-related construct can potentially highlight. In terms of analysing the questionnaire it indicates that the questionnaire potentially has advantages over teachers when estimating autonomy.

Student	Linguistic Confidence	Social Comparison	Categories Average	Confidence reported in ELTCS year-3 teacher interview		ELTCS year-3 teacher’s estimate
				People	Learning	
T1-St-j	48.57	66.67	63.69	Negative		3.50
T1-St-h	60.00	61.90	66.20	Negative?		6.00
T1-St-a	51.43	71.43	x	Negative	Positive	8.00
T1-St-f	62.86	71.43	66.34	Positive		6.00
T1-St-d	65.71	90.48	75.87	Negative	Positive	7.50
T1-St-e	40.00	38.10	61.27		Negative	2.50

Table 6.17: Comparison of ELTCS year-3 students’ questionnaire results with teacher estimates of autonomy and confidence

	People confidence	Learning-confidence
T1-St-a	right, starting with [T1-St-a], well the major thing to say about her is that she is a bit different from the rest of the group and in class is quiet and withdrawn (lines 1-3) (NEGATIVE)	her, you know, self-esteem is quite high I think (line 12) (POSITIVE)
T1-St-d	...she's just shy ... in terms of other people (lines 136-139) she's not an outgoing person (line 123) she's very shy and she's finding this quite hard in her, in her flat because it means that she's very shy to talk to the other people in her flat (lines 111-113) I think she comes apart in social situations so she, her presentation was awful it was the worst of the group (lines 109-111) (NEGATIVE)	[T1-St-d] has self-belief in her abilities she knows that she is quite good at English (line 136-137) although she knows that she is good, she's not good enough for herself (line 140) (POSITIVE)
T1-St-e	-not commented on	Then we come to [T1-St-e], now she's somebody who is hard-working and extremely weak and you know I mean there is this kind of sense of where is autonomy linked to confidence and ability because she's got very little confidence and very little ability but lots and lots of effort (lines 129-133) [little] self-belief in her abilities (line 136) she's got a little bit of self-belief in the sense that she knows that she can work hard (lines 153-154) (NEGATIVE)
T1-St-f	And then we have [T1-St-f] and yes she comes over as someone who is very together, very, yes quite confident socially and doesn't seem to be sort of you know addicted to her Chinese group (lines 163-166) (POSITIVE)	-not commented on
T1-St-h	whereas the others [i.e. the ones she rated more highly in autonomy]are yes probably quite confident although [T1-St-h] comes over as shy but she is, you know by dint of her you know actions she's obviously broken away from the Chinese group which is very very brave (lines 369-371) (NEGATIVE?)	-not commented on
T1-St-j	she is so quiet (line 300) (NEGATIVE)	-not commented on

Table 6.18: Confidence categories in ELTCS year-3 teacher's interview

6.3.4 Movement of teacher estimates

This section looks at some smaller scale data drawn from a sample which due to its size will not give statistically significant results. However, it illustrates in principle a way in which the teacher estimates and questionnaire results can be compared which reveals if they are converging over time, and is therefore a further assessment of the questionnaire, as convergence in a larger sample would indicate that the questionnaire was providing useful information more quickly than were teacher estimates.

In this analysis scatterplots were made showing each student's results from the six questionnaire categories and the Categories Average plotted against my estimates from the ELTCS preessional class. As there were two sets of teacher estimates, (i.e. beginning (P1) and end (P2) of the ELTCS preessional this resulted in 14 scatterplots which are shown below in Figure 10.1 to Figure 10.7 in Appendix 10.5. In each scatterplot a linear fit line has been plotted automatically by SPSS to show the trend of the data points. It can be seen that in all cases the fit line in the scatterplots has rotated anticlockwise from P1 to P2. This indicates a closer correlation between my estimates and the questionnaire data in the second round of questionnaire and estimates (P2) than in the first (P1). This is very much what would be expected when a teacher (in this case myself) has had some time to increase his or her knowledge of a class. The hypothesis suggested by these figures is that increased exposure to a group of students will increase the accuracy of estimates made about those students. This will result in movement of the estimates, and if that movement is towards the questionnaire result then the questionnaire has been shown to be faster at

finding a level than the teacher. If this result were achieved at a larger scale it would be a positive result for the questionnaire.

6.3.5 Two students

In this section I will explore the information I have gathered on two of the ELTCS students. The purpose of this is to look at the questionnaire data with reference to individual learners to gauge whether the different sources of data converged and to see what indications they provide on the validity and reliability of the questionnaire and its categories. At this small scale, dealing as it does with data from only two individuals, the purpose is not to establish statistically significant correlations with the questionnaire data. Rather, I will investigate how individual learners' questionnaire data reflect the indications from other sources to explore whether questions are raised about the suitability of the questionnaire for use with learners.

Two students are involved in this stage of the analysis, T1-St-b and P1-St-a. T1-St-b was in the ELTCS year-3 group (which I did not teach) and she was the only one of the group to come forward to be interviewed after the questionnaire administration. She is also described by the ELTCS year-3 teacher in her interview with me, which means that for this student I have more material than for any of the others. P1-St-a was one of my students in the preessional group. She completed both the initial administration (P1) and the end of course administration (P2) of the questionnaire. Table 6.19 below summarises the questionnaire results of these two students.

	ID Group Teacher Estimate	T1-St-b	P-St-a	
		EY3	P1	P2
		75.00	90.00	60.00
Linguistic Confidence		74.29	62.86	57.14
150. I worry if I don't understand all the grammar in a text. R		5	3	3
147. I worry if I don't understand all the words in a text. R		5	5	2
112. I worry if I don't understand everything when I listen. R		5	3	5
086. When I read an Eng' text I need to und' every word in it. R		6	5	5
110. Every word is important for understanding a listening text. R		5	6	5
Information Literacy		78.57	89.29	69.64
254. I know the parts of a book [..].		6	7	5
253. I know how to use English language reference books [..].		5	3	5
256. I know how to find the information I need on the Internet.		5	7	5
252. I know how to find information in a library.		5	6	5
251. I use real English texts [..] in my learning.		7	6	5
175. I look at causes and effects logically.		5	7	2
196. I am confident I can learn English well.		6	7	7
125. I change the way I write according to who will read it.		5	7	5
Social Comparison		71.43	90.48	66.67
194. The other students are more confident [..] at speaking Eng'. R		5	7	4
193. The other students know English better than me. R		5	5	3
187. [..] learning Eng' is more diffi' for me than for the ave' l'rner. R		5	7	7
Locus of Control		68.57	77.14	82.86
230. My way of learning will never change. R		5	7	7
140. To remember vocabulary you need to be talented. R		4	7	7
246. Memorizing answers is the best way to learn. R		6	7	7
189. I learn English because I have to. R		4	1	1
095. To read you must proceed word by word. R		5	5	7
Metacognition		74.29	60.00	52.86
237. I am an active dynamic person.		6	6	2
234. It is my job to check my work for mistakes.		6	7	5
212. I talk to others about how I feel about learning English.		4	7	1
138. I know techniques to help me remember vocabulary.		6	3	4
109. I predict the content before I listen.		5	3	6
049. I have changed the way I learn after thinking about it.		5	7	2
046. I can describe the learning strategies I use.		5	3	5
142. I fix my problems in vocabulary.		5	2	4
130. My writing is better now than it was a year ago.		6	3	5
238. I choose the exercises I work on.		4	1	3
Self-Reliance		75.51	79.59	51.02
231. I can study independently.		5	6	4
236. I am good at making choices.		5	3	3
229. I can choose the method of learning that suits me best.		6	7	4
220. I am ready to learn in unfamiliar ways.		4	7	2
205. I notice how other people use English.		6	7	7
203. I organise my time for studying.		5	3	2
008. I am good at studying on my own.		6	6	3
Categories Average		73.78	76.56	59.75
Other items				
021. All lessons are equally valuable. R		4	1	1
023. Students should always do what their teacher says. R		4	5	7
026. I feel lucky when I get good marks. R		5	5	6
030. The teacher's job is to give me all the information. R		4	1	7
055. I know why I did well or did badly.		5	5	4
076. I rely on the teacher when learning. R		6	7	7
195. I worry that other students will laugh [..] when I speak Eng'. R		5	7	7
243. I decide what I need to read.		4	1	5

R= Reverse Coded. EY3= ELTCS year-3. Items scored 0-7. Category scores are %.

Table 6.19: Questionnaire results for students T1-St-b and P-St-a

6.3.5.1 Student T1-St-b

The table below (Table 6.20) shows this student's questionnaire scores in the context of her group. She is above the mean in all categories and has the highest score in two, Linguistic Confidence and Metacognition.

Categories	N	Minimum	Maximum	Mean	T1-St-b
Linguistic Confidence	10	40.00	74.29	60.29	74.29
Information Literacy	9	62.50	83.93	73.61	78.57
Social Comparison	10	38.10	90.48	62.38	71.43
Locus of Control	10	51.43	88.57	66.57	68.57
Metacognition	10	54.29	74.29	67.15	74.29
Self-Reliance	10	51.02	83.67	67.75	75.51
Categories Average	9	61.27	75.87	66.44	73.78

Table 6.20: Average questionnaire results for ELTCS year-3 students with T1-St-b's results

A prominent feature of my interview with this student was the importance she gave to what she herself referred to as confidence. I will discuss this and then move on to look at the other questionnaire categories.

6.3.5.1.1 Confidence

The word "confidence" occurs ten times in the 5,160 words she speaks, which is over five times the frequency of this word in spoken English according to the British National Corpus (Leech, Rayson & Wilson 2001). For example:

INTERVIEWER: Yes. Okay, 22 [i.e. item 187] "I think learning English is more difficult for me than for the average learner" disagree

INTERVIEWEE: I think I have the confidence

INTERVIEWER: Confidence, now that's interesting.

INTERVIEWEE: Yes

INTERVIEWER: Do you think confidence is...

INTERVIEWEE: I think it's really important (Lines 605-612)

This was particularly interesting as her first use of the word was unprompted. In fact, she uses the word more frequently in her interview than some other common terms associated with language learning, as shown in Table 6.21 below.

These data suggest that she was thinking of confidence as an important element in learning English. Her questionnaire results (in Table 6.19 above) seem to confirm her beliefs about confidence, as she has scored well in the two categories which appear to relate to confidence (i.e. Linguistic Confidence and Social Comparison), and she in fact has the highest score for Linguistic Confidence in her group. The interview finding independently lends support to the appearance of confidence in two of the questionnaire categories found by factor analysis.

Order	Frequency	Word
27	38	English
33	29	teacher
38	26	remember
40	25	word
46	20	learn
57	18	understand
59	18	write
81	12	vocabulary
82	12	words
90	11	study
91	10	confidence
93	10	exam
107	8	dictionary
109	8	homework
116	8	writing
123	7	grammar
126	7	listening
131	7	speak
132	7	students
140	6	learning
174	5	reading
242	3	exercises
328	2	motivation

Table 6.21: Common terms associated with learning English from T1-St-b's interview, with frequencies and frequency order

In the second instance of the word she attributes confidence at school to obtaining good marks:

INTERVIEWER: Where does it [i.e. confidence] come from? Does it come from being successful, or are you confident and then because you are confident you become successful?

INTERVIEWEE: I think if you are in school I think the confidence just comes from your score, the English mark (Lines 612-615)

This appears to be an idea which is close to the questionnaire category of Linguistic Confidence. However, she then goes on to describe a classmate:

INTERVIEWEE: ... one of my friends [T1-St-d] also came here [Warwick University] I think she she is a very strange person I think she lacks lots of confidence [unclear] she never talks to her flatmates and sometimes she is so hungry but she doesn't go downstairs to cook and we say why and she says lots of people in the kitchen

...

INTERVIEWER: So you think is very important to have this, do you think confidence with other people is the same as confidence about learning English?

INTERVIEWEE: Yes (Lines 615-623)

Here she is clearly referring to confidence with other people, which seems to equate to the idea of people confidence which was found in the ELTCS year-3 teacher's interview (see Section 6.3.3). It also appears to have a relation to the questionnaire Category 3 Social Comparison. This category appears to involve ideas of self-belief in relation to others, and possibly not being inhibited by worries about the judgement of others, and so confidence in relation to others.

In the above quote from the interview, student T1-St-b appears to feel that the "people confidence" is the same as the confidence gained from successful work, though she is not suggesting, I feel, that her friend is lacking in confidence because she is not getting good marks.

Her fourth and fifth uses of confidence:

INTERVIEWER: [item 193] "the other students know English better than me". I think

INTERVIEWEE: [unclear] difficult. This is still about confidence I think

INTERVIEWER: Yes

INTERVIEWEE: I think there must be some students better than me

INTERVIEWER: Well yes

INTERVIEWEE: I just want to have some confidence about myself

INTERVIEWER: I think I agree with you, it is confidence isn't it. In a way it doesn't matter whether they are better than you or not it's that you feel...

INTERVIEWEE: Yes, exactly (Lines 642-651)

Here she seems to be saying that for her it is important to maintain her inner sense of confidence and not think about whether other students are better at English than her. This suggests a conscious management of her reactions to her context to maintain her confidence and so increase her chances of performing well. This interpretation is supported by her next use of confidence:

I think no matter good or bad [unclear] they will speak English but I think I have the confidence. I don't care about the result I just want understand me (Lines 669-670)

She is here saying that concern for accuracy in speech should not inhibit one from speaking, even if other students might laugh. Her seventh mention of confidence is in the same vein:

I think I got the confidence and I speak out and I just try my best to explain my English to you (Lines 671-672)

In her eighth and ninth uses she goes back to her idea of the good students being the confident ones:

I think in my secondary school the good students always had confidence, only if you have confidence you can get the very high mark (Lines 679-680)

However, here she implies that confidence precedes good marks, which is apparently the opposite of what she said earlier in lines 614-615.

On the tenth occasion she uses confidence, she says:

when [unclear] my classmates also have sometimes the speaking [unclear] someone else will laugh at [unclear] I think most my classmates still have the confidence to keep them alert, don't stop (Lines 687-689)

She is making the point that her classmates, like her, have sufficiently high levels of confidence to overcome adverse reactions from their peers. This appears again to relate to Social Comparison. Her use of the term, however, appears to vary and she appears to contradict herself. This might be because she has not distinguished between types of confidence. My tentative interpretation would be that Social Comparison and Linguistic Confidence interact. If a fully validated version of the questionnaire were to put a learner at very different levels in the two confidence-related categories, then it could be a result of a recent positive experience in one of them; it could also be a sign that a learner is below the language level for the course resulting in a lower level of Linguistic Confidence. Alternatively, it could be a warning sign that a student is responding to the questionnaire in a biased way i.e. putting answers which appear to be the “best” ones rather than ones which are true of him or her. Thirdly, the answers could be ones which the respondent believes, wrongly, to be true of him- or herself. The questionnaire may not reveal some types of problem when the learner does not see them. For example, an inadequate ability to reflect (i.e. Metacognition) could result in answers in other categories being misleading. If a respondent gained high scores in the questions “I notice how other people use English” (item 205) and “I organise my time for studying” (item 203) due to a mistaken ability to assess themselves this could lead to a higher than warranted score for Self-Reliance. This suggests that a single questionnaire result viewed in isolation or out of context could be misleading and therefore that questionnaire results should be seen as indications which should be followed up with other questionnaire administrations and interpretation by a teacher who has some

knowledge of the individual concerned. This possible issue with the reliability of self-report questionnaires is not demonstrated here, but is the type of problem which should be investigated when developing an instrument. I will now look at the remaining questionnaire categories and examine what she said in the interview about her answers.

6.3.5.1.2 Information literacy

She scored 78.57% in this making it her highest category. I asked her, regarding item 253 “I know how to use English language reference books” (her questionnaire response was Agree), how she knew how to use reference books, and she said that students are taught this in primary school in China.

Item 252 “I know how to find information in a library” (Agree). This part of the interview is a little unclear, but she appears to say that she has used the university library a lot and has learned how to find things so that for her it is now easy, but this is not the case for everyone (lines 501-549). She had used libraries in China, but mainly for leisure, preferring to use Internet sites such as Wikipedia for reference. She only responded to the item with an “agree”, which either indicates her assessment of her ability or is a “safe” answering style.

Item 256 “I know how to find the information I need on the Internet” (Agree). She says “I think it’s everybody can find the information on the Internet” (line 592) and “I think it’s [unclear] an easy skill. Even my mum can use the Internet” (lines 594-595). I wondered whether this was a little complacent or naïve and asked her why she had only said agree to this, not strongly or very strongly agree and she thought she could change her answer to strongly agree.

Item 251 “I use real English texts (i.e. not made for students) in my learning” (Very strongly agree). She read English websites in China such as the BBC because she wanted to improve her English, and that this had originally been suggested by a teacher. She thinks that “everyone does that” (line 53). This is the first marked response in this category, indicating that she is more comfortable with this item.

Item 125 “I’ve changed the way I write according to who will read it” (Agree):

If I’d just writes the academic writing I really need to focus on the word vocabulary and the register ... but if I just write for my friends, just some MSN talk and then I think I can [unclear] whatever I want. (Lines 310-314)

She says that she thinks it’s “natural” (line 318) to do this, though she has again only gone as high as “agree” in her answer, which suggests either more doubt than she is saying, or a cautious response style.

Item 175 “I look at causes and effects logically” (Agree). Verbally, she essentially agrees, but adds that “Sometimes a little luck is good for me” (line 996).

Item 254 “I know the parts of a book (index, glossary, contents, chapters)”. She strongly agrees in her questionnaire response. She says that she was not taught how to do this, but found it out for herself. She says:

... the first time I need to write an assignment finish assignment so I went to the library and I found a lot of books about the topic ... but I don’t know how to, and I can’t read it in the library ... for the whole book so I just go back the book and find the index I want to find “education” so I find “E” ... and I just go back and I go back go back which page and find [unclear] me or not, I just go skim (Lines 574-584)

This hints at an independent and resourceful attitude, suggestive of autonomy, and it is supportive of the higher score which she obtained in this category.

6.3.5.1.3 Locus of Control

Item 95 “To read you must proceed word by word” (Disagree). She demonstrates that she understands the point of the question, for example she says “I think if in exams I must understand almost all and not every one word ... And if it’s everyday life I just need to understand the general idea (lines 57-60). She says that she does stop and check in a dictionary if “I encounter a word it is very familiar but I can’t remember the meaning so I check this with my e-dictionary” (lines 63-64). If she sees a word that she does not know she “Sometimes ignore oftentimes just to guess” (line 75).

Item 246 “Memorising answers is the best way to learn” (Strongly disagree). In the interview she seems to be clear in her mind about memorising answers not being a good learning method:

I think sometimes when we do the about mathematics [unclear] teacher will always tell you to learn a method not just for this question not just for this answer [unclear] you need to learn a method how to work out this mathematic problem and not just learn the answers (Lines 955-958)

She appears to be substituting one kind of memorisation for another, i.e. memorising a method rather than understanding or finding out for oneself. This illustrates to me that a short or closed answer to a single item cannot reliably probe all the details of a respondent’s thinking. However, I also feel that an appropriate degree of her thinking has been recorded for the purposes of a general overview of the respondent.

Item 140 “To remember vocabulary you need to be talented” (Neither agree nor disagree). She is non-committal because:

I think everyone can remember the vocabulary but some maybe someone is with the talented [unclear] remember it faster (Lines 435-436)

I think her response reflects accurately her feelings about the question, though another person with the same idea but a stricter interpretation of the question could have put one of the disagree responses (varying individual response styles like this are one of the features of questionnaires for small scale research).

Item 230 “My way of learning will never change” (Disagree). She is able to reflect on her past way of learning:

I think the way of learning in my primary school and now is really different I think everyone will change and we encounter new things [unclear] really one day we will find the best way of learning (Lines 937-939)

Again her questionnaire response is more muted than her ideas expressed in the interview might suggest.

6.3.5.1.4 Self-Reliance

Item 203 “I organise my time for studying” (Agree)

I think in China I think my only job is to study and I don't need to worry about what I'm going to eat today and what I'm going to worry about ... but I came here a lot of life problems I need to find what to eat today. I need to go to supermarkets, so it's a lot of problems so I need to organise my time for study (Lines 716-721)

She has been pushed into planning her time now, so perhaps her “agree” is not a sign of being a pro-active learner. On the other hand in this area she seems to be responding to a new situation for herself and has worked out a way to cope. This appears to show a degree of self-reliance.

Item 205 “I notice how other people use English” (Strongly agree).

Yes I think especially the native speaker and you need to try to learn their, copy their into-, intonations ... and the tone and the way of their speaking (Lines 1075-1078)

She is not just saying strongly agree to please me – her response style has been to agree – and she mentions the specific area of intonation which gives credence to her answer.

Item 220 “I’m ready to learn in unfamiliar ways” (Neither agree nor disagree). This item was her lowest scoring in this category. She was initially hesitant in the interview, reflecting her questionnaire response “I think it’s a little difficult with unfamiliar ways” (line 765), but when asked if she would try a new method that was recommended to her she said “I think I would try it, but if it’s not suitable for me I would just give it down” (lines 769-770). She tells a story about a method she tried in China called “Crazy English”, which involved shouting out in public and did not suit her at all so she stopped. Perhaps this experience made her cautious about trying new ways of learning and so prevented her from agreeing with the item. The contingencies of individual’s experience will influence their answers; however, this answer indicates a conservative learner.

Item 236 “I am good at making choices” (Agree). She was not sure how to understand the item, wondering whether spelling could be the kind of matter of choice intended by the question, and this doubt made her want to respond with “disagree”. This illustrates the difficulty of maintaining clarity across language and culture boundaries, especially in the context of a questionnaire. However, in the interview she said how she interpreted the item:

I think I can make the choices, but if it is the homework I have to do it but it is not homework I can decide which one I want to learn today and [unclear] tomorrow (Lines 753-755)

She was still doubtful about her answer and thought of changing it to disagree. I think, in the light of her above quote, that she was right to put agree for this item.

6.3.5.1.5 Metacognition

Her questionnaire score for this category was 74.29%, her third highest behind Information Literacy and Self-Reliance.

Item 138 “I know techniques to help me remember vocabulary” (Strongly agree).

She gives some examples in the interview, such as “... sometimes the teacher told us a lot of technical skills for example D-I-S means [unclear] is a negative way” (lines 386-387). So, she is remembering the techniques from school. She also thinks that it is helpful to try to guess new words. I asked her if I gave her a list of new words to learn, how would she go about it, and her reply was “I think I just study, try to recite”. She says that in China everyone learns English in this way. It involves writing a word repeatedly while chanting the spelling out loud.

Item 49 “I have changed the way I learn after thinking about it” (Agree).

I think sometimes my way just I think usual way it's really a little bit wasting time it takes a long time to take the notes or something and just to pick out the keywords it really takes time so sometimes I will learn from everyone else and to try to highlight something in the handouts (Lines 985-988)

She shows that she is reflecting on her way of learning, though it is not perhaps an example of a fundamental change in her way of learning; she did respond to the item to an appropriate degree with her unmarked “agree” response.

Item 212. “I talk to others about how I feel about learning English” (Neither agree nor disagree). She said “Sometimes if somebody ask me sometimes I will tell them,

but actually I never talk about this topic with someone else” (lines 1010-1011). This, she says, is the case generally, the only exceptions being when class activities call for it, an interview with other students for example. She has responded accurately, in that if she was to decide herself she would never talk about her feelings about learning English, though technically she does occasionally in class, so she was forced to answer with “neither agree nor disagree”. This is another case of an item being hard to answer for certain respondents. I think it is impossible to cater for all contexts at the same time within an item of manageable length. I think it is a limitation that has to be accepted for some students some of the time.

Item 142. “I fix my problems in vocabulary” (Agree). She says that “Every time I encounter a word I’m very familiar with but I can’t remember I will check it” (lines 471-472). She uses an electronic dictionary for this. This appears to be another example of quite a limited thought behind the questionnaire answer, though it is fully accurate.

Item 109 “I predict the content before I listen” (Agree). She mentions exams where the instructions are to read the questions before listening (lines 241-242), and she refers to television trailers (lines 244-246) which help her to predict the content.

Item 234 “It is my job to check my work for mistakes” (Strongly agree). She admits that “it’s really boring I think assignment will take a long time and when you finish you don’t want to touch it again I don’t want to open the Word again” (Lines 911-912). However, she says that she always does check her work. This suggests that she is a diligent student, perhaps with a developed sense of her own responsibility for her learning, though the source of the motivation for this does not come out in the interview.

Item 130 “My writing is better now than it was a year ago” (Strongly agree). When asked how she knew that she had improved she said that the marks her teachers are giving now are better. She also gave examples of how she judges for herself that her writing is better:

And now the thing [unclear] writing when I’m in China we always write 200 words or 500 words passages [unclear] just kind of writing academic writing but I really can [unclear] I will write for one day a whole day, this takes a long time. [unclear] I came here sometimes it’s a rush to write an essay before the deadline so sometimes I can write 2000 words in one day (Lines 348-352)

She is aware of how she feels, saying “I think the speed can also reflect something about my writing I think sometimes it’s natural to write some things academically” (Lines 356-357). She also reflects on her mental process “I don’t need to ... write the sentence in my head before I type it into the computer ... In China I think sometimes I need to think of the sentence” (Lines 359-362). Finally, she even mentions ownership of her writing “...it’s MY English ... It’s not THEIR English ... It’s not for my father mother or my English teacher...” (Lines 371-380). This all indicates that her “strongly agree” response is well founded and that her answer correctly reflects her level of metacognition regarding her writing.

Item 238 “I choose the exercises I work on” (Neither agree nor disagree). She talks about this in the interview as if she has interpreted the question as applying only to a situation where the teacher has given a choice of exercises. This was not what was intended by the item. I was looking for a higher level of control of learning, which could be shown by a motivated student working for him- or herself. I think the wording is too loose here; it should be “I choose some of the exercises I work on”, or “I choose exercises to work on in my own time”.

Item 237 “I am an active dynamic person” (Strongly agree). She interprets the item much as I had intended:

I think it means another English for study by myself or by my parents or by my teachers so I want to learn [unclear] so if there are no teacher no parents [unclear] I also will try to hold down the English book (Lines 1003-1005)

This question was included to find the more pro-active learners. In this case her interview and response confirm each other.

This category was her joint third ranking in terms of questionnaire score. It leaves me with the impression of an earnest and hardworking student. She is speaking in a foreign language for her, and one should not read too much into a sometimes very unclear interview. It does not seem to me that she is quite 74.29 per cent of the way to perfect metacognition, though she does have some nice insights, especially regarding her reflections on her writing such as her feeling of ownership. If I were to estimate her level I would say about 55 to 65 based on the interview. It is true that the questionnaire does have a quite narrow range of scores, and here I feel that a cautious multiple-choice answering style in the questionnaire combined with a remembered knowledge of classroom English lessons has resulted in her obtaining a misleading score. This also indicates an issue with the type and wording of the items, which is information which can contribute to a development cycle.

6.3.5.1.6 Autonomy

I avoided using the word “autonomy” both in the questionnaire and in the interview. She does not speak about autonomy, but she does refer to working on her own, or independently, especially as a contrast to group work:

I think lots of people like teamwork I think I prefer to do the homework or do the study by myself ... and lots of people [unclear] it's not easy to concentrate [unclear] it's only chat, chat about other things ... and although I know that teamwork is really good and it's a lot of different ideas and I think it is waste [unclear] time (Lines 865-878)

She seems to be aware of learning habits and learning styles. She often alludes to hard work and diligence, for example:

I think for example I try so I deserve to get the mark if I've not tried of course I can't get the good marks (Lines 992-995)

She seems to have rather conservative ideas about learning. She describes learning at school in China:

... you need to recite every word we can't guess any word so we checked every word and we know every word's meaning ... I think it's a good way ... You remember lots of English lesson it's easy for you to use (Lines 94-100)

However, she does strongly disagree that she relies on the teacher when learning (item 76). She makes a distinction between following the teacher and relying on the teacher, saying:

Yes I think I really follow the teacher's way, but I don't need to rely on the teacher ... I don't need to follow everything he told me (Lines 851-861)

The interview with the ELTCS year-3 teacher had a number of points which also came up in the interview with T1-St-b. She is hard working and diligent, but is not adapting as well as she could to the new environment in the UK. She prefers to work on her own and perhaps this is related to staying inside the "Chinese bubble", and also not finding her place in the academic field (both of which the ELTCS year-3 teacher mentions). The ELTCS year-3 teacher says that she does not think that T1-St-b is naturally autonomous, and this is not at odds with the impression I obtained from my interview with her. The questionnaire results are mostly in line with my

interviews with the ELTCS year-3 teacher and T1-St-b (though necessarily less rich in details).

6.3.5.2 Student P-St-a (Preessional)

I made observations of all the students during the ELTCS preessional course and wrote them down. I wrote about this student's confidence and how she had impressed me with her spoken English and "Western" air. This had led me to think of her as very autonomous. For the initial teacher estimate of her autonomy I gave her 9/10, but this had fallen to 6/10 when I wrote up her entry in my observations of the students in the fifth week of the course. I think the main reasons why I marked her down were her disappointing level of written English and her passive in-class behaviour; she did not seem to have a dynamic attitude to making progress, though she was always one of the more diligent and cooperative students. For the initial assessment I had only used her spoken English which seemed very fluent and natural, with appropriately-used colloquial phrases. Then over the course of the preessional I had to revise my judgement.

Self-Reliance went down from 79.59 to 51.02. She had large reductions in score in three items. The fall in item 229 suggests that she was no longer agreeing that she could choose the learning method that suited her best. In item 220, she went from very strongly agreeing that she could learn in unfamiliar ways, to strongly disagreeing, and in item 8 she no longer felt that she was good at studying on her own. It appears that she had reconsidered her self-evaluations, perhaps as a result of her experiences. This probably does not mean that she was less autonomous at the end of the preessional course. It probably indicates that she is more autonomous, in the sense that she has a better idea of herself and a better ability to assess her learning

at the time of the second administration (P2) than at the first (P1). Her P1 self-assessment was too high and was adjusted to a more realistic position in P2. This process could have affected her morale and caused her to lose confidence in herself as a learner, and in relation to her classmates. An increase in honesty for the second administration due to feeling more relaxed about the marks not contributing to her course grades would also produce this effect.

She also seems to be realising that she cannot rely on the teacher and will be expected to be able to put more emphasis on study on her own (items 23, 30, 229, and 231) and this has made her worried about her capabilities and how she compares to other students (items 8, 55, 193, 194, and 229). Her Locus of Control score has fallen from 77.14 to 54.29, suggesting that she feels less confident about her abilities to be in control of her learning, perhaps because she is becoming aware that more responsibility is being transferred from the teacher to her. (This kind of reorientation is part of the purpose of the preessional and is why I have interpreted the result in this way.)

Her Linguistic Confidence score was only down slightly, going from 62.86% at the start to 57.14% at the end. Her Social Comparison though went down more, from 90.48% to 66.67%. This at face value would be because she had changed her view of the other students, deciding they were more confident speakers and knew English better than her. Her Metacognition score fell from 60.00 to 52.86 which will probably be connected in some way to her loss of confidence. Her Information Literacy score has also dropped, from 89.29 to 69.64, possibly this is related to her experience in the preessional where information-finding tasks formed a quite demanding part of the course.

The score reductions after a period of new input suggest to me that she is rethinking her assessments of herself in the light of a period of change and challenges (see Section 7.4.5 in the Discussion Chapter). What she considered to be high abilities at the start of the course have become less adequate in her assessment by the time she has reached the end of the course. All this resulted in her Categories Average falling from 76.56 to 59.75.

At the start of the preessional I asked each of the students to write me a letter saying what they hoped to learn from the course. Student P-St-a wrote that she wanted to study grammar and vocabulary to improve her reading, and that she thought that speaking was not a problem for her. Verbally, she told me she thought I would be disappointed with her writing, because she felt herself to be weak in written expression. After seeing some of her writing I felt she was right in her assessment. In her end of course revision test she got 53/90, doing badly at combining sentences, using linking words, and using the passive, all of which are important for academic writing.

I am happier with the assessments of her which came at the end of the course, and these were very different from those at the beginning, including the questionnaire scores. However, the teacher estimate at P1 was 90% compared to the questionnaire result of 76.56%, and the teacher estimate at P2 was 60% compared to 59.75% from the questionnaire which shows that the teacher estimate has been adjusted down by 30% whereas the questionnaire result has moved down by the smaller figure of 16.81% suggesting that the questionnaire had given a better idea of the student than the initial teacher estimate.

The interview with T1-St-b clearly revealed more details and meant that speculation about what scores mean can be reduced. However, I feel that Student P-St-a's questionnaire results have led me to notice more and ask myself more questions about this learner than I would have been able to without the questionnaire. My experience of using the questionnaire has therefore been a very useful and positive one. If I had continued to be her teacher I feel I would have been able to support her autonomous learning in a much more effective way due to the questionnaire data.

6.4 *Validity and reliability*

The purpose of this section is to summarise the data which have been presented to examine the construct validity of the questionnaire. The groupings which have emerged from factor analysis are not the same as the groupings by which items were initially chosen (Section 6.2.5). The questionnaire thus embodies a different picture of autonomous learning which has been substantially produced through factor analysis of empirically obtained data from a quite large sample.

The factor analysis produced six clear factors which were identified and comparison of these with related literature in Section 6.2.4 showed marked similarities. Section 6.2.5 showed that the groupings were not the same as the Short List's item area coverage. This suggests that these data have produced findings which are worthy of further consideration, and this will be addressed in the Discussion Chapter (Section 7.3).

As stated in Section 3.8.1 reliability is indicated when the scores produced by an instrument are longitudinally consistent with the sample's treatment (Dörnyei 2007: 50), and internal reliability is shown by the Cronbach's alpha of the factor groupings,

which should be above .7 and not below .6 (Dörnyei 2007: 207). According to Field (2005: 668) Cronbach's alpha figures are usually higher with increasing numbers of items in a scale. For the large scale data the internal reliability is demonstrated in Table 6.22 which shows the Cronbach's alpha figures for the scales, and it can be seen that they are all above .7 even when there are only three items in the scale.

At the smaller scale, data from individuals or small groups, such as those obtained from the ELTCS preessional group, can be examined for consistency with the "treatment" which the student has undergone. In the case of the ELTCS preessional group the treatment was the five-week preessional course.

Factor	Identification	Cronbach's alpha	Number of items in the scale
1	Linguistic Confidence	.712	5
2	Information Literacy	.749	8
3	Social Comparison	.718	3
4	Locus of Control	.642	5
5	Metacognition	.764	10
6	Self-Reliance	.709	7

Table 6.22: Internal reliability of scales

What emerged from my consideration of Student P-St-a in Section 6.3.5.2 was that the questionnaire reliability was dependent on the purposes to which it would be put. As I was not testing the learners and as autonomous learning is dynamic and variable the questionnaire results were not written in stone but demonstrated how changeable it is over a relatively short space of time. Formal summative testing of autonomy with its focus on reliability over extended periods appears, after this experience with the present questionnaire, to be unnecessary. I feel now that a lighter more nimble instrument could be more appropriate for a more "real-time" picture of the changes in autonomous learning.

6.5 Formative benefits of the questionnaire

In Table 6.23 below I present the feedback concerning how respondents felt about the questionnaire. The vast majority of the feedback concerned the questionnaire's length (for the Long List), or comments about individual items which I have not included here. The feedback indicates that a closed-item questionnaire of this type can help learners to reflect on their learning, and this clearly has benefits in relation to supporting learners with autonomous learning. The questionnaire will therefore be able to fulfil a formative function quite apart from any value it has as a measure of autonomy and as a source of information for teachers. The formative aspect will be particularly appropriate if the questionnaire is used by learners who are studying independently.

Format	Source	Comments
A		-
B	English Internet	"just did your survey! Really made me think as I am just wondering whether to start serious study of Japanese again. I tend to be plunged into language using situations so far above my ability level (at work) that it feels like whatever I do wouldn't show any results for a long time, so I need a lot of energy and consistency but lack it. The survey made me realise I am illiterate as regards libraries and reference works in Japanese. It was a stage in taking myself seriously as a J-learner again, so thank you!"
C		-
D	BNU	"it really helps me to think carefully about my own studying habits." "It is very detailed, and it makes me think over my learning style. I think it is very good." "The questionnaire is very detailed. I haven't think about some of them before. Thank you for you questions that made me understand myself and my study more. " "I think the questionnaire is very useful and helpful for me to rethink my learning methods, study attitude, and some of my believes." "very comprehensive; very carefully designed"
E		-
F		-

BNU=Beijing Normal University

Table 6.23: Respondent feedback from data gatherings

6.6 Conclusion

In this chapter I have presented the findings from the analysis of the questionnaire data, interviews with a teacher and a student, and observations of students. I have also considered the responses of learners to their experience of completing the questionnaire.

Analysis of the questionnaire data at the larger scale was used to look at the construct validity of the instrument using factor analysis to find the model embodied in the questionnaire items. This had encouraging results as a model (which was not the same as the areas used in the initial item selection) was found which compared favourably with Cotterall's (1995) model of autonomous learning. Cronbach's alpha for internal reliability of scales was also satisfactory.

Confidence emerged from Section 6.2.3.1 as an important aspect of the autonomy-related construct, and with this the emergence of the unexpected factor Social Comparison leads to the thought that autonomous learning can be dependent on a learner's morale or psychology, or perceptions of their place in the group, and that this could make autonomous learning highly variable in level over time. Both teachers (subjectively at a qualitative micro level without statistical generalisability) independently of the questionnaire have highlighted confidence as a key element in autonomous learning. These quantitative and qualitative data on individual students have suggested that a "nimble" questionnaire rather than more heavyweight instruments is more appropriate for practical purposes.

At the small scale, data analysis has suggested (see Section 6.3.2) that the teachers estimated the autonomy of their students mainly based on language ability and

confidence, but that the teacher estimates and questionnaire average are close. Some positive indications of potential correlations of the instrument have been presented and techniques used which show promise of being effective for investigating the questionnaire at a larger scale.

In the next chapter I will explore in greater depth the important issues arising here such as the use of teacher estimates in the research and the questionnaire's conceptual model. I will return to reconsider my original aims in the light of the research, and I will also look at how the questionnaire may be used. The issue of questionnaire translation in applied linguistics research will also be discussed.

7 DISCUSSION

7.1 Introduction

The chapter can be broadly divided into four parts. Firstly, in Section 7.2 I look at the use of teacher estimates in the validation of the questionnaire. Secondly, in Sections 7.3 and 7.4 I explore the questionnaire's conceptual model, compare it to the literature, and so examine the questionnaire's validity and theoretical reliability. Thirdly, in Section 7.5 I reassess the original aims of the research in the light of the data and the experience of using the questionnaire with learners. In Section 7.6 I look at how the present questionnaire may be used. Finally in Section 7.7 I consider the advantages and disadvantages of translating questionnaires.

7.2 Teacher estimates

As part of the validity check of the questionnaire I compared its results with teacher estimates of autonomy. This involved only two teachers (including myself) and so the evidence is, statistically, suggestive only. However, it is worth examining the principle of using teacher estimates to validate the questionnaire, and it is directly relevant to the aims of the present research which involve exploring the possibility of using a closed-item questionnaire to provide a measurement of autonomy.

For the two groups ELTCS year-3 and ELTCS preessional P2 there was a significant correlation ($p = < .05$) between Teacher Estimates and the Categories Average (see Section 6.3.2.1), which suggests that the questionnaire could match (these) teachers in producing a general overview grading of learners' autonomy levels. This was a promising initial indication.

Teacher estimates correlated most with the areas of Linguistic Confidence, Social Comparison, and Locus of Control (see Section 6.3.2). Within the questionnaire Social Comparison and Locus of Control are also the categories which correlate most with the Categories Average (see Table 6.7). This correlation is an indication of a successful modelling of teachers' concepts of autonomy, and can be seen as an additional form of validation. The questionnaire may be able to improve on teachers by producing reliable and valid findings earlier than can a teacher and potentially with a more detailed breakdown of the result into the categories.

The idea of assessing the questionnaire's validity by comparing its results with teacher estimates may be called into question since teacher estimates may not be acceptable as a valid standard for assessing autonomy and therefore they would not be appropriate for validating an autonomy measuring instrument. However, since the questionnaire is intended to complement or improve on the estimates made by teachers, correlation with teacher estimates is an important stage in establishing the questionnaire's functionality for this purpose.

If the questionnaire could be demonstrated as being a little more accurate than teachers, or a little quicker, then it would be of practical use for a teacher. This shows that it is not necessary to compare the questionnaire more directly (in some way) with autonomy for it to be a useful tool. The question is not whether the measure is accurately measuring autonomy, it is whether it is emulating teachers, and can therefore help teachers to know their students more quickly and support their autonomy or autonomous learning more efficiently. In fact the aim of the questionnaire research can be recast in terms of this function; rather than being understood in simplistic terms as involving an instrument to measure autonomy it

can be seen as a teacher-estimate emulator or a way of improving on and complementing estimates with some diagnostic capability to help teachers know students better (the latter is an issue which is mentioned by the ELTCS year-3 teacher in her interview – see Section 6.3.2.3).

7.3 The questionnaire's construct

7.3.1 Introduction

In this section I examine the construct embodied by the questionnaire by developing the picture so that in the process I can further probe its construct validity.

		1	2	3	4	5	6
1	Pearson Sig. (2-tailed)						
2	Pearson Sig. (2-tailed)	-.023 .764					
3	Pearson Sig. (2-tailed)	.346(**) .000	.326(**) .000				
4	Pearson Sig. (2-tailed)	.423(**) .000	.309(**) .000	.415(**) .000			
5	Pearson Sig. (2-tailed)	-.005 .949	.614(**) .000	.322(**) .000	.265(**) .000		
6	Pearson Sig. (2-tailed)	-.012 .880	.492(**) .000	.248(**) .001	.166(*) .030	.593(**) .000	
CA	Pearson Sig. (2-tailed)	.376(**) .000	.661(**) .000	.751(**) .000	.709(**) .000	.665(**) .000	.569(**) .000

** Correlation significant at 0.01 level (2-tailed). * Correlation significant at 0.05 level (2-tailed).

1. = Linguistic Confidence. 2. = Information Literacy. 3. = Social Comparison. 4. = Locus of Control
5. = Metacognition. 6. = Self-Reliance. CA = Categories Average

Table 7.1: Inter-Category correlations

The six categories found (Table 7.1) appear to fall into two broad areas, Technical (knowledge and skills), and Psychological (including affect, which can block or promote the actualisation of autonomy). The first can be seen as what the learner knows about learning, and so would be the potential for autonomy given the lack of any blocks to its actualisation. This area is composed of two factor groups, Metacognition and Information Literacy. The second area, Psychological, is

composed of the factors named Linguistic Confidence, Social Comparison, Locus of Control, and Self-Reliance.

In this interpretation of the picture emerging from the questionnaire, the technical area would be the knowledge and skills which are necessary for autonomy.

Information Literacy appears to be one area of knowledge of how to go about learning, while Metacognition is the reflection necessary to understand and internalise the knowledge about learning.

7.3.2 Technical factors

7.3.2.1 Information Literacy

The category was identified in Section 6.2.3.2 and is composed of the following items (in order of loading):

- 254. I know the parts of a book (index, glossary, contents, chapters).
- 253. I know how to use English language reference books (encyclopedias, dictionaries, etc.)
- 256. I know how to find the information I need on the Internet.
- 252. I know how to find information in a library.
- 251. I use real English texts (i.e. not made for students) in my learning.
- 175. I look at causes and effects logically.
- 196. I am confident I can learn English well.
- 125. I change the way I write according to who will read it.

Candy et al. (1994: 43) characterise information literacy as:

- knowledge of the major current resources available in at least one field of study
- ability to frame the searchable questions in at least one field of study
- ability to locate, evaluate, manage and use information in a range of contexts
- ability to retrieve information using our variety of media
- ability to decode information in a variety of forms: written, statistical, graphs, charts, diagrams and tables
- critical evaluation of information

Candy et al. did not obtain this characterisation from factor analysis of data, but from submissions from senior university personnel, employer and professional associations, course documentation and institutional publications, interviews with staff, students, and graduates in different disciplines in higher education in Australia. Their characterisation matches well with some of the items in the questionnaire category, covering much the same ground. Critical evaluation of information is not overtly stated in the questionnaire category, but two items, 125 and 175, do seem to embody this idea. Item 125, which addresses writing rather than finding information, can be seen as evidence of a critical skill, albeit more productive than receptive. Its interpretation in the context of information literacy is that the writer can be seen as mentally putting him or herself in the position of the reader and evaluating what is being written (or will have been written), and so in order to empathise it is necessary to be aware of how communication can be made effective. Item 175 fits in with the critical aspect of information literacy as it concerns thinking clearly and logically.

The questionnaire appears to be independently and empirically converging with the conclusions of Candy et al. with regard to the make-up of information literacy. It is the knowledge and ability to find and interpret information for one's own learning purposes. It is a form of knowledge or skill which can be learned and improved through practice or learner training.

7.3.2.2 Metacognition

This is the questionnaire Category 5, identified as "Metacognition" in Section 6.2.3.5, and it is composed of the items:

- 237. I am an active dynamic person.
- 234. It is my job to check my work for mistakes.

- 130. My writing is better now than it was a year ago
- 212. I talk to others about how I feel about learning English.
- 138. I know techniques to help me remember vocabulary.
- 109. I predict the content before I listen.
- 049. I have changed the way I learn after thinking about it.
- 046. I can describe the learning strategies I use.
- 142. I fix my problems in vocabulary.
- 130. My writing is better now than it was a year ago.
- 238. I choose the exercises I work on.

It may be seen as linked to psychology as it can be a personality trait to step back, but mainly it is the skill of reflecting. It is much emphasised as being a key area of autonomy (e.g. Cotterall 2009; Sinclair 2000). In the literature, metacognitive knowledge is generally categorised into three types (Wenden 1998: 518-519), person knowledge, task knowledge, and strategic knowledge (see Section 2.4.5). Person knowledge, i.e. the knowledge and beliefs learners have about themselves and their ability as learners (in general and for particular tasks) appears to be addressed by two items in this category, 237 and 130. Task knowledge, i.e. knowledge of the purpose and demands of the task, appears to be addressed by item 238. Strategic knowledge, i.e. the awareness of strategies and how and when to apply them, appears to be involved in five of the items: 212, 138, 109, 49, and 46. The remaining items also involve the use of metacognition: item 142 is person knowledge and also implies task and strategy knowledge; and item 234 implies a belief that the learner has sufficient task knowledge to be independent about self-correction, and person knowledge in that the learner is seen as responsible for the learning.

Category 5 therefore appears to correspond to the idea of metacognition to a large extent. This tends to support the idea that metacognition is a concept that reflects a class of beliefs or behaviour that is present in learners. Reciprocally, this means that the literature provides a level of support for the idea that the questionnaire is reliable in that it has independently grouped items which are also grouped in the literature.

This is the area of overlap between the technical area of skills and knowledge with the psychological area. It is the conscious awareness of skills, and outcomes, and strengths and weaknesses, and awareness of one's own psychology which permits the management of one's own learning, or self-direction.

According to Holec (1981: 3), to be autonomous the learner needs:

...to have, and to hold, the responsibility for all the decisions concerning all aspects of this learning, i.e.:

- determining the objectives;
- defining the contents and progressions;
- selecting methods and techniques to be used;
- monitoring the procedure of acquisition properly speaking (rhythm, time, place, etc.);
- evaluating what has been acquired.

This requires self-awareness combined with the knowledge and skills necessary.

Metacognition is the awareness of one's knowledge, skills, and potentials based on reflection on past performance combined with an assessment of one's present psychological states and the context in which one finds oneself. Metacognition, then, appears to be the keystone which holds together all the other areas and is an essential component in autonomy. Without it the other areas are merely mechanical and not directed with conscious understanding (Lai 2001); it is what makes the learner a responsible and active participant in his or her own learning.

7.3.2.3 Self-Reliance

This is the questionnaire Category 6, composed of the items:

- 231. I can study independently.
- 236. I am good at making choices.

- 229. I can choose the method of learning that suits me best.
- 220. I am ready to learn in unfamiliar ways.
- 205. I notice how other people use English.
- 203. I organise my time for studying.
- 008. I am good at studying on my own.

The category was identified as “Self-Reliance” in Section 6.2.3.6. There are no items which mention working with others or place the learner in a social setting, so this category seems to address independence rather than autonomy and interdependence.

This seems to have an equivalent in Candy et al.’s “sense of personal agency”, which has two sub-areas (Candy et al. 1994: 44):

- a positive concept of oneself as capable and autonomous
- self-organisation skills (time management, goal-setting etc.)

Self-concept, self-efficacy, and self-determination are combined with the skills that support and justify the individual’s sense of capability. Candy et al. are specifically addressing the idea of lifelong learning, which is a concept that comes to autonomy through humanistic psychology, notably Rogers (1969). To be a lifelong learner a person needs to have independence, creativity, and self-reliance.

7.3.3 Psychological Factors

7.3.3.1 Categories 1 and 3

In this section I will attempt to clarify the identification of both questionnaire Categories 1 and 3. Category 1 was labelled “Linguistic Confidence” in Section 6.2.3.1, and is composed of the following items:

- 150. I worry if I don’t understand all the grammar in a text.
- 147. I worry if I don’t understand all the words in a text.
- 112. I worry if I don’t understand everything when I listen.
- 086. When I read an English text I need to understand every word in it.

110. Every word is important for understanding a listening text.

Category 3 “Social Comparison” is composed of these items:

- 194. The other students are more confident than me at speaking English.
- 193. The other students know English better than me.
- 187. I think learning English is more difficult for me than for the average learner.

These two categories will be discussed together as they were both linked with concepts of confidence in Chapter 6.

I want to look at category 3 Social Comparison (and therefore also at Category 1 Linguistic Confidence) more closely to see if the two types of confidence which came out of the interview with the ELTCS year-3 teacher (see Section 6.3.3) match.

The analysis of the interview suggested that there were two kinds of confidence being talked about: (i) confidence about the ability with the language, and (ii) confidence regarding relations with other people. Do these two types of confidence correspond to the questionnaire Categories 1 and 3? Confidence type (i) in the interview appears to be a good match with the questionnaire’s Category 1 Linguistic Confidence, but a match is not safe between the interview type (ii) confidence and Category 3 Social Comparison. In the interview with the ELTCS year-3 teacher she spoke of behaviours which were introverted or extroverted, but Social Comparison was not about introversion or extroversion; it was about the confidence or insecurity which can result from making comparisons of oneself with peers. This could still involve an element of confidence, as favourable comparisons would lead to the learner having confidence and being less inhibited in learning in a class situation.

James (2001/1892) divided the self into two main components, the “Me” and the “I”. He had “Me” as, among other constituents, the social self (2001/1892: 46) compared

with “I” for the volitional self. This division is followed by Damon and Hart (1982) in their review of research into child psychological development which showed that a child gains a more sophisticated sense of a social self as he/she develops. The word “me” occurs five times among the full set of items in all the questionnaire categories (compared to “I” occurring 57 times). Of the five instances of “me”, three are in Category 3, which is also the only category to have more than one instance, suggesting that this category is the individual more objectified, that is, seen in a social context. This category has indications of a social-, comparative-, and confidence-related underlying connection.

Turning now to Category 1 “Linguistic Confidence” (often termed “language learning confidence”); its contribution to autonomy is much hinted at in the autonomy literature, confidence often being mentioned in passing in an incidental or peripheral way, but it is not systematically examined as a component of autonomy and it does not seem to have been fully explored in many models of autonomy (e.g. Oxford 1990; Cotterall 1995; Victori and Lockhart 1995; Wenden 1995; Murray 1999). Littlewood (1996), who sees it as one of the four main components of autonomy, does not discuss it in any depth. The connection between confidence and autonomy is often mentioned as being a necessary foundation for autonomy and Littlewood specifically includes it in his anatomy of autonomy (1996). He believes that actualised autonomy requires ability and willingness, the latter being composed of motivation and confidence, while ability is composed of knowledge and skills. The analysis of the categories in the questionnaire is very similar and it tends to support Littlewood’s analysis.

Cotterall (1995) identified one of the factors which she found in her research as “learner confidence in study ability”. It was composed of two items: “I know how to study languages well” and “I know how to study other subjects well” (see discussion in Section 2.5.1.3). Being composed of only two items, it is not a conclusive identification, and confidence is not overt in the wording of the items, but it does mean that at least one other researcher has identified a confidence factor which, while not corroborating the findings of the present research, does at least make it not entirely unprecedented.

Stanton (1988: 127) points out that “...if a person believes he is anxious or lacking self-confidence, or anything else, he is likely to behave as if that perception were true.” If they can be helped to improve their confidence they will feel “more competent to transcend the limits they have been placing on themselves” (1988: 131). However, it seems that for learning confidence to be able to enter the beneficial cycle there has to be a recognition on the part of the learner that it is their responsibility that they have succeeded, and that they do not see it as luck or the result of unstable, accidental or external influences (Dickinson 1995: 166). Deakin-Crick & Wilson (2005) maintain that confidence is required to accept the responsibility to learn and this confidence depends in large measure on the individual’s relationship with others in the learning community. The converse situation would lead to “Fragility and dependence” (Deakin-Crick & Wilson 2005: 372). The relationship with others brings us back to the idea of social comparison being related to confidence, a link which can be seen in the correlations in Table 7.1 above, and is also presented visually in Figure 7.3 (which can be found in Section 7.3.4 below).

Williams and Burden (1997: 97) review the multifaceted make-up of self-concept, and their description is summarised in Figure 7.1 below. From the perspective of self-concept, confidence could be an outcome of self-esteem, itself resulting from either perceived achievements or from favourable comparisons and feedback from others, and therefore social comparison will affect confidence. Failures are more likely to result in lack of confidence when learners are motivated by the (external) goal of gaining peer approval, rather than having a more internal focus on effort and strategy use (Dweck 1986: 1046) which is perceived as being their responsibility. This is in line with other research into the influence of the learner’s group on the individual’s autonomy (for example Chang 2007). If a learner perceives him/herself to be in harmony with the others in the learning environment this “promotes student involvement and activity while moderating anxiety and promoting self-confidence” (Clément et al 1994: 442).

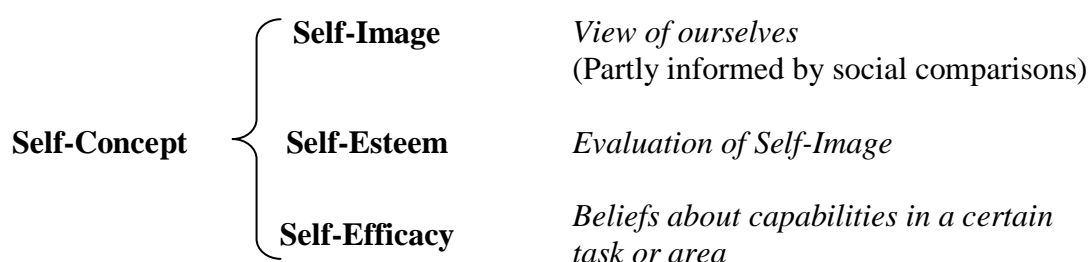


Figure 7.1: Self-Concept

The level of confidence one feels in self-comparisons with the group will affect the quantity and quality of interaction. This is important because, as Arthur (2001: 43) indicates:

...by taking part in pair or small group work learners can [...] develop the confidence to ‘let go’, to make mistakes in front of others, to take the initiative and to experiment with new language structures and hence experience a sense of enjoyment and achievement.

This will potentially lead to a virtuous cycle of confidence feeding in to achievement and in to confidence again. Without engaging with the group, as may be the case with learners who perceive themselves as comparing negatively with their peers, learners who give such comparisons too much importance may be held back. This can work in two ways, both of which will take away from the sense of control and therefore be harmful to learners' perceptions of autonomy. They are either (a) that the inhibitions tend to make the learner perceive the locus of causality as being external; or that (b) these negative external comparisons may result in an internal attribution of low ability which could lead to low self-esteem.

There are strong correlations between Locus of Control and both Social Comparison and Linguistic Confidence (see Table 7.1 and Figure 7.3). An autonomous learner will need to be able to control his/her learning (to the extent it is practicable). Social Comparison, which can be useful, could also be harmful to confidence and control as one's perceived status in relation to others is linked with one's feelings of self-worth and will have an effect on decision making. Learning involves accommodating new ideas into one's own matrix or schemata which therefore requires that an individual plays an active role in learning. If this part is insufficiently played and inhibited by social comparisons, then a social comparative insecurity or inhibition of confidence results in making the learner less well equipped for autonomy.

There is in the literature a link between confidence and motivation, for example Vandergrift (2005: 83) says that a lack of self-confidence and self-efficacy leads to lowered motivation to act, and Burt (2004: 7) says that self confidence in their ability and their perception of control of their learning play an important role in learners' motivation.

Motivation has not overtly figured in the questionnaire analysis, yet it should play a part in actualising autonomy, which is something which the questionnaire is concerned with measuring. In relation to autonomy it is intrinsic motivation which is relevant, as it is the intrinsic rewards of learning which make it autonomously sustainable (Ushioda 1996: 22). An attempt to account for the apparent absence of motivation in the questionnaire model can be found in 7.3.4 below.

7.3.3.2 Locus of Control

This is the questionnaire Category 4, composed of the items:

- 230. My way of learning will never change.
- 140. To remember vocabulary you need to be talented.
- 246. Memorizing answers is the best way to learn.
- 189. I learn English because I have to.
- 095. To read you must proceed word by word.

Locus of Control would appear to be an important area in autonomous learning (see Section 2.4.7). Autonomy involves making one's own choices and taking responsibility for managing one's own learning. An internal locus, where the learner feels able to control, is preferable for autonomy to an external locus, where the learner feels less able to influence their own learning.

Looking at each item it is clear that they do address an area which is at least closely comparable to the idea of locus of control (all the items were reverse coded for the questionnaire). For example, learners agreeing with items 95, 230, and 246 would appear to be non-exploratory; and learners agreeing with items 140 and 189 would appear to see learning as beyond their control; both of these are areas specifically mentioned by Williams and Burden (1997: 102) as indicative of locus of control. Research indicates a learner's locus of control or the way he or she makes

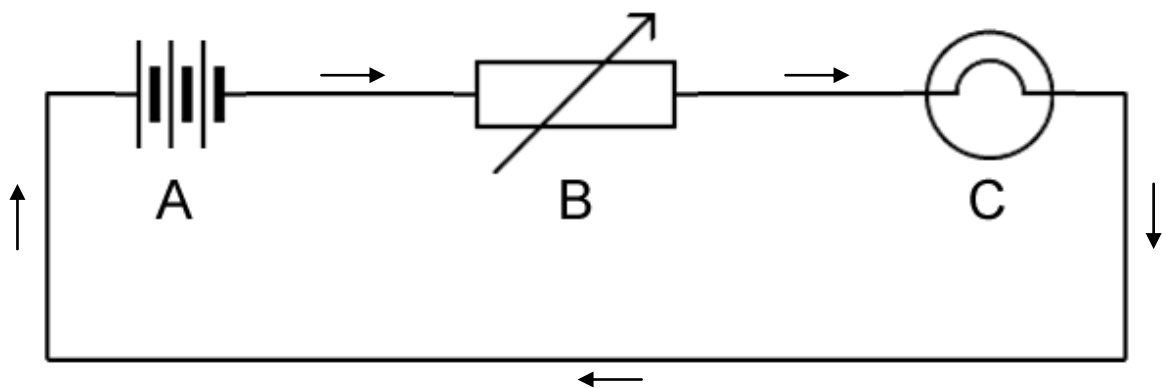
attributions can be improved by training, and it is therefore a component of autonomy which can be said to have levels. This has implications for teaching, and Williams and Burden (1997: 101-103) cite studies which have reviewed ways of changing a learner's locus of control. It can be done by teaching learners to take control of their learning through, for example, practice in planning or finding and organising information.

7.3.4 A Model from the questionnaire

Autonomous learning, such as engaging in a learning task, requires a degree of learning skill and/or metacognitive knowledge. However, the presence of the technical ability sufficient for the task is not enough in itself to motivate the learner to act autonomously at any given time. There are helps and hindrances which filter or block the learner's potential, which are such psychological influences as the learner's confidence or sense of control or openness to independent work. The realisation of the task and actualisation of potential for autonomy may be stronger or weaker depending on how much potential there is and how much it is blocked or not. This is my rationalisation of the categories which emerged from the questionnaire factor analysis.

A general picture of autonomy can be represented by an electrical circuit with a battery, resistor, and bulb (see Figure 7.2 below). The potential comes from the battery, this is passed or blocked to varying degrees by the resistor which represents the psychological (motivation, confidence, etc.) and other influences which may interfere with the potential. The bulb represents the task, and the completion of the circuit represents autonomy with all the areas, necessarily, involved. The questionnaire probes the potential, the resistance and the task completion (though all

from the perspective of the respondent’s self-report, not by observation of the student) to give an indication of the learner’s level of autonomy. The completion of the “circuit” is here the feedback to the metacognitive knowledge, which can, as it were, charge the battery with new (process, metacognitive) knowledge, which would mean enhanced ability to overcome resistance, can make the light brighter, or can light a higher wattage bulb in the future, that is, the learner learns from the fact of having done a task which gives the learner enhanced abilities to complete more, or more challenging, tasks in the future.



A = *Battery/Potential*.
 B = *Resistor (or dimmer switch)/areas affecting potential*.
 C = *Bulb/Task*

Figure 7.2: Autonomy seen as an electric circuit

Teachers can promote autonomy by increasing the “power” in the battery (learner training), by reducing the resistance (for example by making the context more supportive of autonomy), and by having a “light bulb” or task that is not too high a “wattage” for the potential in the battery and the level of resistance, i.e. the teacher can grade the task appropriately for the level of development of the learner.

The circuit can be looked at in different ways, according to the point of view:

- as basically two areas, capacity (skills) and action (with feedback from action back to capacity)
- as three areas (potential (A), psychological boosts or blocks (B), and task (C)) with metacognition as the overview
- with metacognition as part of B
- with metacognition as part of A

Figure 7.3 below represents the relations between the categories found by factor analysis of the questionnaire data. The full questionnaire data were used for this model and it is an empirically-based picture of the inter-relationships as they were found. The model shows that Social Comparison, technical skills (Information Literacy), Metacognition, and control (Locus of Control) are key areas, having the most number of strong correlations.

Since this is a product of the questionnaire data it can be seen as a model of what the questionnaire is measuring. I have labelled this “questionnaire autonomy” to distinguish it from theoretical concepts of autonomy (e.g. autonomy as capacity, political autonomy, etc.), and it will need to be investigated with reference to the literature of autonomy to justify its claim to a connection with autonomy.

In this model Metacognition is most strongly correlated with Technical Skills or Information Literacy. Metacognition is one of the necessary areas, along with Technical Skills, which is advanced or retarded by more psychological areas such as Confidence (see Figure 7.2 above). Metacognition is the conscious reflection which allows autonomy to be intentional, and this intentionality is necessary for the learner to be able to take responsibility for his/her learning. I have already discussed metacognition in Section 7.3.2.2.

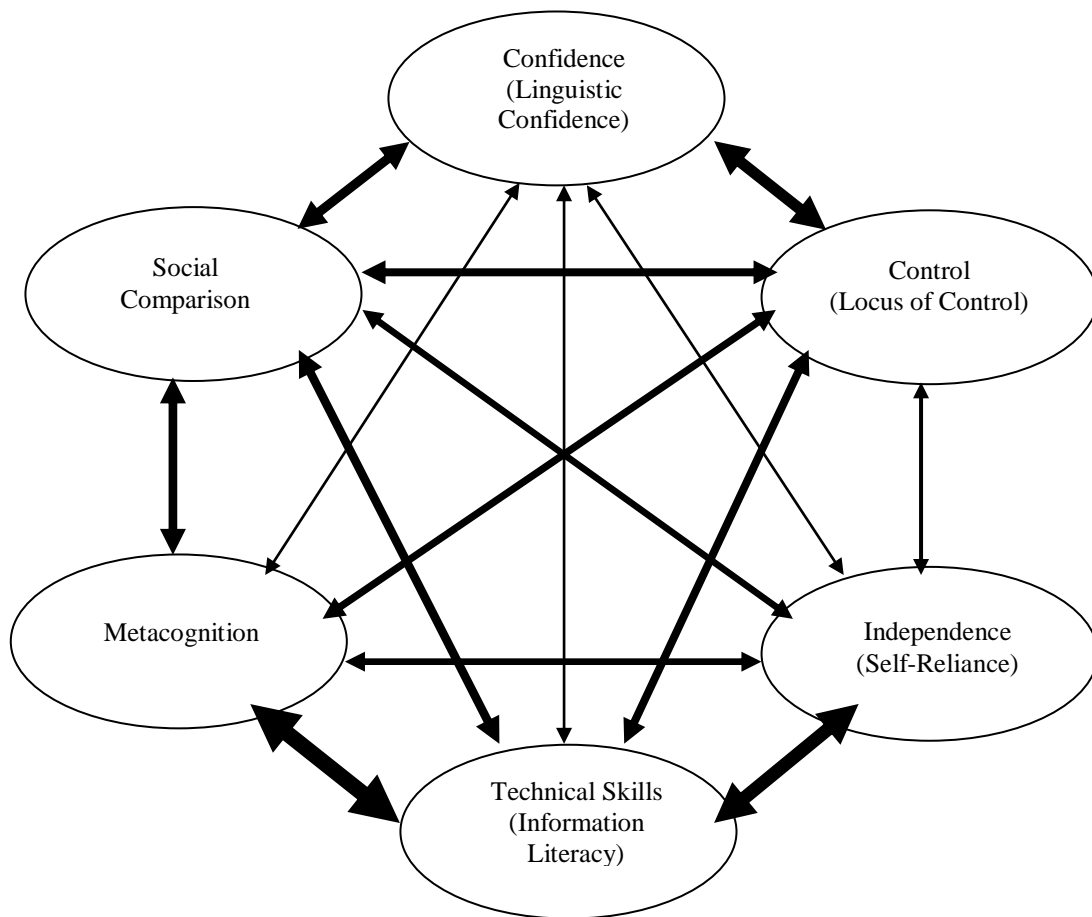


Figure 7.3: A model of Questionnaire Autonomy based on the questionnaire category correlations (N.B. Arrow thicknesses indicate strength of correlation)

In the model above the strength of correlation arrows indicate two areas of groupings bound together with strong correlations, one around Confidence and one around Technical Skills. The correlation arrows link all the areas in the model together, but the arrows indicate weaker correlations between these two nodes and stronger ones within them. The Confidence grouping is linked to Social Comparison and Control, while the Technical Skills area is linked to Metacognition and Independence. This reflects well the division between technical skills and more psychological areas in autonomy found in Littlewood's (1996) model of the components of autonomy (see Figure 7.4 below).

In this model, Littlewood (1996) sees two main components of autonomy, *Willingness* composed of the motivation and confidence necessary to take responsibility for a task, and *Ability* which covers the knowledge and skills required. These correspond well to the questionnaire's model: Confidence, Social Comparison, and Control correspond to Littlewood's *Willingness*, while Technical Skills, Metacognition, and Independence correspond to Littlewood's *Ability*.

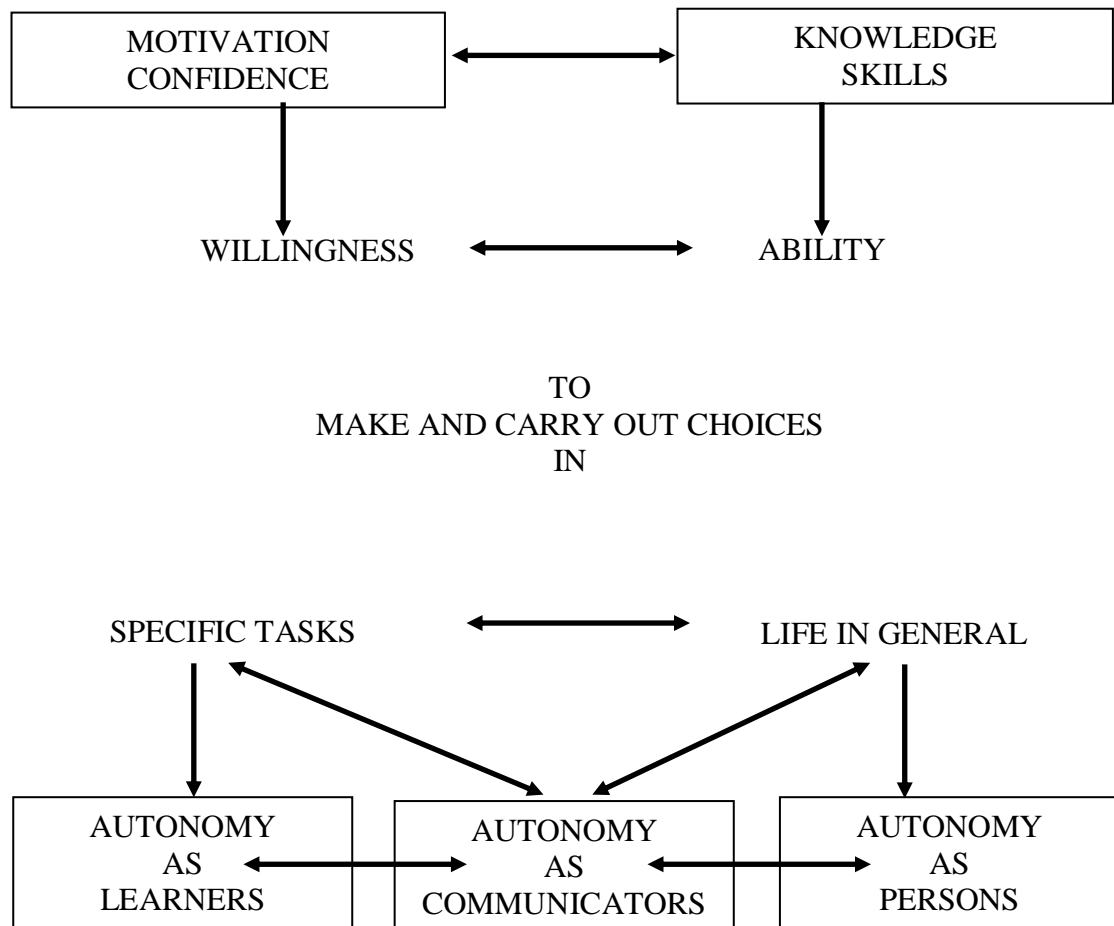


Figure 7.4: Littlewood's (1996: 430) model of the components and domains of autonomy

In this section an arrangement of the categories has been suggested, and how they may logically be seen to interact has been discussed. The suggested model has the potential to be used by teachers to understand (and evaluate) their students' autonomy. Areas which are low, comparatively, can be hypothesised as causing the

“bulb” to burn less brightly. In a classroom situation these areas could be targeted by the teacher in support of autonomous learning. Alternatively, the task can be adjusted to make it more suitable for the particular learners’ “circuit”.

A notable absence from the picture from the questionnaire data is the area of motivation. This is notable, especially since it was highlighted in the review of the literature (Section 2.4.7) as an important area, and as such was included in the areas to be covered by the Long List (Section 4.2.3). Motivation will be discussed in Section 7.3.5 which follows.

7.3.5 Autonomy and motivation

Motivation is a psychological area and is related to confidence, self-esteem, self-efficacy, control etc. (see Section 2.4.7), and is therefore in a sense hinted at if not actually overtly present. The lack of a clear motivation element could appear to be a limitation of this study. However, it is not due to a lack of motivation-related items in the initial 256 as there were for example the items:

- 226. I need tests to motivate me
- 227. I need praise to motivate me
- 228. I motivate myself
- 248. I am motivated by making progress in learning
- 188. I am motivated to learn English

Another 30 items can be interpreted as containing ideas relating to motivation, for example:

- 13. When I learn something new I feel satisfaction in myself
- 113. I look for opportunities to speak English outside class
- 78. I enjoy making my own choices about learning
- 151. I try to find ways of practising grammar outside class
- 225. Praise from the teacher is important to me

Most of these items were eliminated during the selection process (see Section 5.1), but even in the preliminary factor analyses there was no evidence of motivation emerging as a factor. Perhaps it is a “can’t see the wood for the trees” situation.

Fazey & Fazey (2001: 345-346) describe the key features of autonomous people as:

- Intrinsic motivation
- Perception as at the locus of control
- Take responsibility for their actions
- Have confidence in themselves

In the questionnaire model (Figure 7.3) categories are there which match with Fazey & Fazey’s description: Locus of Control could match with Fazey & Fazey’s locus of control and taking responsibility (since the two are logically connected); Linguistic Confidence could match with confidence; motivation does not have a direct match. The presence of these in both the literature and in the questionnaire model is promising. What is missing from the otherwise rather good fit between the questionnaire model and the literature is motivation. However, motivation as described in terms of Deci & Ryan’s Self-Determination Theory, or “SDT” (see Section 2.4.7) may be discernible. In SDT terms motivation involves the three broad areas of competence, relatedness, and self-determination.

These can be expressed respectively as skills, social connections, and intrinsic reasons for action. Matches for these appear to be present, or at least suggested, in the questionnaire model in the form of the categories named respectively Information Literacy/Metacognition (in combination), Social Comparison, and Locus of Control (see Table 7.2).

The questionnaire (model)	Fazey & Fazey (2001) (autonomy)	SDT (motivation)
Information Literacy		Competence (skill)
Metacognition		
Locus of Control	Perception as at the locus of control Take responsibility for their actions Intrinsic motivation	Self-determination (intrinsic reasons for action)
Social Comparison		Relatedness (social connection)
Linguistic Confidence	Have confidence in themselves	
Self-Reliance		

Table 7.2: Possible equivalences between the questionnaire, Fazey & Fazey, and SDT

In Littlewood's (1996) model of autonomy motivation is part of "Will"; the questionnaire, as with the Fazey & Fazey (2001) and Cotterall (1995) pictures of autonomy, is not explicit about motivation, but the SDT elements of motivation, i.e. competence, self-determination, and relatedness do appear to have matches in the questionnaire. Based on empirical findings, the questionnaire seems to be in accord with views in the literature (e.g. Dickinson 1995; Ushioda 1996; Benson 2001) that autonomy and motivation are closely related.

7.4 The questionnaire and its model's relation to themes in the autonomy literature

In the previous section I related the areas embodied in the questionnaire with the literature. In this section I will consider important areas in the literature and investigate whether the questionnaire's model has accounted for them.

7.4.1 Autonomy and responsibility

As seen in Section 2.4.6 responsibility is associated with autonomy by many authors (Holec 1981: 3; Boud 1988: 23; Dickinson 1987: 15; Little 1996: 203-204; Little et al. 2002: 17). It is therefore an area which would be expected to appear in a model of autonomy. Responsibility implies a sense of agency, seeing oneself as having some significant control. In the questionnaire model taking responsibility is represented by Locus of Control. In the model this area is most strongly linked to Linguistic Confidence, and then equally to Social Comparison, Metacognition, and Technical Skills (Figure 7.3). In the model, then, taking responsibility is associated with the supporting areas of Technical Skills and Confidence. Motivation has recognised links with the sense of being an agent and so to taking responsibility (Ushioda 2003; Spratt et al. 2002). This was discussed earlier in Section 7.3.5. In the area of responsibility then, the model compares well when viewed in relation to the literature.

7.4.2 Autonomy as capacity and behaviour

A key point in Holec's (1981) view of autonomy is that of "capacity". Autonomy is "the ability to take charge of one's own learning", and the skills that this involves including determining objectives, selecting methods, and evaluating what has been acquired). Littlewood (1996) analyses capacity into two distinct elements, ability and willingness (Littlewood 1996: 428). Ability is the technical skills, and willingness is the motivation and confidence. This is a picture which maps onto the questionnaire model very well. For a person to be successful in acting autonomously all the four components (knowledge, skills, motivation, and confidence) need to be present together (Littlewood 1996: 428).

A measure of autonomy should not be based on learner behaviour alone (see Section 2.4.4) because behaviour can be misinterpreted, and does not reliably indicate the underlying intentionality, and may not reveal the levels of knowledge, skills, motivation, and confidence. A self-report questionnaire should help the teacher to see beyond the behaviour, by probing it, not by direct observation but based on respondent perceptions which may be accurate but may also be misleading. Self-report and observation can both be seen as problematic ways of measuring autonomy. However, the questionnaire model which has emerged from the data collected by it has notable similarities with the literature, such as Littlewood's model, and therefore it is suggestive that the self-report format is producing data which are related to autonomy. This in turn indicates a degree of reliability in the questionnaire's ability to probe what underlies learners' behaviour.

7.4.3 Autonomy and social interaction

There are seven items in the Short List which refer to interactions with others, but they did not form a scale. This raises questions about both the usefulness and validity of the questionnaire in this area. If the questionnaire is not reflecting an area of current concern in the autonomy literature then the construct validity is called into question. If this area is not represented in the questionnaire then the questionnaire will not be useful for diagnosing problems in this area. However, the Social Comparison category does look at one aspect of social interaction, and the Metacognition category contains item 212 "I talk to others about how I feel about learning English". The Social Comparison category and item 212 both indicate levels of awareness of being "a participant in a social milieu" (Esch 2009: 33).

7.4.4 Autonomy is variable

There are four reasons which have been given for autonomy being seen as variable (see Sections 2.4.1 and 2.4.2):

- there are stages in the development of autonomy
- it is multidimensional
- it is sensitive to short term peripheral conditions
- it is sensitive to context

In one sense autonomy can be seen as variable because it has stages in development, expressed in the literature with band descriptors (Breen & Mann 1997: 143; Nunan 1997; Littlewood 1996). Different approaches are adopted: Breen & Mann (1997: 143) use levels of dependency; Nunan's (1997: 195) model is expressed in terms of levels of "learner action"; and Littlewood's (1996) levels of autonomy are defined by the choices which are possible for the learner. These are not practical measures of autonomy; they are theories about how autonomy may develop in different dimensions.

After having studied the data from the questionnaire the band descriptors seem to be in comparison more generalised and abstract. The questionnaire has provided a breakdown of specific areas which affect the autonomous learning of the individual, but has not suggested levels. The six components in the questionnaire model were seen (in the case of the ELTCS preessional class) to rise and fall individually over time which suggests that using broad phases to describe levels of autonomy will not be appropriate at the level of describing individual learners' autonomy in real classroom situations. I had seen these abstract models as a potential route into

operationalizing a measure of autonomy (see Section 2.4.2) but I now see, based on my experiences with learners and with the data provided by the questionnaire, that an individual learner's autonomy is too complex to be usefully expressed in these terms and that they have limited use for practical teaching purposes.

Autonomy can be seen as variable due to its multidimensional character (Benson 2001: 47). This is a concern which I raised in Section 2.4.1 and which implies that autonomy may not be accessible by means of a single simple instrument. This concern has been alleviated by:

- Specifying the purposes of the instrument (see Chapter 1)
- Specifying the spectrum of autonomy to be covered (see Section 2.4.1)
- Linking its validity to that of teacher estimates (see Sections 6.3.2 and 7.2).

With this clarification of the initial aims it has not been necessary to confront the problem of designing a universal autonomy measuring instrument which would overtly include all dimensions of autonomy.

In a third way, autonomy can be seen as variable due to its sensitivity to short term contingent peripheral conditions such as mood, environment etc. which affect individuals and their willingness to engage in autonomous learning tasks (Carr & Claxton 2002: 12; Sinclair 2009: 185). These influences can be very short term, such as hunger or tiredness, or they can be longer term, such as the environment of learning.

Longer term influences can be seen as forming part of the context, and this leads to the idea that autonomy is necessarily situated. This situated nature of autonomy would imply that it is inherently variable as it is not separable from all relevant

contextual (and contingent) influences. Seen in relation to this, a measure of autonomy would need to look either at autonomy over a long time period and not show its variability, such as Ravindran's (2000) method which, as described in Section 2.5.1.2, takes three years to complete; or be quick and simple enough to capture short term "snapshots" (Benson 2001: 54) which can then be interpreted in relation to other current information regarding the learner and context, and would permit repeated administrations without becoming onerous for learner or teacher. This type of data has great potential, for example, repeated applications of the instrument may reveal tendencies in individual students or classes which can then be targeted appropriately by the teacher.

Changes in the context (different tasks, times, places, etc.) will influence the level of autonomy, but the variability will depend on the transferability of autonomy (see Section 2.4.2) about which the literature is ambivalent. It is reasonable to speculate that an instrument such as a version of the one under investigation here may have a role to play in shedding light on this question (see Section 7.6).

In conclusion, a practical instrument to help teachers in class must show the variations in autonomy in sufficient detail. An instrument such as the one under investigation in the present research when fully developed should be appropriate for this due to its ease of application and its division of autonomy into categories, assuming that its potential in the present research can be corroborated in the field.

In the following section I will look at an aspect of variability which concerns how learners self-evaluate and how this relates to self-report questionnaires aimed at probing autonomy.

7.4.5 Self-evaluation in self-report autonomy questionnaires

It is not always clear what has caused a change in a questionnaire result: in item 237 (I am an active dynamic person) P-St-a's response changed from strongly agree to strongly disagree between the two administrations of the questionnaire. In this case there are three explanations which are the most probable.

Firstly (a), the questionnaire is, in part, a self-assessment of one's ability to self-assess, and a product of this reflexivity is that, should a student's ability to self-assess improve over time, s/he can be expected to give him/herself a lower, but a more accurate, assessment than previously. Alternatively (b), at the beginning of the course she felt ready for new learning experiences, but after five weeks' exposure to unfamiliar ways of learning (in a new country and in a new learning culture) she felt less enthusiastic and as a result of this she felt less motivation. Thirdly (c), she misread the question as a negative thus transposing strongly agree and strongly disagree.

If I were still her teacher I would talk to her; in a normal classroom situation this would be the response, and it would not be necessary to speculate as I have here. A use of the questionnaire would thus be to indicate possible problems and enable an appropriate response. This indicates a function of the questionnaire in facilitating a more targeted support of learners' autonomy.

This discussion raises two important points: firstly, the questionnaire will not be, and its results should not be treated as indicating, an absolute measure of autonomy; monitoring changes should be the primary purpose for using such a questionnaire. Secondly, the questionnaire will not in fact measure an abstract autonomy, but

aspects related to autonomous learning as found in the questionnaire model. The questionnaire looks at six areas, plus the overall score (Categories Average) and with the six areas it is possible to isolate and inspect areas individually.

Possibility (a) above was a change in the learner's ability to self-assess. A pattern which may illustrate whether this is occurring would be when the questionnaire scores start relatively high, then dip, and later start a rising trend, as illustrated in Figure 7.5 below.

The "U-shaped" pattern in the questionnaire results would be:

- Phase 1 (1 and 2 in Figure 7.5), steady or rising score
- Phase 2 (3 in Figure 7.5), dip in score due to reappraisal of own learning
- Phase 3 (4+ in Figure 7.5), score recovering as student adapts to new standards

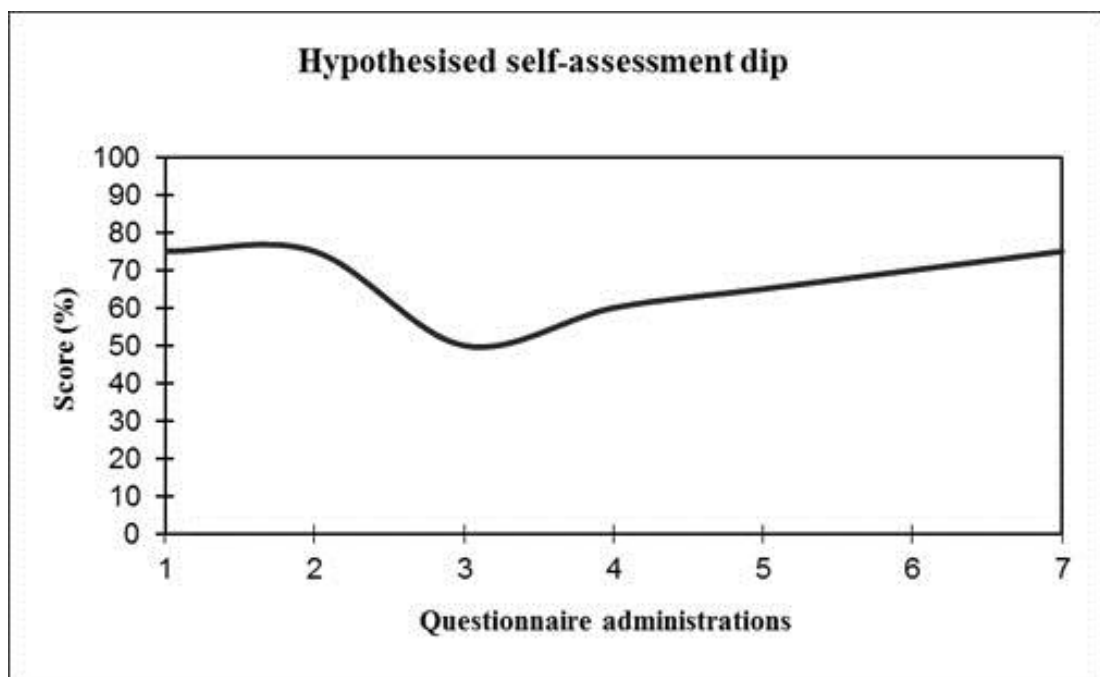


Figure 7.5: Example of the hypothesised dip in student questionnaire results

By this hypothesis, student P-St-a is entering Phase 2. A prediction from this would be that a third administration would show changes in her scores which fit in to the development curve. A further prediction would be that measurement taken prior to starting a course where autonomous learning is encouraged will usually be higher than later on in the course.

The concept of a “dip” as indicating progress does have a precedent in the literature; something analogous is proposed by Breen & Mann (1997: 142-144) concerning the disruption caused by the renegotiation of roles in a class when autonomy is introduced (see Section 2.4.2.1). In the wider language learning literature it also not unprecedented for a dip in performance to be a sign of progress, for example:

... the acquisition of forms such as ‘went’ follows a U-shaped pattern of development, with children first using it correctly (for example, ‘went’) and then incorrectly (for example, ‘goed’) before they finally once again produce the correct form (‘went’). (Ellis 1994: 77)

It indicates the use of the rule which is evidence of progress beyond the level of copying overheard utterances, but it would be judged as incorrect in an assessment.

The parallel with the “autonomy dip” is that something that has been learned (improved self-assessment) has had the effect of changing an indicator of progress so that it appears to have relapsed but a more important advance has in fact been achieved. This underlines a point made above that the questionnaire results should be seen in context with other indications (e.g. homework, participation etc.), and with communicating with the learner.

7.4.6 Summary

In this section I have looked at important areas in the literature and investigated whether the questionnaire and its model have accounted for them. The questionnaire model was investigated with reference to standards from the literature and was found to be acceptable. The important areas of responsibility, capacity, behaviour, and social interaction were all found to be present in the model. Littlewood's (1996) framework was shown to correspond well to the questionnaire model. The variability of autonomy which has been viewed as a threat to measurement was accounted for by clarifying that the questionnaire should not be used as a measure of absolute or abstract autonomy, but as a source of data which are related to autonomy for use in a primarily comparative measure for monitoring and facilitating autonomy in the classroom. Some limitations or features of self-report questionnaires were discussed and it was underlined that the questionnaire should be primarily comparative and used in conjunction with other sources of data, not least the learners themselves.

7.5 What the questionnaire is measuring

As a result of the research I have been led to question and clarify the original concepts of measurement and autonomy which were used in formulating the aims.

The two questions which I will explore in this section are:

1. Is the questionnaire measuring autonomy?
2. What concept of measuring is appropriate to the questionnaire?

Regarding the first question, Dam (2000: 48-49) was careful to specify that she was evaluating "autonomous learning" rather than autonomy. This distinction of a less abstract, more practical and more "in-class" concept fits well with my initial

motivation for this project, a practical tool to help teachers support their students. It seems clear now after engaging in the research that the questionnaire should be aimed at measuring a form or aspect of autonomy which has practical relevance.

The questionnaire model is clearly autonomy-related as shown by the exploration of its construct validity in previous sections. However, it would not be justifiable to maintain that the model is a new model of autonomy, and that the questionnaire is therefore measuring autonomy. It can be said that it is measuring a construct and components of that construct which are relevant to autonomous learning. This is a change of emphasis which clarifies the original expression of the purpose of the instrument by focussing on its practical function rather than on a more abstract notion of autonomy. The instrument then would be aimed at classroom use, and would need to have advantages over the present methods of measuring autonomy, but it would not need to be a measure of the abstract notion of autonomy, but of areas relevant to autonomous learning in class. This meant that it would then be possible to compare the questionnaire results with teacher estimates based on observations and knowledge of the learners rather than an ideal measure of abstract autonomy.

My research has tried to deal seriously with checking the construct validity of the questionnaire. The model does appear to match areas of the literature of autonomy, and therefore the questionnaire is measuring something related to autonomy, a construct which should be useful to teachers in the support of autonomous learning.

The analysis of the data has found that the questionnaire does have potential for helping teachers to judge their learners' autonomy. Cofield (2002: 40) has pointed out that:

... the inability to define a concept precisely does not have to act as a barrier to measuring it; indeed empirical research into a complex concept may lead in time to a more precise definition or definitions

The questionnaire is not an attempt to operationalise any one concept of autonomy, but the factors which have been found are variable and hence the construct which has been found is of the measurable kind, and has been given the name Questionnaire Autonomy (see 7.3.4). It emulates teacher estimates of autonomous learning. The model it presents appears to be an autonomy-like construct. It arose from a long list of autonomy-related statements.

Regarding the second question (What concept of measuring is appropriate to the questionnaire?), the idea of a “measure” has been clarified and become more nuanced in the course of the research. The questionnaire provides information on the six areas in its model which can be monitored separately rather than as one uninformative score which would be of limited practical use since autonomous learning is multifaceted. Secondly, it is important to use the data comparatively and relatively in a context and with a more holistic attitude to the individual learner. The measures produced by the questionnaire are of the nature of snapshots of a dynamic system rather than being similar to an IQ score. When the context changes it should not be assumed that this will have no influence on the individual’s autonomy. It is not expected that the questionnaire will give absolute measures of autonomy. However, it may be useful for comparative measures, i.e. within a group, and for groups and individuals longitudinally.

Benson says that snapshots are misleading (2001: 54). In the questionnaire this can be helped by longitudinal measurements, and by combining questionnaire indications with other knowledge which the teacher or researcher has about the respondent.

The components of autonomy may be stable or unstable in learners; if they are stable then a snapshot will be less limited. What will be important is that the snapshot is representative of that moment, and is not extended to be interpreted as a permanent or long term judgement of the person. An advantage of this questionnaire is that it is quick enough to provide many such snapshots to build up a more informative picture of a learner's development over time.

A norm-referenced rather than a criterion-referenced interpretation of the questionnaire scores is very much indicated (Messick 1975: 957), as it is important not to link the questionnaire scores to an independent standard, but look at them in comparison with others from the same or a comparable class or with the same student's previous results. The questionnaire provides useful information about an individual student if the scores are seen in context. Comparison with the respondent's previous results could potentially indicate when a student is not progressing well and act as a warning signal to the teacher who can then communicate with the individual learner and ascertain whether it is necessary to consider taking some remedial action.

The questionnaire should be treated not as a final assessment of a learner but as a way of raising questions to be asked of the learner. This means that the questionnaire can provide indications about how the learner is placed regarding their levels in the categories, but that these are better seen as initiating or contributing to a more holistic dialogue about the learner's autonomy.

The development and use of this closed response quantitative questionnaire has vividly illustrated to me how autonomy benefits from a social and interactive context. It is interesting to be shown this by a quantitative instrument, and it illustrates that there is a place for a quantitative instrument in the support of autonomy. Interacting with learners on a day-to-day basis in a normal class and in more detail in the one interview carried out, and then the interaction of these experiences with the questionnaire data has shown me that the questionnaire adds a useful dimension to my understanding of my students.

I have speculated about the reasons for changes in an individual learner's questionnaire results (see Section 6.3.5.2), but in a normal teaching context a teacher would be able to interact with the learner. This points to the importance of the questionnaire being used by a teacher who knows the class and is more aware therefore of what has been happening in class. It also points to the importance of knowing one's class, that one cannot support autonomous learning only in an abstract way but one has to be involved with the class and individual learners in order to be aware of their needs.

7.6 Suitable uses for the questionnaire

The use made of the questionnaire by a teacher would need to be similar to the use that is made of teacher estimates. By this I mean that if the questionnaire can be shown to be as valid as teacher estimates it should not be used to make decisions about teaching methods or contents which go beyond what teacher estimates may justifiably be used for.

A number of roles for the questionnaire present themselves. The instrument is quick and can be administered in class, so that teachers will be able to see with only a small investment of time whether their assumptions are correct. They may see that their students have more potential than they imagined, and may feel, as a result of having this evidence, more able to modify their teaching to help learners actualise their potentials for autonomous learning. Viewed in this light, the questionnaire is as much a tool for teacher support as it is for the learners' benefit. Eventually, it could be an important tool for the review of teacher estimates and provide empirical evidence which could help to influence institutional beliefs and practices which may have been blocking the promotion of autonomy.

The questionnaire could contribute to the understanding of the theory of autonomy, with repeated applications over long periods to establish which categories are quicker to change and which tend to be more stable. This will help with the understanding of how autonomy develops and so indicate ways to enhance its promotion. As Benson (2001: 68) says:

To date, however, research does not provide conclusive evidence on the mutability of individual variables in learning, their interrelationships, or the role of experience, training and self-control in change. Learner control over individual variables in language learning is therefore an important research area for the theory of autonomy

Responding to the questionnaire may in itself be formative as has been indicated by feedback received (see Section 6.5), i.e. the learners will be thinking about areas they may not have considered before, and this will potentially feed back into their reflections. The questionnaire results could be presented as a report with suggestions for how to enhance autonomous learning by category.

As the questionnaire may reveal comparative weaknesses in the respondents' component areas, this information could be used for diagnostic purposes to see whether learner training or other interventions are required. Appropriate steps can then be taken at individual or class level to help learners improve in the indicated areas and so enhance their autonomous learning ability. Take, for example, weakness in the area of confidence/social comparison. Teachers are able to influence the learners' confidence (Gan et al. 2004: 401), and it is widely seen as important (Ellis & Sinclair 1989:3; Oxford 1990: 1; Victori & Lockhart 1995: 223; Wenden 1995: 192). One way of doing this is for the teacher to arrange "study buddies" or peer tutoring schemes; Mynard & Almarzouqi (2006: 14) point out that peer tutoring programmes are beneficial as they lead to "making friends; building confidence and self-esteem; enhancing team-working skills" and other benefits. Helping others to learn by involvement in peer teaching or tandem learning can be beneficial to self-esteem and self-confidence, even for the more inhibited learners (Walker 2001: 92). Materials can be chosen or designed which support the confidence of struggling learners (Wagman 2005: 71).

The questionnaire may be able to indicate if a learner will perform better in a different context, for example self-access learning might be indicated for a low Social Comparison score. One of the advantages of self-access learning is that it allows those who are shy to work at their own pace without having to be judged publicly. If this interpretation is correct then SAC or independent work could be a boon to those whose questionnaire scores indicate that they are inhibited in relation to others, which may hold them back in class.

The questionnaire could be used to augment teacher evaluations or to enrich the information from learners' logbooks. Logbooks are not always clear, they depend on the level of the learner, and the learner may not automatically cover all the component areas.

The questionnaire could be a tool for teacher development. Chan (2003: 43) found among teachers "a strong preference for a relatively dominant teacher role" which suggested that teachers did not expect learners to take on an autonomous role.

Teachers did not believe that students were ready to accept responsibility for their learning, and it was also thought that learners saw it as the teacher's role to make the decisions about learning. In addition, teachers also thought that it saved time if the teacher made the decisions (Chan 2003: 49). Such attitudes sound very familiar and I have encountered them in my teaching in the Middle East. There is a clear mismatch between what teachers understand as autonomy and what they see as realistically achievable in their cultural and institutional contexts. If such views are as widespread as it appears, teacher estimates of their learners' autonomy can be expected to be relatively low. Davis (2003: 212) reports that "Findings suggest teachers' beliefs can shape both the quality of their interactions with students as well as the quality of their instruction". It may be that the questionnaire could be used to change teacher beliefs about their learners and so potentially improve the environment for autonomy in their classrooms.

The questionnaire was not intended, and is not designed, to be appropriate for use as a high stakes test, in contrast to Ravindran's (2000) CILL (see Section 2.5.1.2). The characteristics of the present questionnaire mean that it can be used in different ways from such instruments as CILL. The present questionnaire differs from the methods

adopted by Ravindran (2000), Lai (2001), and Dam (2000) in that it can be used within a pre-existing course without the need for a new curriculum. This is an advantage as in many institutions it will not be possible to introduce new courses or curricula. It is therefore possible for an individual teacher to use the questionnaire informally in his or her own classes. The questionnaire can be administered in a short time and can provide more immediate information than the long term methods and without the need for a well-trained team. The questionnaire could provide data, in a quick and economical way, which could support and complement teacher estimates. The questionnaire will also not depend on the commitment of dedicated and expert staff which is a weakness in many autonomy supporting schemes.

7.7 Translating questionnaires

In Section 3.13 I explained my reasons for translating the questionnaire and gave the translation procedure which was used. Checking the translation using the technique of back translation proved to be an important stage. In Table 7.3 I have classified the types of problems found in the back translation of the Long List from Chinese, and give some examples (for the full list of items and back translations see Appendix 10.1). Sometimes there is one clear source of error, in other cases there are combinations of problems, for example item 203 shows two problem areas, *Emphasis Added and Related Concept*.

Feedback suggests that the translation was a success; prior to translation there was a rate of 14.52 comprehension-related queries per 100 respondents which contrasts with the rate after translation of only 1.85 per 100 respondents. The fact that there are fewer queries strongly suggests that respondents were better able to understand the questionnaire. Translation clearly had benefits for the reliability of the instrument

and the quality of the data gathered compared to the untranslated instrument, and therefore the data indicate that translation should be considered for questionnaire-based research where the original questionnaire is not in the respondents' first language.

Problem	Original item	Back translation
False assumption	071. If I learn something well, it is because I <u>studied well</u> 004. Learning continues all your life	If I learn well, that is because I <u>work hard</u> . It is never too old to learn.
Superficial similarity	024. The student's job is to develop <u>as a person</u>	The task of the students is to develop <u>individually</u> .
Trans-position	220. I am ready to learn in unfamiliar ways	I have made preparations for learning by using unfamiliar methods.
Related concept	022. I know how to <u>study</u> 074. <u>Memorization</u> is the best way to learn 102. My <u>general knowledge</u> helps me to understand texts I read 113. I <u>look for</u> opportunities to speak English outside class 193. The other students <u>know</u> English better than me	I know how to <u>learn</u> . <u>Reciting</u> is the best learning method. My <u>basic common sense</u> helps me with the comprehension of the articles . I <u>find</u> opportunities to practice my oral English in the spare time. Other students <u>learn</u> English better than me.
Emphasis added	203. I organise my time for studying	I arrange my learning time <u>soundly</u> .
Under-specified	211. I avoid <u>situations where there is a chance of</u> making mistakes	I avoid making errors.
Over-specified	233. I know how to check my own <u>work</u> for mistakes	I know how to examine my <u>homework</u> and find the errors out.
Addition	219. I reflect on my learning	I <u>summarize</u> and reflect on my learning.

Table 7.3: Classification of translation problems in the Long List with examples

Efforts were made to avoid cultural assumptions or impositions. These were considered, and the translation process into Chinese revealed some assumptions in items, which were then changed. At this level of attention to clarity across cultures the design of the questionnaire was quite painstaking.

There is another level to be considered, that of the appropriateness of Western interpretations of learner autonomy in non-Western contexts. The questionnaire is very much based on ideas of learner autonomy which are thought to originate in the West. No feedback on the questionnaire has been received from respondents which criticises it in this area, which is an indication that the questionnaire is not causing respondents, consciously, to feel imposed upon. The suggestion from this is that the questionnaire, though produced in the West, is acceptable in the East.

The translation issues involved in questionnaire development are seldom discussed in books on research methods in general (for example Nunan 1992; Punch 2005) or in volumes focusing on questionnaires (e.g. Dörnyei 2003). I had to look beyond the language teaching field to find most of the information I used for guidance (see Section 3.13).

Most papers in the language teaching-related field which use questionnaires in the research do not mention translation (e.g. Cotterall 1995), but a minority of them do refer to the translation of the instrument, and some describe the procedure and rationale. Mynard & Almarzouqi (2006: 22) briefly state in a footnote that “Tutees were given [...] this questionnaire with an Arabic translation on the reverse side” and Prodromou (1992: 43) mentions parenthetically that “The questionnaire was given to beginners in a Greek translation”. Li (2005: 3) describes a basic procedure:

For this study, SILL (Oxford, 1990) was translated into Chinese by the author first, and then was checked by a professional translator from the New Zealand Translation Centre

Spencer-Oatey & Xiong (2006: 41) and Tseng et al. (2006: 87-88) are moving towards committee translation (see Section 3.13.4.3) as both pieces of research use

several well-qualified individuals to translate and check the translation. Gan (2004: 407) describes using back translation (see Section 3.13.4.2), describing how:

The questionnaire was issued in Chinese. A preliminary version of the questionnaire items was initially formulated in English. These items were then professionally translated into Chinese. To further ensure the validity of the questionnaire, the Chinese version was translated back into English to see whether anything could be misinterpreted

Sometimes bilingual instruments are used (Spencer-Oatey & Xiong 2006; Mynard & Almarzouqi 2006) which have both languages present but this is not the straightforward choice it may appear to be as it can cause confusion (see Section 3.13.4.6).

The approaches used in the examples above are all well-intentioned but give the impression of being, to varying degrees, quite informal or based on common sense rather than being informed by an awareness of the issues and procedures which have been discussed earlier in this thesis (see Section 3.13). I maintain that there is a need to translate (and gave my reasons in Section 3.13.2). If translation is to be used it should be an integrated part of the plan for the research and not a late addition because it can be an involved process to organise and carry out, especially if the translation approach calls for a team to be gathered.

7.8 Conclusion

In this chapter I began by discussing and justifying the use of teacher estimates for validation. I then examined the questionnaire's construct and presented and probed this model in relation to the literature, which produced encouraging indications of close links. I have also developed and clarified the initial aims of this research with regard to the concepts of autonomy and measurement. I then discussed appropriate

uses for the questionnaire. Finally I assessed the advantages of translating the present questionnaire.

In the next chapter I will return to the research questions to consider whether they have been answered and I will assess the contribution made by this research. Limitations will be discussed and areas for further research will be considered.

8 CONCLUSIONS

8.1 Introduction

In this chapter I will answer the research questions posed in Section 3.3 and look at some of the implications of the research. I will consider what contribution has been made, will consider its limitations, and suggest further research which may develop what has been started here. Finally I will give my concluding remarks.

8.2 Answers to research questions

8.2.1 Research question 1

The first research question was:

Can a closed-item questionnaire be used to provide a practical and viable measure of autonomy? (*What are the issues involved and can they be overcome? Is there a place for quantitative techniques in the support of autonomy?*)

My findings suggested that the closed-item questionnaire in the present research cannot be claimed to be a measure of autonomy, and it appears doubtful that learner autonomy in an abstract sense can be measured at all; the questionnaire can, however, measure dimensions relevant to autonomous learning as it has shown promise in matching well with previous published work in the field, such as Littlewood's (1996) model of autonomy (see Section 7.3.4). My research suggests that autonomous learning has multiple dimensions which vary, and that indications about these can be obtained using a questionnaire which can thus serve a useful purpose in the classroom when the data it provides are viewed in context and in consultation with the learner.

Comparison of the limited data from teacher estimates with the questionnaire results has shown correlations between them (see Section 6.3). It has therefore been indicated in this research that the questionnaire is measuring a construct which is relevant to what the two teachers (myself and the year-3 ELTCS teacher) involved understand as autonomy, and therefore, seen in these terms, the questionnaire shows that a quantitative instrument does have some potential to measure a construct of autonomy which is recognisable to some teachers.

The variability of autonomy is reflected in the instrument's results (see for example Section 6.3.5.2) which means that the data the questionnaire provides regarding a learner must be seen as a snapshot which may soon be out of date; but such snapshots, when interpreted as being context-sensitive and appropriately treated as time-limited, have the potential to be of value to teachers for prompting and informing a current and relevant engagement with learners which can enhance the teacher's interaction with the learner and so enable the teacher to better support the learners' autonomy (see Research question 2 in Section 8.2.2 below). This shows that there is a place for quantitative techniques in the support of autonomy.

I would specify that the questionnaire provides information relevant to autonomous learning and cannot make any claim (and nor is it necessary to) to measure autonomy in a more abstract sense, such as described by Benson (2001: 44) as an "aspect of well-being deserving of protection in its own right".

I set out to see whether it was possible to have a measure of autonomy which was quick and practical. I now have discovered that expressing this aim with the indeterminate use of the word autonomy is too abstract to be useful for the practical instrument I (really) wanted. Autonomous learning is situated in a real-world context,

but “autonomy” is not, it is a “Platonic Idea”, a concept without substance until given a shape in a context such as a classroom where it can inform events, and be part of the interaction; an autonomy-related measuring instrument can have a role in this environment.

8.2.2 Research question 2

The second research question was:

What are the uses of a closed-item-questionnaire autonomy-measuring instrument? (*What are the advantages and disadvantages? If it works is it just a mass statistical tool or can it be useful for individuals?*)

My research has highlighted that the published techniques for measuring autonomy require much investment of time, reorganisation of teaching or curricula or the introduction of new courses designed around them. The present questionnaire is much quicker than the alternatives found in the literature and requires no reorganisation of teaching. It can be introduced easily into existing structures, and can even be used informally by individual teachers as it takes only a few minutes of class time. The literature describes methods which require weeks, or even years. The main competitor with the present questionnaire in terms of speed and low investment is teacher estimates and in this area the questionnaire, or questionnaires designed on the same quantitative principle, may (based on the limited data so far) have the potential to offer general indications about learners more immediately than the teacher estimates of their autonomy (see especially Sections 7.2 and 6.3.4). The comparison of the questionnaire with teacher estimates suggests that the questionnaire has advantages over initial teacher estimates, providing more detail of

category, and teacher estimates appeared to move towards the questionnaire results over time (see Section 6.3.4).

The questionnaire could also act as a standard procedure which would not have the variations which individual teacher judgements might bring. It could be a procedure which is repeatable for a teacher in his or her context and so would help teachers and learners to make internal comparisons to, for example, track changes.

A closed-item questionnaire could inform the dialogue between teacher and learner, raising questions and indicating possible problems which can then be followed up. Giving a score to variable dimensions can be seen as similar to asking a question with a time-limited answer, but an answer which is still important and useful for purposes of building and maintaining the quality of interaction between teacher and learner, and this would be particularly advantageous with a large class, as was indicated by the interview with the year-3 teacher (see Section 6.3.2.3).

A limitation which teachers must be aware of is that the questionnaire cannot ask and answer all the questions about an individual's autonomy and the questionnaire results should not be looked at in isolation. However, the data suggest that it does have the potential to be a useful addition to the autonomy-related teacher-learner dialogue (see Section 7.6).

Feedback from respondents (see Section 6.5) strongly suggests that the questionnaire has a potential to help individual learners reflect on their learning in ways they have not done before and it could therefore be useful as a formative aid in developing their autonomy. This indicates that a closed-item autonomy-measuring instrument will not only be a general statistical tool but could be useful for individuals.

The use of the questionnaire will not be to measure autonomy (still less to test it), but to gain a measure which is useful for supporting the development of learner autonomy. It can perhaps be compared to a small barometer which can be tapped periodically and which can add to the information you obtain by looking out of the window. It is a useful addition to everyday observations, and although it may not be a replacement for a supercomputer in a government meteorological office it is useful none the less, with the advantage of being within the means of an individual teacher or learner. This fits in well with my earlier argument that autonomy should not be measured in a top-down way, and thus a major advantage of the present questionnaire-based approach is that it is an instrument which learners can freely choose to use themselves to self-diagnose and so it can support autonomy from the bottom up.

8.3 Contributions to knowledge

I have divided the contributions to knowledge into three sections, dealing with learner autonomy (Section 8.3.1), innovative approaches to research (Section 8.3.2), and finally with the translation of questionnaires (Section 8.3.3).

8.3.1 In the area of language learner autonomy theory

8.3.1.1 Highlighting the issue of transferability

This thesis has highlighted transferability as an issue in autonomy (see Section 2.4.2) since authors have claimed on the one hand that autonomy is necessarily situated, context dependent, or task dependent (e.g. Dickinson 1987; Carr & Claxton 2002), but on the other hand it is claimed that a major part of the importance of autonomy is that it is a valuable transferable quality to have for life, which is expressed for

example in the Bergen Definition (Dam et al. 1990: 102). Thus a question that needs to be investigated is in what sense is autonomy transferable, which aspects are transferable, and what is the extent of a “situation”, i.e. when does it cease to be the same situation in autonomy terms and become another; is it as broad as life in general, or as broad as second language learning in general, or can it be as narrow as a single task type at a single time?

8.3.1.2 There is a place for snapshots

The word “snapshot” suggests something rapid and casual (Oxford Concise Dictionary, 1982) and instruments to measure autonomy will, says Benson, (2001: 54), provide only misleading snapshots. It would appear from this that such snapshots of autonomy would be unacceptable for use in promoting the development of learner autonomy. However, I have argued (see Section 7.4.4) that they are a highly important source of current information which can be used in the classroom provided that their limits are observed. A learner’s mood, recent experiences, and feelings, which all vary, will have an influence on the choice of response to questionnaire items. The questionnaire need not be criticised for reflecting these transient states, but the use that is made of the data and the conclusions drawn from them must be appropriate to the nature of what is being measured, in particular one snapshot should not be treated as a permanent and unchanging reading of an individual’s autonomy level. It is more akin to a share price than an IQ score. It is still useful information for a teacher to help them engage with learners’ autonomy and so can be added to the range of information which can be used by a teacher to support the learner.

8.3.1.3 Confidence

My research indicates that confidence is an area worthy of more explicit research regarding its influence on the development of autonomy. It is often mentioned in autonomy theory (e.g. Littlewood 1996), especially in relation to motivation.

Possibly it is its association with psychological or individualistic approaches to autonomy which makes it less interesting in the context of current thinking about autonomy. However, the category of Social Comparison which emerged from my research has links to confidence which indicate that confidence is socially situated and forms part of the social context relevant to learner autonomy (see Sections 6.2.3.3 and 7.3.3.1).

8.3.2 Innovative approach to researching autonomy

8.3.2.1 Critical reflexive mixed methods

The approach adopted in this thesis involved pursuing one line of research (the quantitative closed-item instrument) in order to investigate whether it was appropriate to researching autonomy. This was combined with the use of more qualitative research and reflection in a critical reflexive mixed methods approach by which I have explored how autonomy can be investigated, and in particular whether and how a quantitative questionnaire can be used to aid in supporting the development of learner autonomy.

This innovative approach enabled the exploration of the use of a quantitative approach to investigating autonomy without presupposing that such a positivist approach was an appropriate means of researching autonomy, but rather was a means for me to investigate for myself the limits of positivism in the field of promoting

learner autonomy. This critical reflexive approach has allowed me to experience in a practical, vivid, and first hand way, from the bottom up, what had hitherto seemed to me to be an abstract, theoretical, and ideologically-based division between inflexible proponents of conflicting paradigms. Adopting an open-minded critical reflexive approach has allowed me to see, in a way that has been very effective, that positivist tools used appropriately and sensitively can complement what should be a primarily social constructivist approach to developing language learner autonomy.

8.3.2.2 Comparison of questionnaire data with teacher estimates

This research has shown that in principle it is not necessary to compare questionnaire results against some objective standard of autonomy which does not exist. The lack of an alternative measure to compare against has been a major problem in attempts to establish the construct validity of measurements of autonomy in the past, and I have highlighted this in relation to Ravindran's (see Section 2.5.1.2) and Lai's (see Section 2.5.1.4) schemes. In my method the idea of comparing with the current de facto method (i.e. teacher estimates) has been proposed as it is more relevant to the realities of actual class teaching, and because estimation is the existing method of measuring autonomy in class, and it is thus appropriate to compare the instrument with estimates rather than with a (non-existent) "objective" measure.

It is not necessary to prove that the other measure (in this case teacher estimates) is actually measuring autonomy, only that it is the accepted way, even if it is accepted only for want of a better method. Since I was looking for an instrument for practical use I was able to justify using comparison with teacher estimates of autonomy in the classroom and did not need to find an objective measure of abstract autonomy. The question is not whether the measure is accurately measuring autonomy, it is whether

it is accurately emulating teachers, and can therefore help teachers to know more quickly their students and interact more effectively with them. If the new measure can be shown to have advantages (such as providing equivalent results but delivered more quickly) over the estimates the analysis will have been productive, even though it has not proved that the new instrument is actually measuring autonomy.

8.3.2.3 Measuring distinguished from testing

It has been illustrated in this thesis that there is a useful distinction to be made between measuring and testing (see Sections 2.1.2 and 2.3); one can measure without testing and this can remove apparent objections and inhibitions to the use of quantitative instruments to obtain useful information about learners which can help teachers support their learners' autonomy. Learners' autonomy should not be constrained (Champagne et al. 2001: 49) and clearly distinguishing the present quantitative measure from a test has avoided this type of objection. It may be a useful distinction to observe in future autonomy research into such questions as how autonomy develops over time, how it is affected by context, and whether it is transferable.

8.3.2.4 Not predefining the construct under investigation

I have illustrated that the idea that it is necessary to define the construct before it can be measured is not necessarily universally true; one can measure and then find out what it is that is being measured. Furthermore, one can use a measure even if one does not formally identify its subject; it can be identified in terms of its functions or usefulness. This is important with autonomy because it is still so much debated what it is. This difficulty with defining autonomy in advance does not need to inhibit us because it is not necessary to the function of the measure. In fact, giving a definition

in advance can close off the possibilities of discovery of new dimensions or techniques: as Cofield (2002: 40) points out “empirical research into a complex concept may lead in time to a more precise definition or definitions”. It is not, therefore, unprincipled to adopt a more flexible and open-minded approach to defining the construct, especially in terms of practical uses.

This illustrates another point made by Cofield (2002: 40), i.e. that a definition is not required for measurement. This view appears to challenge ideas of the type that “before developing a test of any construct, one should clearly and explicitly express what one wants to test” (Most & Zeidner 1995). It may be that the clear disassociation I make between the concepts of “testing” and “measuring” will allow the measurement of autonomy-related constructs in a principled way. I propose that not predefining a construct of autonomy, and thus going against what has been the norm for previous authors, is important to the more general acceptance of the measurement of autonomy-relevant dimensions. Covering a broad range of items early in the development of an instrument and then using exploratory factor analysis after data has been collected can reveal the construct which the instrument is measuring and so avoids making definitions prior to data collection.

8.3.2.5 The factor analysis stage of identifying the unifying concept of a scale

The use of factor analysis in my research meant that I studied the statistical literature regarding its procedures and also papers in applied linguistics which made use of it. This reading has made me realise that the stage in factor analysis where each scale is named by the researcher (see Section 3.12.3.3.4) based on its unifying concept did not yet have a set procedure and was thus susceptible to subjective interpretation by

the researcher. In Section 3.12.3.3.4 I presented a procedure for this interpretation of factor groupings, and this can contribute to research in future because it highlights both the lack of and the necessity for a standardised procedure for this step in factor analysis, which is, after all, a widespread method of analysis.

8.3.3 Translation of questionnaires

I have highlighted that there are translation issues involved in questionnaire development which are seldom discussed in the field of language teaching research. Translation is little used in research in our field, perhaps due to the ethic or habit of using the target language as much as possible in language teaching. However, this study argues for increased use of translation in questionnaire research in our field.

It is not handled consistently (see Section 7.7) in research at present, and as I discovered, an uninitiated or casual approach may be counterproductive (see Section 3.13.1). The benefits of the correct use of translation (see Section 3.13.2) in the present research ranged from improved response numbers, to greater “customer satisfaction”, and more reliable responses as respondents will understand their own language more consistently than they do English. Further advantages which I experienced were: firstly I found that translation is useful in considering the wording and clarity of one’s items, even for the original language version; secondly, translation forces engagement with the subjects’ world, culture, way of thinking, and enriches the research itself by informing one’s understanding of what one is trying to research.

I have looked beyond the language teaching field and surveyed the literature of translation with particular attention to the translation of questionnaires, and this

review (see Section 3.13) should be a useful resource for future researchers in the field of language teaching research. The implications are explained in Section 8.5.

8.4 Limitations

As has been noted throughout the thesis and in particular in Sections 3.13.5 and 5.8 the research encountered a number of obstacles and it was not possible to follow the original plan. This has had an impact on the nature and quantity of data collected and consequently has had a significant effect on what can be concluded. In this section I explain these limitations.

The results of the factor analysis must be treated with caution. Factor groups cannot be formed of items which were not included in the initial Long List selection and items which were de-selected in the data reduction process (Dörnyei 2007: 234). This is a feature of factor analyses and means that conclusions about the questionnaire model should not be treated as final. The ratio of respondents to items of a little over 3:1 could be improved. Longer multi item scales would also be preferable (see Section 4.2.1.8 above), but were not achievable as the very large number of items would be impractical. With the quantity of data at present available the sample (see Section 3.7.3) for factor analysis had to be a pooled non-specific representation of the language learner population. The items used in the factor were the same 50 items for all respondents, though some are drawn from the Long List and some were translated into Chinese. This was a necessary compromise to achieve factor analysis of the data. This means that data for specific situations and nationalities has not been available.

The small scale data do not allow statistical significance to be demonstrated. This affects what can be concluded quantitatively from the comparison which took place between the questionnaire data and the year-3 teacher and myself (see Section 5.2). Consequently, quantitatively speaking, conclusions drawn from this comparison regard the year-3 teacher and myself only. Further research with larger samples would be necessary in this area to achieve more generalisable and statistically significant results.

The two interviews (Sections 5.2.2 and 5.2.3) which were carried out are insufficient for drawing wide conclusions. They serve to illustrate the use of the questionnaire and show possible areas meriting further investigation. However, they served to raise considerably my awareness of issues of relevance to the use of the instrument at an individual level and have deeply enriched my own understanding of autonomy.

The quantity of data from the small scale research is therefore insufficient to confirm or otherwise other empirical research or theory, and I have not done this. The larger scale data from the standardised and pooled questionnaire responses which was used for factor analysis is more acceptable and has produced scales with acceptable Cronbach's alpha figures and construct validity has been suggested by the comparison of the questionnaire model with the autonomy literature. However, these data must also be treated with caution and will need to be confirmed with larger scale research (see Section 8.6).

8.5 Implications

Closed-item questionnaires should not be written-out of the promotion of autonomy as a casualty of the paradigm wars. They do serve a purpose if used appropriately.

They are quick to administer and can supply relevant information which can act as a prompt for interaction and understanding of learners by teachers, and if self-administered can aid reflection on learning. These benefits are likely to aid the development of autonomous learning skills and the development of a context where autonomy can be expressed in a practical form. The speed and convenience of a “nimble” autonomous learning-related measure have the potential to encourage the introduction of autonomous practices in classrooms where it was previously seen as too ambitious and a cause of upheavals. Teachers often wish to introduce more autonomy into their classes, but the institutional environment is not conducive to this; an unobtrusive instrument will be one small way for teachers to introduce autonomy into their interaction with their students. An instrument such as the one being developed for this thesis should be able to assist teachers with their understanding of the needs of a class and achieve this in advance of teacher estimates, which is an advantage at the beginning of a course in particular.

This research provides some empirical evidence to suggest that confidence plays a role in autonomy, and this is an idea which has been found in other reports of research and theoretical papers. The implication is that by supporting confidence a teacher can indirectly support autonomy. This however would need to be carried out appropriately so as not to adversely influence the other five areas found in the present research. The model of the electric circuit (Section 7.3.4) may be useful here.

This thesis has highlighted the area of questionnaire translation in second language research and has proposed that the issues which it presents are neglected in books on research methods and that researchers often overlook the importance of translating instruments in their projects, and when they are translated it is often quite informally

done. The reasons for and benefits of translating an instrument have been presented in Sections 3.13 and 7.7. Table 7.3 showed some of the problems which came to light in my own questionnaire after it had been translated informally without the level of attention to details of translation procedure described in Section 3.13. My study argues for the increased use of translation in questionnaire-based research and a more principled and consistent approach to it. This means that more time and resources will need to be dedicated to it. For this to be achieved the profile of translation needs to be raised: it needs to be described in guides to research as an essential part of the reliability and validity of cross-linguistic research, which in the field of second language learning is a large proportion of the research.

Comparisons made with data from cross-linguistic instruments need to be handled with caution, and it is advisable to interpret results with a team which represents all language groups involved.

8.6 Further research

Further research can be broadly divided into two areas: (a) consolidation, by remedying the limitations of the present research (see Sections 5.8 and 8.3.3); and (b) extension, to explore further.

Firstly, as regards consolidation, in the present research only two teachers made estimates and only one of them was properly independent of the research (as the other was myself). This means that there is a need for further and more extensive gathering of teacher estimates (from those not directly involved in the research) with a wider range of classes and teachers.

At the micro scale of comparing my own estimates with the questionnaire results there was over time a convergence that suggests that estimates move towards the questionnaire's reading (see Section 6.3.4). If this analysis could be carried out with a larger statistically significant sample it would confirm or deny that estimates and questionnaire have a correlation.

It would be very useful to involve other researchers and experts in the research. The choice of items was carried out fairly independently, as was the identification and naming of the factors. The contribution of others would have obvious benefits in the areas where interpretations can be unwittingly subjective or tied to specific favoured views of autonomy. A panel of experts with varied positions to make the initial choice of items and to identify the scales produced by factor analysis would strengthen the research.

For the questionnaire it would be important to gather new data from large specific samples, perhaps with as many as ten individuals for each item, so as to enable factor analysis of the Long List of items (which was beyond the resources of the present research). Interviews with respondents should also be carried out. This is necessary to confirm the results of the factor analysis which was carried out with a pooled sample in the present research due to the low number of respondents to any one format of the questionnaire.

Secondly, regarding extension, the present research has explored the methodology necessary for researching a quantitative instrument for supporting autonomy in the classroom. In future, if consolidation produces positive results, it will be necessary to gather data to shed light on how practically useful the instrument is for teachers.

In terms of researching the development of autonomy, the fully validated instrument may be able to probe the categories of the questionnaire's autonomy-related construct to ascertain whether, as seems likely, some are more dispositional and some more variable over time. Repeated administrations of the questionnaire may help to distinguish which of the categories are more variable, and which are more stable.

In Section 7.4.5 I proposed that there would be a “dip” in learners' self-assessment as they initially become better able to reflect on their own learning. This is a testable hypothesis and so the data gathered for consolidation purposes could also be used to research this area of the development of learner autonomy and so start to identify the developmental processes over time as hoped for by Benson (2001: 51; 2010: 78).

Finally, as a self-access centre (SAC) coordinator I had hoped to provide an instrument which could be used by individuals in a SAC or working at home on the Internet. It would provide a formative experience through the process of reflecting on the items in order to answer them, but it would also be enhanced if it offered feedback and support. This interactive online questionnaire would also be an ambition for future research and development to produce a version of the instrument which would be more targeted at self-supporting learners.

8.7 Concluding remarks

Five years ago at the beginning of this research I wanted (I thought) a simple solution to my need for a measure of autonomy to justify my Independent Learning Centre to management in a way that they could readily understand and which would relate to the stated goals of the ILC and the college. In the course of my research this

aspiration has been considerably modified and the question quickly became whether such a thing was possible, and to find out by attempting it critically.

The critical reflexive aspect of my investigation was very interesting. I feel that I went about it in the right way, in principle, though in the event the project was too ambitious for the resources available and when unforeseen difficulties arose there was little margin of safety. Qualitative research requires rich data and much time, and quantitative research needs substantial resources and guaranteed access to large numbers of respondents. Mixed methods research would ideally combine these, making it a very ambitious choice for a PhD student with limited time and resources. I still, however, think it is the best way to conduct research where possible as it combines the advantages of both. Valuable lessons have been learned which I can carry forward professionally.

I hope I have indicated that, though the problems raised regarding quantitative methods in the area of autonomy research are challenges and do restrict what can be usefully investigated by quantitative methods alone, they are not conceptual barriers to it. There is a place for the contributions of quantitative research allied with qualitative methods, especially if they can result in a practical and viable tool for use in the classroom (rather than a universal measure of autonomy).

I did not find the autonomy measuring instrument, but this is not a negative result as I have learned in the process that I did not want it, and that there is a far more satisfying and useful function to be had from an instrument, that of enhancing the autonomy in a classroom by contributing to the understanding between learner and teacher, and helping to clarify problems by initiating a dialogue rather than being a way of labelling a learner as being autonomous at a certain level. I have been able to

understand the impact of the emphasis on the social and situated nature of autonomy because I have worked it through for myself in a way that I hope comes across in this thesis. I have also learned that understanding autonomy is a long and open-ended process.

9 REFERENCES

- Ager, D. E. (1996) *Language policy in Britain and France: The processes of policy*. London: Continuum
- Aiken, L. R. (1997) *Questionnaires and Inventories: Surveying opinions and assessing personality*. New York: John Wiley
- Arthur, L. (2001) Understanding language learners. In *Supporting lifelong language learning: Theoretical and practical approaches*. CILT in association with The Open University
- Atkinson, D. (2002) Toward a sociocognitive approach to second language acquisition. *The modern language journal*. 86/4 525–545
- Bailey, P. D. & Onwuegbuzie, A. J. (2002) The role of study habits in foreign language courses. *Assessment and evaluation in higher education*. 27/5 463–473
- Bandura, A. (1997) *Self-efficacy: the Exercise of control*. New York: W. H. Freeman
- Barnett, L. (1993) Teacher off: Computer technology, guidance and self-access. *System*. 21/2 295–304
- Baumann, U. (2007) Developing autonomy and intercultural competence in a distance learning environment. In Gardner, D. (ed.) *Learner autonomy 10: Integration and support*. Dublin: Authentik
- Benson, P. (1997) The philosophy and politics of learner autonomy. In Benson, P. & Voller, P. (eds.) *Autonomy and independence in language learning*. London: Longman
- Benson, P. (2001) *Teaching and researching autonomy in language learning*. Harlow: Pearson Education
- Benson, P. (2006) State-of-the-art article: Autonomy in language teaching and learning. *Language teaching*. 40 21–40

- Benson, P. (2009) Making sense of autonomy in language learning. In Pemberton R., Toogood, S. & Barfield, A. (eds.) *Maintaining control: Autonomy and language learning*. Hong Kong: HKUP
- Benson, P. & Lor, W. (1998) *Making sense of autonomous language learning*. *English Centre Monograph No. 2*. Hong Kong University English Language Centre
- Benson, P. & Voller, P. (1997) *Autonomy and independence in language learning*. Harlow: Pearson
- Biesta, G. (2009) Good education in an age of measurement: On the need to reconnect with the question of purpose in education. *Educational assessment, evaluation and accountability*. 21 33-46
- Blais, A. & Gidengil, E. (1993) Things are not always what they seem: French-English differences and the problem of measurement equivalence. *Canadian Journal of Political Science*. 26/3 541-555
- Blin, F. (2004) CALL and the development of learner autonomy: Towards an activity-theoretical perspective. *ReCALL*. 16/2 377-95
- Blin, F. (2005) *CALL and the development of learner autonomy: an activity theoretical study*. PhD Thesis. The Open University
- Boud, D. (ed.) (1988) *Developing student autonomy in learning – Second edition*. London: Kogan Page
- Bourque, L. B. & Fielder, E. P. (1995) *How to conduct self-administered and mail surveys*. Thousand Oaks: Sage
- Breen, P. & Mann, S. (1997) Shooting arrows at the Sun: Perspectives on a pedagogy for autonomy. In Benson, P. & P. Voller (eds.) *Autonomy and independence in language learning*. London: Longman

- Bruner, J. S. (1966) *Toward a theory of instruction*. Cambridge MA: Harvard University Press
- Bruner, J. S. (1974). *Beyond the information given*. London: Allen and Unwin.
- Candy, P.C., Crebert, G. & O’Leary, J. (1994) *Developing lifelong learning through undergraduate education* (Canberra, National Board of Employment, Education and Training). Retrieved May 2009 from
http://www.dest.gov.au/sectors/training_skills/publications_resources/indexes/documents/94_21_pdf.htm
- Carr, M. & Claxton, G. (2002) Tracking the development of learning dispositions. *Assessment in education*. 9/1
- Champagne, M-F., Clayton, T., Dimmitt, N., Laszewski, M., Savage, W., Shaw, J., Richmond, S., Thein, M. M. & Walter, P. (2001) The assessment of learner autonomy and language learning. In Dam, L. (ed.) *Learner autonomy: New insights. AILA Review*. 15 45-55
- Chan, V. (2001) Readiness for learner autonomy: What do our learners tell us? *Teaching in higher education*. 6/4
- Chan, V. (2003) Autonomous language learning: The teachers’ perspectives. *Teaching in higher education*. 8/1
- Chan, V., Spratt, M. & Humphreys, G. (2002) Autonomous language learning: Hong Kong tertiary students’ attitudes and behaviours. *Evaluation and research in education*. 16/1
- Chang, L. Y. (2007) The influences of group processes on learners’ autonomous beliefs and behaviours. *System*. 35 322–337
- Chen, C., Lee, S. & Stevenson, H. W. (1995) Response style and cross-cultural comparisons of rating scales among East Asian and North American students. *Psychological science*. 6/3

- Cohen, L., Manion, L. & Morrison, K. (2000) *Research methods in education*.
London: RoutledgeFalmer
- Converse, J. M. & S. Presser (1986) *Survey questions: Handcrafting the
standardized questionnaire*. Beverly Hills: Sage
- Costello, A. B. & Osborne, J. W. (2005) Best practices in exploratory factor analysis:
Four recommendations for getting the most from your analysis. *Practical
assessment research and evaluation*. 10/7
- Cotterall, S. (1995) Readiness for autonomy: Investigating learner beliefs. *System*.
23/2 195-205
- Cotterall, S. (1999) Key variables in language learning: What do learners believe
about them? *System*. 27 493-513
- Cotterall, S. (2000) Promoting learner autonomy through the curriculum: Principles
for designing language courses. *ELTJ*. 54/2
- Cotterall, S. (2009) Learner autonomy in a mainstream writing course: Articulating
learning gains. In Pemberton R., Toogood, S. & Barfield, A. (eds.) *Maintaining
control: Autonomy and language learning*. Hong Kong: HKUP
- Crabbe, D. (1993) Fostering autonomy from within the classroom: The teacher's
responsibility. *System*. 21/4 443-452
- Dam, L. (2000) Evaluating autonomous learning. In Sinclair, B., McGrath, I. and
Lamb, T. (eds.). *Learner autonomy, teacher autonomy: Future directions*. Harlow:
Longman
- Dam, L. (2009) The use of logbooks – A tool for developing learner autonomy. In
Pemberton R., Toogood, S. & Barfield, A. (eds.) *Maintaining control: Autonomy
and language learning*. Hong Kong: HKUP

- Dam, L. & Legenhausen, L. (1996) The acquisition of vocabulary in an autonomous learning environment – The first months of beginning English. In Pemberton, R. et al. (eds.) *Taking control: Autonomy in language learning*. Hong Kong: HKUP
- Dam, L., Eriksson, R., Little, D., Miliander, J. & Trebbi, T. (1990) Towards a definition of autonomy. In Trebbi, T., (ed.) *Third Nordic workshop on developing autonomous learning in the FL classroom, Bergen, August 11-14, 1989: Report*. Bergen, Norway: Institutt for praktisk pedagogikk, Universitet i Bergen. (Mimeo)
- Damon, W. & Hart, D. (1982) The development of self-understanding from infancy through adolescence. *Child development*. 53 841-864
- Davis, H. A. (2003) Conceptualizing the role and influence of student-teacher relationships on children's social and cognitive development. *Educational psychologist*. 38/4 207-234
- Deakin-Crick, R. & Wilson, K. (2005) Being a learner: A virtue for the 21st century. *British journal of educational studies*. 53
- Deci, E. L., Vallerand, R. J., Pelletier, L. G. & Ryan, R. M. (1991). Motivation and education: the self-determination perspective. *Educational psychologist*. 26/3
- Dickinson, L. (1987) *Self-instruction in language learning*. Cambridge: CUP
- Dickinson, L. (1992). *Learner autonomy 2: Learner training for language learning*. Dublin: Authentik
- Dickinson, L. (1995) Autonomy and motivation: A literature review. *System*. 23/2
- Dörnyei, Z. (1998) Motivation in Second and Foreign Language Learning. *Language teaching*. 31 117–35
- Dörnyei, Z. (2001) *Teaching and researching motivation*. Harlow: Pearson
- Dörnyei, Z. (2003) *Questionnaires in second language research: Construction, administration, and processing*. New Jersey: Lawrence Erlbaum

- Dörnyei, Z. (2007) *Research methods in applied linguistics*. Oxford: OUP
- Dweck, C. S. (1986) Motivational processes affecting learning. *American psychologist*. 41/10 1040-1048
- Ellard, J. H., & T. B. Rogers (1993) Teaching questionnaire construction effectively: The ten commandments of question writing. *Contemporary social psychology*. 17/1 17-20
- Ellis, R. (1994) *The study of second language acquisition*. Oxford: OUP
- Ellis, G. & B. Sinclair (1989). *Learning to learn English: A course in learner training*. Cambridge: CUP
- Esch, E. (1996). Promoting learner autonomy: Criteria for the selection of appropriate methods. In Pemberton, R. et al. (eds.) *Taking control: Autonomy in language learning*. Hong Kong: HKUP 35-48
- Esch, E. (2009) Crash or clash? Autonomy 10 years on. In Pemberton R., Toogood, S. & Barfield, A. (eds.) *Maintaining control: Autonomy and language learning*. Hong Kong: HKUP 27-44
- Field, A. (2000) *discovering statistics using SPSS for windows: Advanced techniques for beginners*. London: Sage
- Flavell, J. H. (1979) Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American psychologist*. 34/10 906-911
- Flavell, J. H. (1987) Speculation about the nature and development of metacognition. In: F. E. Weiner & R. H. Kluwe (eds.). *Metacognition, motivation and understanding*. Hillsdale, NJ: Lawrence Erlbaum 1-29
- Fowler, F. J. (1993) *Survey research methods*. Newbury Park: Sage
- Gan, Z. (2004) Attitudes and strategies as predictors of self-directed language learning in an EFL context. *International journal of applied linguistics* 4/3

- Gan, Z., Humphreys, G. & Hamp-Lyons, L. (2004) Understanding successful and unsuccessful EFL students in Chinese universities. *The modern language journal*. 88/2
- Gardner, R. C., & MacIntyre, P. D (1993) On the measurement of affective variables in second language learning. *Language learning* Vol. 43 No. 2 157–194
- Gillham, B. (2000) *Developing a questionnaire*. London: Continuum
- Green, J. M. & Oxford, R. L. (1995) A closer look at learning strategies, L2 proficiency, and gender. *TESOL quarterly*. Vol. 29 261–97.
- Gremmo, M. & Riley, P. (1995) Autonomy, self-direction and self access in language teaching and learning: The history of an idea. *System*. 23/2 151–164
- Grice, H. P. (1975) Logic and conversation. In Cole, P. & Morgan, J. L. (eds.) *Syntax and semantics: Speech acts. Volume 3* 41–58 New York: Academic
- Harkness, J. A. & Schoua-Glusberg, A. (1998) Questionnaires in translation. In Harkness, J. A. (ed.) *ZUMA-Nachrichten spezial no. 3*. Mannheim: Zentrum für Umfragen, Methoden und Analysen
- Holec, H. (1979) *Autonomie et apprentissage des langues étrangères*. Strasbourg: Council of Europe
- Holec, H. (1981) *Autonomy and foreign language learning*. Oxford: Pergamon
- Holec, H. (2008) Foreword. In Lamb, T. & Reinders, H. (eds.) *Learner and teacher autonomy: concepts, realities and responses*. John Benjamins: Amsterdam
- Holliday, A. (2003) Social autonomy: addressing the dangers of culturalism in TESOL. In Palfreyman, D. & Smith, R. C. (eds.) *Learner autonomy across cultures: Language education perspectives*. Basingstoke: Palgrave Macmillan
- Hughes, A. (2003) *Testing for language teachers – Second edition*. Cambridge: CUP

- Hurd, S. (1998) "Too carefully led or too carelessly left alone"? *Language learning journal*. 17 70-74
- Hurd, S (2004) Autonomy and the distance language learner. In Holmberg, B., Shelley, M. & White, C. (eds.) *Distance education and languages: Evolution and change*. Clevedon: Multilingual Matters. Accessed June 2009 from http://oro.open.ac.uk/622/1/Hurd_Autonomy_chapter.pdf
- Jones, F. R. (1998) Self-instruction and success: A learner-profile study. *Applied linguistics*.. 19/3 378–406
- Johnson, T. P. (1998) Approaches to equivalence in cross-cultural and cross-national survey research. In Harkness, J. A. (ed.) *ZUMA-Nachrichten spezial no. 3*. Mannheim: Zentrum für Umfragen, Methoden und Analysen
- Karlsson, L., Kjisik, F. & Nordlund, J. (2005) Language counselling: A critical and integral component in promoting an autonomous community of learning. *System*. 35/46-65
- Kenny, B. (1993) For More Autonomy. *System* Vol. 21 No. 4 431-442
- Kohonen, V. (1992) Experiential language learning: Second language learning as cooperative learner education. In Nunan, D. (ed.) *Collaborative language learning and teaching*. Cambridge: CUP
- La Ganza, W. (2002). Rethinking a topic: Challenges in researching autonomy and creativity in language teaching for business students. In Benson, P. & Toogood, S. (eds.), *Learner autonomy 7: Challenges to research and practice*. Dublin: Authentik.
- La Ganza, W. (2008) Learner autonomy – teacher autonomy: Interrelating and the will to empower". In Lamb, T. & Reinders, H. (eds.) *Learner and teacher autonomy: Concepts, realities, and responses*. Amsterdam: John Benjamins

- Lai, J. (2001) Towards an analytic approach to assessing learner autonomy. *AILA review 15 learner autonomy: New insights*. 34-44
- Lamb, T. & Reinders, H. (eds) (2006) *Supporting independent language learning: Issues and interventions*. Frankfurt am Main: Peter Lang
- Lamb, T. (2009) Controlling learning: learners' voices and relationships between motivation and learner autonomy. In Pemberton R., Toogood, S. & Barfield, A. (eds.) *Maintaining control: Autonomy and language learning*. Hong Kong: HKUP
- Lantolf, J.P. (2000) Introducing sociocultural theory. In Lantolf, J.P. (ed.) *Sociocultural theory and second language learning*. Oxford: OUP
- Leech, G., Rayson, P. & Wilson, A. (2001) *Word frequencies in written and spoken English: Based on the British National Corpus*. Longman: London
- Legutke, M. & Thomas, H. (1991) *Process and experience in the language classroom*. London: Longman
- Li, A. (2005) A look at Chinese ESL students' use of learning strategies in relation to their English language proficiency, gender and perceived language difficulties – A quantitative study. In Anderson, H. Hobbs, M., Jones-Parry, J., Logan, S., & Lotovale, S. (eds.) *Supporting independent learning in the 21st century. Proceedings of the second conference of the Independent Learning Association, Auckland NZ, 9-12 September 2005*. Auckland: Independent Learning Association Oceania.
- Little, D. (1991) *Learner autonomy 1: Definitions, issues and problems*. Dublin: Authentik
- Little, D. (1996) Freedom to learn and compulsion to interact: Promoting learner autonomy through the use of information systems and information technologies. In

- Pemberton, R. et al. (eds.) *Taking control: Autonomy in language learning*. Hong Kong: HKUP
- Little, D. (2000) Learner autonomy and human interdependence: Some theoretical and practical consequences of a social-interactive view of cognition, learning and language. In Sinclair, B., McGrath, I. & Lamb, T. (eds.) *Learner autonomy, teacher autonomy: Future directions*. Harlow: Pearson
- Little, D. (2003) Learner autonomy and second/foreign language learning. Subject centre for languages, linguistics and area studies. Retrieved May 2007 from: <http://www.llas.ac.uk/resources/goodpractice.aspx?resourceid=1409>
- Little, D. (2005) The common European framework and the European Language Portfolio: Involving learners and their judgements in the assessment process. *Language testing*. 22/3. 321–336
- Little, D. (2009) Learner autonomy, the European Language Portfolio, and teacher development. In Pemberton R., Toogood, S. & Barfield, A. (eds.) *Maintaining control: Autonomy and language learning*. Hong Kong: HKUP
- Little, D., Ridley, J. & Ushioda, E. (2002) *Towards greater learner autonomy in the foreign language classroom*. Dublin: Authentik
- Littlewood, W. (1981) *Communicative language teaching: An introduction*. Cambridge: CUP
- Littlewood, W. (1996) “Autonomy”: An anatomy and a framework. *System*. 24/4
- Littlewood, W. (1999) Defining and developing autonomy in East Asian contexts. *Applied linguistics*. 20/1 71-94
- Macaro, E. (1997) *Target language, collaborative learning and autonomy*. Clevedon: Multilingual Matters

- Macaro, E. (2008) The shifting dimensions of language learner autonomy. In Lamb, T. & Reinders, H. (eds.) *Learner and teacher autonomy: Concepts, realities and responses*. John Benjamins: Amsterdam
- Messick, S. (1975) The standard problem: Meaning and values in measurement and evaluation. *American psychologist* October 1975 955-966
- Miles, M. B. & Huberman, A. M. (1994) *Qualitative data analysis – Second edition*. Thousand Oaks: Sage
- Miller, R. L., Maltby, J., Fullerton, D.A. & Acton, C. (2002) *SPSS for social scientists*. Basingstoke: Palgrave
- Moreira, M. A. (2007) On democracy and learner autonomy in initial teacher education: Guerrilla warfare. In Andrew Barfield, A. & Brown, S. H. (eds.) *Reconstructing autonomy in language education: Inquiry and innovation*. Basingstoke: Palgrave Macmillan.
- Morrison, B. (2005) Evaluating learning gain in a self-access language learning centre. *Language teaching research*. 9/3 267–293
- Most, R., & Zeidner, M. (1995) Constructing personality and intelligence instruments. In Saklofske, D. H. & Zeidner, M. (eds.), *International handbook of personality and intelligence*. 475-504. New York: Plenum Press
- Muijs, D. (2004) *Doing quantitative research in education*. London: Sage
- Munday, J. (2001) *Introducing translation studies: Theories and applications*. Abingdon: Routledge
- Murphey, T. 2003. Learning to surf: Structuring, Negotiating, and owning autonomy. In Barfield A. & Nix, M. (eds.) *Autonomy you ask!* 1–10. Tokyo: JALT Learner Development Special Interest Group.
- Mynard, J. & Almarzouqi, I. (2006) Investigating peer tutoring. *ELTJ*. 60/1 13-22

- Nakatani, Y. (2006) Developing an oral communication strategy inventory. *The modern language journal*. 90/2 151–168
- Newmark, P. (1981) *Approaches to translation*. Oxford and New York: Pergamon
- Newmark, P. (1991) *About Translation*. Clevedon: Multilingual Matters
- Nida, E. (1964) Principles of Correspondence. In Venuti, L. (2000) *The translation studies reader*. London, New York: Taylor & Francis
- Nunan, D. (1992) *Research methods in language learning*. Cambridge: CUP
- Nunan, D. (1997) Designing and adapting materials to encourage learner autonomy. In Benson, P. & Voller, P. (eds.) *Autonomy and independence in language learning*. London: Longman
- Ohta, A. S. (2000) Rethinking interaction in SLA: Developmentally appropriate assistance in the zone of proximal development and the acquisition of L2 grammar. In Lantolf, J. P. (ed.) *Sociocultural theory and second language learning*. Oxford: OUP
- Oppenheim, A. N. (1992) *Questionnaire design, interviewing and attitude measurement*. London: Continuum
- Oxford, R. L. (1990) *Language learning strategies: What every teacher should know*. Boston, Massachusetts: Heinle and Heinle
- Oxford, R. L. (2003) Toward a more systematic model of L2 learner autonomy. In Palfreyman, D. & Smith, R. C. (eds.) *Learner autonomy across cultures: Language education perspectives*. Basingstoke: Palgrave Macmillan
- Page, B. (1992) *Letting go – taking hold: A guide to independent language teaching by teachers for teachers*. London: CILT

- Pavlenko, A. & Lantolf, J. P. (2000) Second language learning as participation and the (re)construction of selves. In Lantolf, J. P. (ed.) *Sociocultural theory and second language learning*. Oxford: OUP. 155-177
- Pemberton, R. (1996) Introduction. In Pemberton, R. et al. (eds.) *Taking control: Autonomy in language learning*. Hong Kong: HKUP
- Phillipson, R. (1992) *Linguistic Imperialism*. Oxford: OUP
- Prodromou, L. (1992) What culture? Which culture? Cross-cultural factors in language learning. *ELTJ*. 46/1
- Punch, K. F. (2005) *Introduction to social research: Quantitative and qualitative approaches – Second edition*. London: Sage
- Reeves, T. C. & Okey, J. R. (1996) Alternative assessment for constructivist learning environments. In Wilson, B. G. *Constructivist learning environments: Case studies in instructional design*. Englewood Cliffs, NJ: Educational Technology.
- Reinders, H. (2007) Big brother is helping you: Supporting self-access language learning with a student monitoring system. *System* 35 93-111
- Reinders, H. & Lázaro, N. (2007) Current approaches to assessment in self-access language learning. *TESL-EJ* 11/3
- Richards, K. (2003) *Qualitative inquiry in TESOL*. Basingstoke: Palgrave
- Rivers, W. M. & Melvin, B. J. (1981) Language learners as individuals: Discovering their needs, wants, and learning styles. In Alatis, J. E., Altman, H. B. & Alatis, P.M. (eds.) *The second language classroom: Directions for the 1980's*. New York: OUP
- Rivers, W. P. (2001) Autonomy at all costs: An ethnography of metacognitive self-assessment and self-management among experienced language learners. *The modern language journal*. 85/2 279–290

- Rogers, C. (1969) *Freedom to learn*. Columbus, Ohio: Charles E. Merrill
- Russell, D. W. (2002) In search of underlying dimensions: The use (and abuse) of factor analysis in personality and social psychology bulletin. *Personality and social psychology bulletin* 28/12 1629–1646
- Ryan, R. M & Connell, J.P. (1989) Perceived locus of causality and internalization: Examining reasons for acting in two domains. *Journal of personality and social psychology* 57/5 749-761
- Ryan, R. M & Deci, E. L. (2000) Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary educational psychology*. 25 54-67
- Sakui, K. & Gaies, S. J. (1999) Investigating Japanese learners' beliefs about language learning. *System* Vol. 27 473 – 492
- Samuelsson-Brown, G. (1996) Working procedures, quality and quality assurance. In Owens, R. (ed.) *The translator's handbook 3rd edition*. The Association for Information Management: London
- Schommer, M. (1990) Effects of beliefs about the nature of knowledge on comprehension. *Journal of educational psychology*. Vol. 82 No. 3 498-504
- Schwarz, N. (2003) Culture-sensitive context effects: A challenge for cross-cultural surveys. In Harkness, J. A, Van de Vijver, F. J. R. & Mohler, P. Ph. (eds.) *Cross-cultural survey methods* 93-100 Hoboken: John Wiley
- Shaw, J. (2008) Teachers working together: What do we talk about when we talk about autonomy? In Lamb, T. & Reinders, H. (eds.) *Learner and teacher autonomy: Concepts, realities and responses*. John Benjamins: Amsterdam
- Sheerin, S. (1997). An exploration of the relationship between self-access and independent learning. In Benson, P. & Voller, P. (eds.) *Autonomy and independence in language learning*. London: Longman.

- Sinclair, B. (1999) More than an act of faith? Evaluating learner autonomy. In Kennedy, C. (ed.) *Innovation and best practice*. Harlow: Pearson
- Sinclair, B. (2000) Learner autonomy: The next phase. In Sinclair, B., McGrath, I. & Lamb, T. (eds.) *Learner autonomy, teacher autonomy: Future directions*. Harlow: Pearson
- Sinclair, B. (2009) The teacher as learner: Developing autonomy in an interactive learning environment. In Pemberton R., Toogood, S. & Barfield, A. (eds.) *Maintaining control: Autonomy and language learning*. Hong Kong: HKUP
- Smith, R. (2008) Key concepts in ELT: Learner autonomy. *ELTJ* Vol. 62 No. 4
- Smith, R. & Ushioda, E. (2009) Autonomy: Under whose control? In Pemberton R., Toogood, S. & Barfield, A. (eds.) *Maintaining control: Autonomy and language learning*. Hong Kong: HKUP
- Smith, R. C. (2003) Pedagogy for autonomy as (becoming-) appropriate methodology. In Palfreyman, D. & Smith, R. C. (eds.) *Learner autonomy across cultures: Language education perspectives*. London: Palgrave Macmillan 129-46
- Smith, T. W. (2004) Developing and evaluating cross-national survey instruments. In Presser, S., Rothgeb, J. M., Couper, M. P., Lessler, J. T., Martin, E., Martin, J. & Singer, E. (eds.) *Methods for testing and evaluating survey questionnaires*. 431-452. John Wiley and Sons: Hoboken
- Spencer-Oatey, H. & Xiong, Z. (2006) Chinese students' psychological and sociocultural adjustments to Britain: An empirical study. *Language, culture and curriculum* Vol. 19 No. 1
- Spratt, M., Humphreys, G. & Chan, V. (2002) Autonomy and motivation: Which comes first? *Language teaching research* Vol. 6 No. 3 245–266

- Stansfield, C. W. (2003) Test translation and adaptation in public education in the USA. *Language testing* Vol. 20 No. 2 189–207
- Stanton, H. 1988. Independent study: A matter of confidence? In Boud, D (ed.) *Developing student autonomy in learning* (second edition) (118-131). London: Kogan Page
- Sudman, S. & Bradburn, N.M. (1982) *Asking questions: A practical guide to questionnaire design*. San Francisco: Jossey-Bass
- Teddlie, C. & Tashakkori, A. (2009) *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences*. Thousand Oaks: SAGE
- Toohey, K. (2007) Autonomy/agency through socio-cultural lenses. In Barfield, A. & Brown, S. H. (eds.) *Reconstructing autonomy in language education: Inquiry and innovation*. Basingstoke: Palgrave Macmillan.
- Tourangeau, R., Rips, L. J. & Rasinski K. (2000) *The psychology of survey response*. Cambridge: CUP
- Tremblay, P. F. (2001) Research in second language learning motivation: psychometric and research design considerations. In Dörnyei, Z. & Schmidt R. (eds.), *Motivation and second language acquisition*. Honolulu: University of Hawaii. 239 – 55.
- Tsang, W. Y. (2005) *Developing learner autonomy through self-access*. MA Dissertation. University of Hong Kong
- Tseng, W-T., Dörnyei, Z., & Schmitt, N. (2006) A new approach to assessing strategic learning: The case of self-regulation in vocabulary acquisition. *Applied linguistics* Vol. 27 No. 1 78–102

- UK Higher Education International Unit (2010) *International higher education in facts and figures*. UK Higher Education International Unit. Retrieved on 06/01/2011 from: http://www.international.ac.uk/our_research_and_publications/index.cfm
- Ushioda, E. (1996) *Learner autonomy 5: The role of motivation*. Dublin: Authentik
- Ushioda, E. (2003) Motivation as a socially mediated process. In Little, D., Ridley, J. & Ushioda, E. (eds.) *Learner autonomy in the foreign language classroom: Teacher, learner, curriculum and assessment*. Dublin: Authentik.
- Van de Vijver, F. R. & Leung, K. (1997) *Methods and data analysis for cross-cultural research*. Thousand Oaks: Sage
- Van de Vijver, F. R., & Hambleton, R. K. (1996) Translating tests: Some practical guidelines. *European psychologist*. Vol. 1 No. 2 89-99
- Vickers, C. H. & Ene, E. (2006) Grammatical accuracy and learner autonomy in advanced writing. *ELTJ* Vol. 60 No. 2. 109-116
- Victori, M. & Lockart, W. (1995). Enhancing metacognition in self-directed learning. *System* Vol. 23 No. 2 223-234
- Vygotsky, L. (1978) *Mind in society*. Cambridge, MA: Harvard University Press
- Vygotsky, L. (1994) Academic concepts in school aged children. In van der Veer, R. & Valsiner, J. (eds.) *The Vygotsky reader*. Oxford: Blackwell
- Wagman, J. C (2005) *The effects of an inquiry-internet research project on motivation, self-efficacy, and academic autonomy in heterogenously grouped high school Latin i students*. Ph.D. Thesis, Capella University. Retrieved June 2009 from http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/1b/bf/16.pdf
- Walker, L. (2001) Learning strategies and learner autonomy. In *Supporting lifelong language learning: Theoretical and practical approaches*. CILT in association with The Open University

- Weiner, B. (1979) A theory of motivation for some classroom experiences. *Journal of educational psychology*. Vol. 71 No. 1 3-25
- Wenden, A. L. (1987) Conceptual background and utility. In Wenden, A. L. & Rubin, J. (eds.) *Learner strategies in language learning*. Cambridge: Prentice-Hall
- Wenden, A. L. (1991) *Learner strategies for learner autonomy*. Hemel Hempstead: Prentice Hall
- Wenden, A. L. (1995) Learner training in context: A knowledge-based approach. *System* Vol. 23 No. 2 183-194
- Wenden, A. L. (1998) Metacognitive knowledge and language learning. *Applied linguistics* Vol. 19 No. 4 515-537
- Williams, M. & Burden, R. L. (1997) *Psychology for language teachers*. Cambridge: CUP
- Yang, N-D. (1999) The relationship between EFL learners' beliefs and learning strategy use. *System* Vol. 27 515-535
- Zimmerman, B. J. (1995). Self-efficacy and educational development. In Bandura, A. (ed.) *Self-efficacy in changing societies*. Cambridge: CUP

10 APPENDICES

10.1 Full 256 items with back translation

Translations that seemed problematic are shaded.

No.	Original English Text	Back translation
001	I will be happy when I can stop learning	I want to stop learning.
002	One day I will stop learning	Some day, I will stop learning.
003	I want to learn something new every day	I hope to learn new things every day.
004	Learning continues all your life	It is never too old to learn.
005	I work hard	I am hard-working.
006	When I study I take breaks in order to maintain my concentration	When I am learning, I relax timely to ensure the concentration.
007	When I study I am organised	My learning is ordered.
008	I am good at studying on my own	I am good at learning independently.
009	I know how to find the information I want	I know how to find out the information I need.
010	I give myself targets for studying	I set learning goals for myself.
011	I am good at planning my learning	I plan my learning well.
012	I watch TV or videos in English in my own time	I watch the English programs or videos in my spare time.
013	When I learn something new I feel satisfaction in myself	When I have learned new things, I feel satisfied.
014	When I learn something new I feel good because the teacher is happy	When I have learned new things, I feel well because the teacher is satisfied.
015	When I learn something new I don't feel good	When I have learned new things, I don't feel well.
016	When I learn something new I feel good because I can stop learning it	When I have learned new things, I feel well because I need not learn it forever.
017	If I must finish a job at a certain time I finish early	If I have to finish a task before the deadline, I will accomplish it as early as possible.
018	I meet deadlines	I finish the task before the deadline.
019	The student's job is to remember the content of all lessons	The task of the students is to remember well the contents of all the classes.
020	All teachers are equally good	All the teachers are good.
021	All lessons are equally valuable	All the courses are of the same value.
022	I know how to study	I know how to learn.
023	Students should always do what the teacher says	Students should always do as teacher orders.
024	The student's job is to develop as a person	The task of the students is to develop individually.
025	I know my own abilities	I know my own ability.
026	I feel lucky when I get good marks	When I've got good scores, I feel very lucky.
027	I need a teacher to help me	I need teachers' help.
028	I feel unlucky when I get bad marks	When I've got bad scores, I feel very unlucky.
029	I know which is my best subject	I know my best subject.
030	The teacher's job is to give me all the information	The teacher's task is to provide me with all the relevant information.
031	The teacher's job is to help me learn	The teacher's task is to help me learn.
032	The teacher's job is to control students	The teacher's task is to control and administrate the students.
033	Everybody can make progress if they try	Everyone can make progress if only s/he makes great efforts.
034	If I am not sure what I have to do, I don't worry about it	If I am not clear about what I am to do, I put it aside.

035	If I am not sure what I have to do I ask somebody	If I am not clear about what I am to do, I ask other people.
036	If I find a word that I don't know, I always ignore it and continue reading	When I don't know the words in the reading materials, I put them aside and go on reading.
037	If I find a word that I don't know, I always ask a teacher first	When I don't know the words in the reading materials, I ask teacher at once.
038	If I find a word that I don't know, I look it up in a dictionary first	When I don't know the words in the reading materials, I look them up in the dictionary at once.
039	If I find a word that I don't know, I try to guess it	When I don't know the words in the reading materials, I try to guess their meanings.
040	I usually need the teacher to help me with my learning	I usually need teacher to help me learn.
041	I choose my own ways of studying	I choose my own learning method.
042	I think about different ways of studying	I have thought about different kinds of learning methods.
043	When I have something to learn I try to think of different ways of doing it	When I learn new things, I think about different methods can be employed.
044	I know my strong points and weak points	I know my merits and shortcomings.
045	I think about how I study best	I wonder how I can learn best.
046	I can describe the learning strategies I use	I can describe my learning methods.
047	When I finish something I think about the ways I worked	When I have finished a task, I think about the methods I have employed.
048	When I finish something I think about ways to do it differently in the future	When I have finished a task, I think about how to do it by using different methods in future.
049	I have changed the way I learn after thinking about it.	After I thought about these methods, my learning methods have changed.
050	I have tried different ways of learning	I have tried various kinds of learning methods.
051	I know why I have problems learning	I know the reason why I have difficulties in learning.
052	I try to fix problems I have in learning	I try to solve the problems existing in my learning process.
053	I know some different ways of learning	I know some different kinds of learning methods.
054	I choose the best way to learn something	When I am learning, I choose the best learning method.
055	I know why I did well or did badly	I know the reason why I do well or badly.
056	I have made my own plans for my learning next week	I have already made a plan for my learning of next week.
057	It is important to finish an exercise before my classmates	It is important for me to finish the exercise earlier than my classmates.
058	I use the teacher's comments and corrections in my written work to improve my English	I use the teachers' comments and explanations to improve my English writing.
059	When I like a learning activity, I know why I like it	When I like one learning activity, I know the reason why I like it.
060	I always agree with what a teacher says	I always approve of what the teacher says.
061	I find it difficult to tell facts from opinions when reading	It is difficult for me to distinguish the facts from the author's opinion when I am reading.
062	Sara eats a lot of sweets so she must be fat	Sally eats a lot of sugar everyday, so she is definitely fat.
063	My dictionary is always right with its definitions.	The explanations in the dictionary are always right.
064	Science books contain only facts	The science books only consist of facts.
065	I learn about all kinds of different things outside class	I learn various kinds of things out of class.
066	Knowledge is something a teacher gives me	The knowledge is given by the teacher.
067	Knowledge is something I construct for myself	I cumulate the knowledge by myself.

068	I trust the Internet	I trust the internet.
069	In the last 4 months I have disagreed with something a teacher told the class	During the last four months, I have voiced different opinions for the teachers' utterance.
070	The Great Wall of China can be seen from space	The Great Wall can be seen from the Space.
071	If I learn something well, it is because I studied well	If I learn well, that is because I work hard.
072	If I learn something well, it is because my teacher taught well	If I learn well, that is because the teacher teaches well.
073	The teacher is responsible for my learning	The teacher should be responsible for my learning.
074	Memorization is the best way to learn	Reciting is the best learning method.
075	I am responsible for my learning	I must be responsible for my learning.
076	I rely on the teacher	I rely on my teachers.
077	I am self-reliant	I trust myself.
078	I enjoy making my own choices about learning	I like to make decision on my learning by myself.
079	I want to make my own choices about learning	I want to make decision on my learning by myself.
080	I make my own choices about learning	I make decision on my learning by myself.
081	It is good to make your own choices about learning	It is good to make decision on learning by myself.
082	I like the teacher to make the choices about learning	I like the teacher to make decision on my learning.
083	My work is my own, not my teacher's	My homework is mine, not the teacher's.
084	My work is my teacher's, not mine	My homework is teacher's, not mine.
085	I have a mature attitude to learning	My attitude towards learning is mature.
086	When I read an English text I need to understand every word in it	When I read an English text, I need to make clear the meaning of every word.
087	I need to be sure about instructions	The demands for me must be clear.
088	If I am not sure about something it bothers me	If I am not sure about something, I will be puzzled.
089	There is no one correct way to write an essay	There is not only one correct writing method.
090	I learn exclusively about college subjects	I only learn the subjects provided by school.
091	Learning well is a talent that some people have and others do not have	Learning well is the endowment for some people, but some other people do not possess this natural gift.
092	I can learn how to learn better	I can learn to how to learn better.
093	Reading is a passive activity; the information passes from the page to you	Reading is a passive activity. The information transfers from the book to you.
094	Reading is an active activity	Reading is an active activity.
095	To read you must proceed word by word	When we are reading, we have to read word by word.
096	There is one correct way of reading	There is only one correct reading method.
097	I predict the content of a text (using pictures, headings, the context etc.)	I predict the content of a passage (by pictures, titles or context)
098	I read newspapers in a different way to books	I use different method to read books and newspapers.
099	I sometimes look up words on the internet or in reference books	Sometimes, I look up the unknown words in net or in the reference books.
100	Last time I read an English text I predicted the content of it	When I read the English text last time, I predicted the content of it.
101	Different types of text (novel, newspaper, web site etc.) are read in different ways	We should use different methods to read different kinds of articles (novels, newspapers, websites, etc.).
102	My general knowledge helps me to understand texts I read	My basic common sense helps me with the comprehension of the articles.
103	It is best to read by starting at the beginning and reading line by line to the end	When reading an article, it is better to read from the very beginning, and come to the end word by word.

104	I read in English outside class	I also read the English articles in my spare time.
105	When I read in English I think about what the source of the text is	When I am reading English articles, I consider their sources.
106	I know the sources of the texts I read	I know the sources of my reading articles.
107	When I read I think about the motives of the writer	When I am reading, I consider the author's writing purpose.
108	When I read I start at the beginning and read line by line to the end	When I am reading, I read from the beginning and then read word by word to the end.
109	I predict the content before I listen	Before I listen to one passage, I predict its content.
110	Every word is important for understanding a listening text	Every word is important for comprehending a listening text.
111	The last time I listened to English I tried to predict the content	When I listen to English last time, I tried to predict the content of the passage.
112	I worry if I don't understand everything when I listen	When I listen to English, I will be worried if I cannot comprehend every word.
113	I look for opportunities to speak English outside class	I find opportunities to practice my oral English in the spare time.
114	I have looked for opportunities to speak English recently	I have tried to find opportunities to practice my oral English recently.
115	I enjoy speaking English	I really like speaking English.
116	Accuracy is very important in speaking	When speaking English, accuracy is very important.
117	Making mistakes is OK when speaking	When speaking English, making errors is nothing important.
118	I know different ways of practicing speaking	I know different ways of practicing oral English.
119	I use different ways to practice speaking	I use different ways to practice oral English.
120	Speaking well is a talent some people have, but not all	Speaking English well is the endowment for some people, and not all the people possess this natural gift.
121	I can help myself to improve my level of speaking	I can help myself with the improvement of the level of oral English.
122	It is important to check one's writing	Re-examining my own writing is very important.
123	I check my writing	I examine my writing.
124	Last time I wrote in English I checked it myself	When I wrote in English last time, I examined my article.
125	I change the way I write according to who will read it	I change my writing methods according to different readers.
126	There are different types of writing	There are many kinds of articles.
127	I know my problems in writing	I know the problems existing in my writing.
128	I can help myself to improve my writing	I can help myself with the improvement of writing.
129	If I try my writing will get better	I can improve my writing if only I work hard.
130	My writing is better now than it was a year ago	My writing ability have improved compared with one year ago.
131	I enjoy writing in English	I like to write in English.
132	I guess the meaning of new words	I guess the meaning of new words.
133	I like learning new words	I like to learn new words.
134	I keep a record of new words	I make record of the new words.
135	I choose the best ways for me to learn new words	I choose the most appropriate way for me to learn new words.
136	I try to use new words outside class	I try to use new words out of class.
137	I try to use newly learned words in my essays	I try to use new words in my compositions.
138	I know techniques to help me remember vocabulary	I know the strategies to help me memorize English vocabulary.
139	I use techniques to help me remember vocabulary	I use some strategies to help me memorize English vocabulary.

140	To remember vocabulary you need to be talented	It needs some talent to memorize English vocabulary.
141	I know my problem areas in vocabulary	I know my difficulties in learning English vocabulary.
142	I fix my problems in vocabulary	I can solve the problems in my vocabulary learning.
143	I only learn words that a teacher recommends	I only learn the words recommended by the teacher.
144	I am able to decide myself which words are important to learn	I can decide which words are important and need to learn.
145	My vocabulary is better now than it was a year ago	My vocabulary is richer than one year ago.
146	Different types of text (magazine, letter, recipe etc.) have different vocabulary	Different kinds of articles use different kinds of vocabulary (magazines, letters, recipes, etc.)
147	I worry if I don't understand all the words in a text	I will be worried if I cannot understand the meaning of all the words in the articles.
148	I can help myself to improve my level of vocabulary	I can help myself improve the level of vocabulary.
149	If I try my vocabulary will get better	I will learn vocabulary better if only I work hard.
150	I worry if I don't understand all the grammar in a text	I will be worried if I cannot understand all the grammar in the text.
151	I try to find ways of practising grammar outside class	I try to find opportunities to practice English grammar out of class.
152	I know different ways of practising grammar	I know different methods to practice grammar.
153	Learning grammar is a talent some people have, but not all	Some people have the inborn gift to learn grammar, and not all the people have.
154	I know my problem areas in grammar	I know the problems existing in my grammar.
155	I can help myself to improve my level of grammar	I can help myself improve grammatical level.
156	If I try my grammar will get better	My grammatical level will improve if only I work hard.
157	My grammar is better now than it was a year ago	My grammatical level has made progress compared with one year ago.
158	I guess the meaning of new grammatical structures	I guess the meaning of new grammatical structures.
159	I like learning new grammar	I like to learn new grammar.
160	I keep a record of new grammar	I make record of the new-learned grammar.
161	I choose the best ways for me to learn new grammar	I choose the best way for me to learn grammar
162	I am able to decide which grammar is important to learn	I can decide what grammar is important and needs to learn.
163	I only learn grammar that a teacher recommends	I only learn grammar recommended by the teacher.
164	Different types of text (magazine, letter, recipe etc.) have different grammar	Different kinds of articles use different grammar (magazines, letters, recipes, etc.)
165	Different people have different ways of learning	Different people use different kinds of learning methods.
166	I know which sense to use to learn best (i.e. sight, or hearing, or touch, or physical movement)	I know which sense can best help me learn (e.g. visual, listening, touch or body movements)
167	I select learning techniques (i.e. taking notes, or drawing diagrams, or by listening, etc.) that suit my best way of learning	I choose the most appropriate learning strategies for my learning methods (e.g. taking notes, making diagrams, listening, etc.)
168	I think about the context for something new I am learning	I think about the linkage between the new-learned content and the context, or its background.
169	When I am learning something new I look for similarities with things I already know	When I learn new thing, I find the similarities existing between it and the things I have already known.

170	I use my knowledge from other subjects when I study English	When I am learning English, I employ the knowledge from other subjects.
171	I use my background knowledge when I do something new	I connect my background knowledge with the new-learned things.
172	I relate new things to my own personal experiences	I connect the new-learned things with my own experience.
173	I approach a topic in a careful, step by step manner.	I learn a subject step by step.
174	I consider facts and come to objective conclusions	I get the objective conclusion by considering the facts.
175	I look at causes and effects logically	I check the relations of cause and result by using logical methods.
176	I prefer a structured plan when I study	I am inclined to make an ordered learning plan when I am learning.
177	I follow textbooks as closely as possible	I use my textbook as closely as possible.
178	I collect all necessary information before I start	I collect all the information before I act.
179	I have a general idea for studying, then organise the details later	I have a general idea about learning, and then I organize the details.
180	When studying, I don't plan first	I do not make a plan before I start learning.
181	When I study I only use the textbooks	I only use textbook when I am learning.
182	I need time for personal reflection when I study	When I am learning, I need time to do self-reflection.
183	I am happy to use different worksheets from the rest of the class	It does no matter that the teacher asks me to do homework different from other classmates' .
184	I like negotiating with other students in class	I like to discuss with my classmates.
185	I like class discussions	I like the discussions in class.
186	I like working in pairs or small groups in class	I like pair work or group work in class.
187	I think learning English is more difficult for me than for the average learner	I feel that learning English is more difficult for me than for other students.
188	I am motivated to learn English	I am motivated to learn English.
189	I learn English because I have to	I learn English because I have to.
190	I do extra work	I do extra homework or learning out of class.
191	I think about what I have studied in class	I think about the things learned from the class.
192	I do my English homework	I finish my English homework.
193	The other students know English better than me	Other students learn English better than me.
194	The other students are more confident than me at speaking English	Other students speak English more confidently than me.
195	I worry that other students will laugh at me when I speak English	I am worried that when I speak English, some classmates will laugh at me.
196	I am confident I can learn English well	I am confident that I will learn English well.
197	I am determined about learning English	I am determined to learn English.
198	I hate to study with less than my best effort	I think that one should do his/her best to learn.
199	I always notice my mistakes	I can always be aware of my errors.
200	I try to find out how to learn better	I try to find out the method about how to learn better.
201	I have clear goals for improving my English	I have a very clear goal for improving my English.
202	I look for opportunities to practice English	I look for every opportunity to practice my English.
203	I organise my time for studying	I arrange my learning time soundly.
204	I use my mistakes to help me do better	I learn from the errors in order to do better in future.
205	I notice how other people use English	I pay attention to how other people use English.
206	I try to find the best environment for studying	I try to find the best learning environment.

207	I know the aim of the learning tasks I do	I know the purpose of my homework and exercise.
208	I know how much improvement I have made in the last six months	I know how much progress I have made in the last six months.
209	I think about my progress in learning English	I think about the progress of my English learning.
210	I try to relax when I am nervous about speaking English	When I feel nervous to speak English, I try to relax.
211	I avoid situations where there is a chance of making mistakes	I avoid making errors.
212	I talk to others about how I feel about learning English	I talk with other people about my feeling of learning English.
213	If someone is speaking English too fast I ask him/her to slow down or repeat	If someone speaks English too fast, I will ask her/him to speak more slowly or repeat.
214	I practice English with other students	I practice English with other students.
215	I ask for help from English speakers	I look for help from the English speakers.
216	I am aware of the feelings of others	I am aware of others' feelings.
217	I make learning plans	I make learning plans.
218	I join in with classroom discussions	I participate in the discussion in class.
219	I reflect on my learning	I summarize and reflect on my learning.
220	I am ready to learn in unfamiliar ways	I have made preparations for learning by using unfamiliar methods.
221	I want to learn in a more Western way	I want to learn in a more western-style way.
222	Repetition is important for learning	Repetition is very important for learning.
223	Errors must always be corrected	Errors must be corrected.
224	In learning it is important to work independently	Learning independently is very important.
225	Praise from the teacher is important to me	The approval from teacher is important for me.
226	I need tests to motivate me	I need quiz to motivate myself.
227	I need praise to motivate me	I need approval to motivate myself.
228	I motivate myself	I can motivate myself.
229	I can choose the method of learning that suits me best	I can choose the most appropriate learning method for myself.
230	My way of learning will never change	My learning method will always be unchanged.
231	I can study independently	I can learn independently.
232	My own needs are important to the way I learn	My own desire is important for my learning method.
233	I know how to check my own work for mistakes	I know how to examine my homework and find the errors out.
234	It is my job to check my work for mistakes	It is my responsibility to find errors from my homework.
235	Making mistakes is bad for language learning	Making errors is bad for learning language.
236	I am good at making choices	I am good at making choice.
237	I am an active dynamic person	I am enthusiastic and energetic.
238	I choose the exercises I work on	I choose my exercise.
239	I like to work at my own pace	I like learning in my own speed.
240	If I am not sure about an answer I go to the next question	If I am not sure about the answer to a question, I skip it and come to next one.
241	I do not go on to the next question in an exercise until I am sure about the answer	If I am not sure about the answer to a question, I will not continue to do the next one.
242	I guess answers if I don't know them for sure	If I am not sure about the answer to a question, I guess the answer.
243	I decide what I need to read	I choose my reading materials.
244	I like myself	I like myself.
245	I need the teacher to check my answers	I need teacher to examine my answers.
246	Memorizing answers is the best way to learn	Memorizing the answers is the best learning method.
247	If I do badly in a test I know why	I know the reason why I have not done well in the test.

248	I am motivated by making progress in learning	The progress in my learning motivates me to continue learning.
249	It is necessary to practice using English outside the classroom	It is necessary to practice English out of class.
250	Students can help the teacher choose the subject of lessons	Students can help teacher with the choice of course' s topics and contents.
251	I use real English texts (i.e. not made for students) in my learning	I use the authentic English articles when I am learning (viz. non-specifically written articles for the learners)
252	I know how to find information in a library	I know how to find materials in library.
253	I know how to use English language reference books (encyclopedias, dictionaries, etc.)	I know how to make use of English reference books (encyclopaedia, dictionary, etc.)
254	I know the parts of a book (index, glossary, contents, chapters)	I know each part of the book (index, table, content, chapter)
255	I keep a learning diary	I write learning diary.
256	I know how to find the information I need on the Internet	I know how to search information I need in the internet

Table 10.1: Full 256 items with back translation

10.2 Long List areas covered

Item	Reading	Writing	Listening	Speaking	Grammar	Vocabulary	Attitudes to learning	Social Interaction	Motivation	Confidence	Responsibility	Actions/ Behaviours	Strategies	Meta-cognition	Control	Skills
001. I will be happy when I can stop learning							X		X		X					
002. One day I will stop learning							X		X							
003. I want to learn something new every day							X		X							
004. Learning continues all your life							X		X							
005. I work hard							X		X							
006. When I study I take breaks in order to maintain my concentration													X	X		X
007. When I study I am organised													X	X		X
008. I am good at studying on my own							X	X	X	X						X
009. I know how to find the information I want									X	X			X	X		X
010. I give myself targets for studying							X		X	X	X		X	X	X	X
011. I am good at planning my learning									X	X	X		X	X	X	X
012. I watch TV or videos in English in my own time			X				X		X	X	X	X	X	X		X
013. When I learn something new I feel satisfaction in myself				X			X		X					X		
014. When I learn something new I feel good because the teacher is happy							X	X	X					X	X	
015. When I learn something new I don't feel good							X		X							
016. When I learn something new I feel good because I can stop learning it							X		X							
017. If I must finish a job at a certain time I finish early							X		X		X	X				
018. I meet deadlines							X		X		X	X				
019. The student's job is to remember the content of all lessons							X		X		X		X		X	
020. All teachers are equally good							X							X		
021. All lessons are equally valuable							X		X				X	X		
022. I know how to study										X			X	X	X	X
023. Students should always do what the teacher says							X		X		X				X	

Item	Reading	Writing	Listening	Speaking	Grammar	Vocabulary	Attitudes to Learning	Social Interaction	Motivation	Confidence	Responsibility	Actions/ Behaviours	Strategies	Meta-cognition	Control	Skills
024. The student's job is to develop as a person							x		x		x					
025. I know my own abilities										x			x	x		x
026. I feel lucky when I get good marks							x		x	x	x			x	x	
027. I need a teacher to help me							x		x	x	x			x	x	x
028. I feel unlucky when I get bad marks									x					x	x	
029. I know which is my best subject										x				x		
030. The teacher's job is to give me all the information							x		x		x				x	
031. The teacher's job is to help me learn							x									
032. The teacher's job is to control students							x				x				x	
033. Everybody can make progress if they try							x		x	x	x				x	
034. If I am not sure what I have to do, I don't worry about it							x			x	x				x	
035. If I am not sure what I have to do I ask somebody							x	x			x	x	x	x	x	
036. If I find a word that I don't know, I always ignore it and continue reading	x					x	x			x		x	x			
037. If I find a word that I don't know, I always ask a teacher first						x	x	x		x	x	x	x		x	x
038. If I find a word that I don't know, I look it up in a dictionary first						x	x		x	x	x	x	x		x	x
039. If I find a word that I don't know, I try to guess it						x	x		x	x	x	x	x		x	x
040. I usually need the teacher to help me with my learning							x			x	x				x	
041. I choose my own ways of studying							x		x	x	x		x		x	
042. I think about different ways of studying							x		x	x	x		x	x	x	x
043. When I have something to learn I try to think of different ways of doing it							x		x	x	x	x	x	x	x	x
044. I know my strong points and weak points										x				x		
045. I think about how I study best							x		x	x	x		x	x	x	x
046. I can describe the learning Strategies I use										x			x	x		
047. When I finish something I think about the ways I worked							x		x				x	x		x

Item	Skills	Control	Meta-cognition	Strategies	Actions/ Behaviours	Responsibility	Confidence	Motivation	Social Interaction	Attitudes to Learning	Vocabulary	Grammar	Speaking	Listening	Writing	Reading
048. When I finish something I think about ways to do it differently in the future	X	X	X	X	X	X	X	X		X						
049. I have changed the way I learn after thinking about it.	X	X	X	X	X	X	X	X		X						
050. I have tried different ways of learning	X	X	X	X						X						
051. I know why I have problems learning			X													
052. I try to fix problems I have in learning		X	X		X					X						
053. I know some different ways of learning	X		X													
054. I choose the best way to learn something	X	X	X	X						X						
055. I know why I did well or did badly	X		X							X						
056. I have made my own plans for my learning next week	X	X	X		X					X						
057. It is important to finish an exercise before my classmates					X				X	X						
058. I use the teacher's comments and corrections in my written work to improve my English			X		X				X	X				X		
059. When I like a learning activity, I know why I like it			X							X						
060. I always agree with what a teacher says		X	X		X				X	X						
061. I find it difficult to tell facts from opinions when reading			X	X			X							X		
062. Sara eats a lot of sweets so she must be fat	X															
063. My dictionary is always right with its definitions.	X	X			X	X				X	X					
064. Science books contain only facts	X															
065. I learn about all kinds of different things outside class		X		X		X		X		X						
066. Knowledge is something a teacher gives me		X	X	X			X	X	X	X						
067. Knowledge is something I construct for myself		X	X	X	X	X		X	X	X						
068. I trust the Internet	X															
069. In the last 4 months I have disagreed with something a teacher told the class	X	X	X				X		X							
070. The Great Wall of China can be seen from space	X															
071. If I learn something well, it is because I studied well	X	X	X			X	X									

Item	Reading	Writing	Listening	Speaking	Grammar	Vocabulary	Attitudes to Learning	Social Interaction	Motivation	Confidence	Responsibility	Actions/ Behaviours	Strategies	Meta-cognition	Control	Skills
072. If I learn something well, it is because my teacher taught well							x			x	x			x	x	
073. The teacher is responsible for my learning							x	x		x	x			x	x	
074. Memorization is the best way to learn							x					x	x	x		x
075. I am responsible for my learning							x		x	x	x			x	x	
076. I rely on the teacher							x			x	x				x	
077. I am self-reliant							x	x	x	x	x				x	
078. I enjoy making my own choices about learning							x		x	x				x	x	
079. I want to make my own choices about learning							x		x						x	
080. I make my own choices about learning							x		x	x	x				x	
081. It is good to make your own choices about learning							x									
082. I like the teacher to make the choices about learning							x	x	x	x	x				x	
083. My work is my own, not my teacher's							x			x	x					
084. My work is my teacher's, not mine							x			x	x					
085. I have a mature attitude to learning							x							x		
086. When I read an English text I need to understand every word in it	x					x	x			x			x	x	x	
087. I need to be sure about instructions							x			x					x	
088. If I am not sure about something it bothers me							x			x	x				x	
089. There is no one correct way to write an essay		x					x			x			x	x		x
090. I learn exclusively about college subjects							x		x				x			
091. Learning well is a talent that some people have and others do not have							x								x	
092. I can learn how to learn better							x		x	x	x				x	
093. Reading is a passive activity; the information passes from the page to you	x						x						x		x	x
094. Reading is an active activity	x						x				x		x		x	x
095. To read you must proceed word by word	x					x	x			x			x			x
096. There is one correct way of reading	x						x						x		x	x

Item	Reading	Writing	Listening	Speaking	Grammar	Vocabulary	Attitudes to Learning	Social Interaction	Motivation	Confidence	Responsibility	Actions/ Behaviours	Strategies	Meta-cognition	Control	Skills
097. I predict the content of a text (using pictures, headings, the context etc.)	x											x	x		x	x
098. I read newspapers in a different way to books	x												x		x	x
099. I sometimes look up words on the internet or in reference books						x			x	x	x	x	x		x	x
100. Last time I read an English text I predicted the content of it	x						x		x	x		x	x		x	x
101. Different types of text (novel, newspaper, web site etc.) are read in different ways	x												x			x
102. My general knowledge helps me to understand texts I read	x						x						x			x
103. It is best to read by starting at the beginning and reading line by line to the end	x						x			x			x		x	x
104. I read in English outside class	x						x		x	x		x			x	
105. When I read in English I think about what the source of the text is	x											x	x	x	x	x
106. I know the sources of the texts I read	x						x						x	x		x
107. When I read I think about the motives of the writer	x											x	x	x		x
108. When I read I start at the beginning and read line by line to the end	x						x			x		x	x		x	x
109. I predict the content before I listen			x				x			x		x	x		x	x
110. Every word is important for understanding a listening text			x			x	x			x			x		x	
111. The last time I listened to English I tried to predict the content			x				x					x	x	x	x	x
112. I worry if I don't understand everything when I listen			x				x			x			x		x	x
113. I look for opportunities to speak English outside class				x			x	x	x	x	x	x	x		x	
114. I have looked for opportunities to speak English recently				x			x	x	x	x	x	x	x		x	
115. I enjoy speaking English				x				x	x	x				x		
116. Accuracy is very important in speaking				x	x	x	x	x		x				x		
117. Making mistakes is OK when speaking				x	x	x	x	x		x			x	x		x
118. I know different ways of practising speaking				x								x	x	x		x
119. I use different ways to practise speaking				x			x		x	x	x	x	x	x	x	x
120. Speaking well is a talent some people have, but not all				x			x								x	

Item	Skills	Control	Meta-cognition	Strategies	Actions/ Behaviours	Responsibility	Confidence	Motivation	Social Interaction	Attitudes to learning	Vocabulary	Grammar	Speaking	Listening	Writing	Reading
121. I can help myself to improve my level of speaking	X	X	X	X	X	X	X			X			X			
122. It is important to check one's writing	X	X	X	X		X		X		X	X	X			X	
123. I check my writing		X	X	X	X	X	X	X		X	X	X			X	
124. Last time I wrote in English I checked it myself		X	X	X	X	X	X	X		X	X	X			X	
125. I change the way I write according to who will read it	X	X	X	X	X	X	X		X						X	
126. There are different types of writing	X														X	
127. I know my problems in writing						X									X	
128. I can help myself to improve my writing	X					X	X	X		X					X	
129. If I try my writing will get better		X				X	X	X		X					X	
130. My writing is better now than it was a year ago							X	X							X	
131. I enjoy writing in English					X										X	
132. I guess the meaning of new words		X				X	X	X								
133. I like learning new words			X					X								
134. I keep a record of new words	X		X	X	X	X		X								
135. I choose the best ways for me to learn new words	X		X	X	X	X	X	X								
136. I try to use new words outside class	X				X	X	X	X	X							
137. I try to use newly learned words in my essays	X				X	X	X	X	X					X		
138. I know techniques to help me remember vocabulary	X					X	X									
139. I use techniques to help me remember vocabulary	X						X	X								
140. To remember vocabulary you need to be talented		X														
141. I know my problem areas in vocabulary	X															
142. I fix my problems in vocabulary	X					X	X	X								
143. I only learn words that a teacher recommends	X					X	X	X	X							
144. I am able to decide myself which words are important to learn	X					X	X	X								
145. My vocabulary is better now than it was a year ago							X									

Item	Reading	Writing	Listening	Speaking	Grammar	Vocabulary	Attitudes to Learning	Social Interaction	Motivation	Confidence	Responsibility	Actions/ Behaviours	Strategies	Meta-cognition	Control	Skills
146. Different types of text (magazine, letter, recipe etc.) have different vocabulary						x								x		x
147. I worry if I don't understand all the words in a text	x					x	x			x			x	x	x	
148. I can help myself to improve my level of vocabulary						x	x		x	x	x	x	x	x	x	
149. If I try my vocabulary will get better						x	x		x	x	x			x	x	
150. I worry if I don't understand all the grammar in a text					x		x			x				x	x	
151. I try to find ways of practising grammar outside class					x		x		x	x	x		x	x	x	x
152. I know different ways of practising grammar					x					x			x	x	x	x
153. Learning grammar is a talent some people have, but not all					x		x								x	
154. I know my problem areas in grammar					x					x				x		
155. I can help myself to improve my level of grammar					x		x		x	x	x		x	x	x	x
156. If I try my grammar will get better					x		x		x	x	x			x	x	
157. My grammar is better now than it was a year ago					x					x				x		
158. I guess the meaning of new grammatical structures					x		x		x	x	x		x		x	
159. I like learning new grammar					x		x		x							
160. I keep a record of new grammar					x		x		x		x	x	x	x		
161. I choose the best ways for me to learn new grammar					x		x		x	x	x		x	x	x	x
162. I am able to decide which grammar is important to learn					x		x		x	x			x	x	x	
163. I only learn grammar that a teacher recommends					x		x	x	x	x	x	x	x	x	x	x
164. Different types of text (magazine, letter, recipe etc.) have different grammar					x								x			x
165. Different people have different ways of learning							x							x		
166. I know which sense to use to learn best (i.e. sight, or hearing, or touch, or physical movement)													x	x	x	x
167. I select learning techniques (i.e. taking notes, or drawing diagrams, or by listening, etc.) that suit my best way of learning											x		x	x	x	x
168. I think about the context for something new I am learning							x						x	x		x

Item	Skills	Control	Meta-cognition	Strategies	Actions/ Behaviours	Responsibility	Confidence	Motivation	Social Interaction	Attitudes to Learning	Vocabulary	Grammar	Speaking	Listening	Writing	Reading
169. When I am learning something new I look for similarities with things I already know	X		X	X												
170. I use my knowledge from other subjects when I study English	X		X	X						X						
171. I use my background knowledge when I do something new	X	X	X	X						X						
172. I relate new things to my own personal experiences	X		X	X						X						
173. I approach a topic in a careful, step by step manner.	X	X	X	X			X			X						
174. I consider facts and come to objective conclusions		X	X	X						X						
175. I look at causes and effects logically				X						X						
176. I prefer a structured plan when I study			X	X			X			X						
177. I follow textbooks as closely as possible		X		X	X		X			X						
178. I collect all necessary information before I start				X			X			X						
179. I have a general idea for studying, then organise the details later		X		X			X			X						
180. When studying, I don't plan first		X		X	X		X			X						
181. When I study I only use the textbooks		X		X		X	X	X		X						
182. I need time for personal reflection when I study		X		X						X						
183. I am happy to use different worksheets from the rest of the class				X	X		X	X		X						
184. I like negotiating with other students in class	X			X	X					X						
185. I like class discussions			X		X		X	X		X						
186. I like working in pairs or small groups in class				X	X		X	X		X						
187. I think learning English is more difficult for me than for the average learner			X				X	X		X						
188. I am motivated to learn English			X				X	X		X						
189. I learn English because I have to		X	X			X		X		X						
190. I do extra work		X		X	X		X	X		X						
191. I think about what I have studied in class		X		X	X			X		X						
192. I do my English homework		X		X	X		X	X		X						

Item	Reading	Writing	Listening	Speaking	Grammar	Vocabulary	Attitudes to Learning	Social Interaction	Motivation	Confidence	Responsibility	Actions/ Behaviours	Strategies	Meta-cognition	Control	Skills
193. The other students know English better than me								X	X	X				X	X	
194. The other students are more confident than me at speaking English				X				X		X				X		
195. I worry that other students will laugh at me when I speak English				X				X	X	X				X	X	
196. I am confident I can learn English well							X		X	X				X	X	
197. I am determined about learning English							X		X	X				X	X	
198. I hate to study with less than my best effort							X		X		X	X		X		
199. I always notice my mistakes					X	X					X	X		X		
200. I try to find out how to learn better							X		X	X	X		X	X	X	X
201. I have clear goals for improving my English							X		X	X	X		X	X	X	
202. I look for opportunities to practice English							X		X	X	X	X	X	X	X	
203. I organise my time for studying							X		X	X	X	X	X		X	
204. I use my mistakes to help me do better							X		X	X	X		X	X	X	
205. I notice how other people use English								X					X			
206. I try to find the best environment for studying							X		X		X		X	X	X	
207. I know the aim of the learning tasks I do											X		X	X		
208. I know how much improvement I have made in the last six months														X		
209. I think about my progress in learning English									X		X		X	X	X	
210. I try to relax when I am nervous about speaking English				X									X	X	X	
211. I avoid situations where there is a chance of making mistakes							X	X	X	X	X	X	X			
212. I talk to others about how I feel about learning English				X			X	X		X		X	X	X		
213. If someone is speaking English too fast I ask him/her to slow down or repeat			X	X				X	X	X	X	X	X		X	
214. I practice English with other students							X	X		X		X	X		X	
215. I ask for help from English speakers							X	X	X	X		X	X		X	
216. I am aware of the feelings of others								X						X		X
217. I make learning plans							X		X		X	X	X		X	
218. I join in with classroom discussions				X				X	X	X		X		X		X

Item	Skills	Control	Meta-cognition	Strategies	Actions/ Behaviours	Responsibility	Confidence	Motivation	Social Interaction	Attitudes to Learning	Vocabulary	Grammar	Speaking	Listening	Writing	Reading
219. I reflect on my learning		X				X		X		X						
220. I am ready to learn in unfamiliar ways			X	X	X		X	X		X						
221. I want to learn in a more Western way			X	X	X					X						
222. Repetition is important for learning				X	X		X			X						
223. Errors must always be corrected						X	X			X						
224. In learning it is important to work independently		X			X	X			X	X						
225. Praise from the teacher is important to me		X				X	X	X		X						
226. I need tests to motivate me		X				X	X	X		X						
227. I need praise to motivate me		X				X	X	X	X	X						
228. I motivate myself		X	X	X		X	X	X		X						
229. I can choose the method of learning that suits me best		X	X	X		X	X			X						
230. My way of learning will never change		X		X			X			X						
231. I can study independently		X	X	X	X		X	X	X	X						
232. My own needs are important to the way I learn		X	X				X	X		X						
233. I know how to check my own work for mistakes		X	X	X	X		X			X						
234. It is my job to check my work for mistakes					X		X	X		X	X	X		X		
235. Making mistakes is bad for language learning							X			X						
236. I am good at making choices		X	X				X	X								
237. I am an active dynamic person			X		X		X									
238. I choose the exercises I work on		X				X	X	X		X						
239. I like to work at my own pace		X	X	X												
240. If I am not sure about an answer I go to the next question				X	X		X			X						
241. I do not go on to the next question in an exercise until I am sure about the answer				X	X		X			X						
242. I guess answers if I don't know them for sure				X	X		X			X						
243. I decide what I need to read		X	X	X	X		X	X		X				X		
244. I like myself			X				X									

Item	Reading	Writing	Listening	Speaking	Grammar	Vocabulary	Attitudes to Learning	Social Interaction	Motivation	Confidence	Responsibility	Actions/ Behaviours	Strategies	Meta-cognition	Control	Skills
245. I need the teacher to check my answers							x	x		x	x	x	x	x	x	
246. Memorizing answers is the best way to learn							x			x			x	x		
247. If I do badly in a test I know why										x				x	x	
248. I am motivated by making progress in learning							x		x	x					x	
249. It is necessary to practice using English outside the classroom							x	x	x	x	x	x	x	x	x	
250. Students can help the teacher choose the subject of lessons							x	x		x	x	x			x	
251. I use real English texts (i.e. not made for students) in my learning	x						x			x		x	x			
252. I know how to find information in a library										x		x			x	x
253. I know how to use English language reference books (encyclopedias, dictionaries, etc.)	x					x				x		x			x	x
254. I know the parts of a book (index, glossary, contents, chapters)	x															x
255. I keep a learning diary							x		x			x		x	x	
256. I know how to find the information I need on the Internet	x						x			x		x	x			x
Total number of items in area	24	14	6	15	22	35	194	47	127	157	114	82	142	144	152	90
Percentage of total number of items	9	5	2	6	9	14	76	18	50	61	45	32	55	56	59	35

Table 10.2: The Long List items with areas covered.

10.3 Long List format C with translation

No.	Item	Translation
1	251. I use real English texts (i.e. not made for students) in my learning.	我在学习时，使用‘真实’的英语文章（即：非为学生专门写的）
2	095. To read you must proceed word by word.	在阅读时，我们得一个字一个字地读
3	086. When I read an English text I need to understand every word in it.	当我读一篇英语课文时，我需要弄懂其中每一个词的意思
4	100. Last time I read an English text I predicted the content of it.	我上次读英语文章的时候，我预测了它是在讲什么内容
5	147. I worry if I don't understand all the words in a text.	我如果不能理解文中所有词的意思，就会觉得担心
6	243. I decide what I need to read.	我决定我要读什么
7	109. I predict the content before I listen.	我在听一段课文之前会预测它的内容
8	110. Every word is important for understanding a listening text.	要听懂一段文章的内容，每一个字都很重要
9	111. The last time I listened to English I tried to predict the content.	我上次听英语的时候，我尝试了预测文章的内容
10	112. I worry if I don't understand everything when I listen.	我在听英语的时候，如果不能听懂每个词的意思，我就会很担心
11	125. I change the way I write according to who will read it.	我会根据读者的不同而改变我的写作方法
12	130. My writing is better now than it was a year ago.	我的写作比一年以前有进步
13	138. I know techniques to help me remember vocabulary.	我知道如何帮我记住英语词汇的学习技巧
14	140. To remember vocabulary you need to be talented.	要记住英语词汇，你得有些天赋
15	142. I fix my problems in vocabulary.	我解决我在词汇方面的问题
16	153. Learning grammar is a talent some people have, but not all.	有的人有学语法的天赋，但不是所有人都有
17	150. I worry if I don't understand all the grammar in a text.	我如果不能理解文中所有的语法，就会觉得担心
18	252. I know how to find information in a library.	我知道如何在图书馆找资料
19	253. I know how to use English language reference books (encyclopedias, dictionaries, etc.).	我知道如何使用英语参考书（百科全书，字典等）
20	254. I know the parts of a book (index, glossary, contents, chapters).	我知道书的各部分（索引，术语表，目录，章节）
21	256. I know how to find the information I need on the Internet.	我知道如何在网上搜索我要的信息
22	187. I think learning English is more difficult for me than for the average learner.	我觉得学英语对我来说，比一般的学生要难
23	193. The other students know English better than me.	其他同学英语比我学得好
24	194. The other students are more confident than me at speaking English.	别的同学讲英语时比我更加有自信
25	195. I worry that other students will laugh at me when I speak English.	我担心我讲英语时，别的同学会笑我
26	196. I am confident I can learn English well.	我自信我能学好英语
27	017. If I must finish a job at a certain time I finish early	如果我必须在规定时间内完成一个任务，我会尽早完成
28	203. I organise my time for studying.	我合理安排学习的时间
29	238. I choose the exercises I work on.	我选择我要做的练习
30	229. I can choose the method of learning that suits me best.	我能选择最适合我的学习方法
31	236. I am good at making choices.	我很善于做选择
32	220. I am ready to learn in unfamiliar ways.	我愿意尝试新的方法学习
33	023. Students should always do what their teacher says.	学生应该总根据老师说的做

34	030. The teacher's job is to give me all the information.	老师的任务是给我所有的相关信息
35	076. I rely on the teacher when learning.	我的学习靠老师
36	008. I am good at studying on my own.	我善于独立学习
37	231. I can study independently.	我能独立地学习
38	234. It is my job to check my work for mistakes.	从我的作业中找出错误是我的责任
39	055. I know why I did well or did badly.	我知道我为什么做得好或做得差
40	230. My way of learning will never change.	我的学习方法将永远不会改变
41	026. I feel lucky when I get good marks.	我得到好成绩时，觉得自己很幸运.
42	246. Memorizing answers is the best way to learn.	记住答案是最好的学习方法
43	046. I can describe the learning strategies I use.	我能描述我使用的学习策略
44	049. I have changed the way I learn after thinking about it.	在我思考了学习方法之后我改变了我的学习方法
45	175. I look at causes and effects logically.	我用逻辑的方法来看原因和结果的关系
46	237. I am an active dynamic person.	我是一个积极主动，充满活力的人
47	212. I talk to others about how I feel about learning English.	我与他人谈论对于英语学习的感受
48	189. I learn English because I have to.	我学英语，因为我不得不学
49	021. All lessons are equally valuable.	所有的课都一样有价值
50	205. I notice how other people use English.	我留意别人是怎样使用英语的

Table 10.3: Long List Format C with translation

10.4 Selection Table

Item	Likert Responses as Percentage							Selection Criteria				
	1	2	3	4	5	6	7	A	B	C	D	E
93. Reading is a passive activity; the information passes from the page to you.	3.0	7.9	20.8	6.9	35.6	12.9	12.9	4	13.9	1.597	19.497	4
195. I worry that other students will laugh at me when I speak English.	2.0	5.1	26.3	17.2	27.3	10.1	12.1	5	10.1	1.613	16.713	4
189. I learn English because I have to.	5.8	7.8	27.2	19.4	27.2	1.9	10.7	4	7.6	1.562	13.162	4
194. The other students are more confident than me at speaking English.	4.3	10.0	32.9	17.1	24.3	7.1	4.3	4	7.2	1.545	12.745	4
112. I worry if I don't understand everything when I listen.	0	14.1	24.2	17.2	32.3	7.1	5.1	4	7	1.424	12.424	4
120. Speaking well (in English) is a talent some people have, but not all.	10.7	12.0	34.7	16.0	20.0	1.3	5.3	5	4	1.510	10.510	4
21. All lessons are equally valuable	0	2.7	13.3	12.0	38.7	10.7	22.7	5	1.3	1.387	7.687	4
91. Learning well is a talent that some people have and others do not have.	1.4	6.8	32.4	17.6	18.9	13.5	9.5	4	1.3	1.506	6.806	4
164. Different types of text (magazine, letter, recipe etc.) have different grammar.	0	1.0	10.2	9.2	50.0	15.3	14.3	4	1	1.503	6.503	4
96. There is one correct way of reading.	0	2.7	10.8	2.7	47.3	14.9	21.6	4	8.1	1.293	13.393	3
140. To remember vocabulary you need to be talented.	2.8	8.3	36.1	19.4	22.2	4.2	6.9	3	2.8	1.406	7.206	3
130. My writing is better now than it was a year ago.	0	0	12.7	9.9	40.8	28.2	8.5	3	2.8	1.379	7.179	3
61. I find it difficult to tell facts from opinions when reading.	0	0	23.3	20.5	38.4	9.6	8.2	3	2.8	1.295	7.095	3
243. I decide what I need to read.	3.1	7.1	11.2	14.3	39.8	13.3	11.2	5		1.716	6.716	3
251. I use real English texts (i.e. not made for students) in my learning.	0	0	10.5	23.2	38.9	15.8	11.6	5		1.666	6.666	3
254. I know the parts of a book (index, glossary, contents, chapters).	0	0	12.7	19.7	36.6	12.7	18.3	5		1.384	6.384	3
187. I think learning English is more difficult for me than for the average learner.	1.4	2.7	20.3	20.3	33.8	10.8	10.8	5		1.365	6.365	3
70. The Great Wall of China can be seen from space.	7.6	1.5	12.1	13.6	34.8	9.1	21.2	4		1.998	5.998	3
180. When studying, I don't plan first	1.3	6.7	21.3	20.0	37.3	5.3	8.0	3	1.3	1.349	5.649	3

Item	Likert Responses as Percentage							Selection Criteria				
	1	2	3	4	5	6	7	A	B	C	D	E
30. The teacher's job is to give me all the information.	5.0	14.0	13.0	26.0	28.0	10.0	4.0	4		1.530	5.530	3
111. The last time I listened to English I tried to predict the content.	0	0	11.7	11.7	56.4	14.9	5.3	4		1.503	5.503	3
175. I look at causes and effects logically.	0	0	1.4	26.8	45.1	15.5	11.3	4		1.486	5.486	3
14. When I learn something new I feel good because the teacher is happy	3.9	13.2	15.8	25.0	28.9	9.2	3.9	4		1.441	5.441	3
230. My way of learning will never change.	1.0	1.0	5.0	16.0	44.0	13.0	20.0	4		1.418	5.418	3
17. If I must finish a job at a certain time I finish early	2.7	2.7	18.9	24.3	27.0	14.9	9.5	4		1.416	5.416	3
90. I learn exclusively about college subjects, and nothing else.	1.9	1.9	11.4	13.3	21.9	14.3	5.7	4		1.397	5.397	3
100. Last time I read an English text I predicted the content of it.	0	1.0	13.3	15.3	52.0	12.2	6.1	4		1.394	5.394	3
153. Learning grammar is a talent some people have, but not all.	5.4	10.8	27.0	25.7	23.0	4.1	4.1	4		1.388	5.388	3
236. I am good at making choices.	2.8	5.6	20.8	27.8	27.8	12.5	2.8	4		1.381	5.381	3
32. The teacher's job is to control students in the classroom.	0	1.3	13.2	19.7	30.3	13.2	22.4	4		1.374	5.374	3
26. I feel lucky when I get good marks.	6.7	17.3	24.0	24.0	22.7	4.0	1.3	4		1.368	5.368	3
166. I know which sense is best for me to use when learning (i.e. sight, or hearing, or touch, or physical movement).	0	0	4.2	18.1	40.3	20.8	16.7	4		1.368	5.368	3
116. Accuracy is very important in speaking English.	3.9	9.7	33.0	24.3	15.5	11.7	1.9	4		1.358	5.358	3
205. I notice how other people use English.	0	1.4	4.1	17.8	45.2	20.5	11.0	4		1.330	5.330	3
246. Memorizing answers is the best way to learn.	0	2.7	8.2	12.3	35.6	19.2	21.9	4		1.313	5.313	3
1. I will be happy when I can stop learning	0	2.6	5.3	10.5	36.8	14.5	30.3	4		1.311	5.311	3
237. I am an active dynamic person.	0	0	8.3	26.4	34.7	15.3	15.3	4		1.306	5.306	3
150. I worry if I don't understand all the grammar in a text.	1.0	9.7	39.8	16.5	26.2	5.8	1.0	3	9.7	1.194	13.894	2
124. Last time I wrote in English I checked it myself.	0	0	12.0	5.3	41.3	24.0	17.3	3	6.7	1.183	10.883	2
34. When I study English, if I am not sure what I have to do, I don't worry about it.	2.0	1.0	15.7	10.8	53.9	8.8	7.8	3	4.9	1.294	9.194	2
252. I know how to find information in a library.	1.4	1.4	6.8	4.1	50.7	21.9	13.7	3	2.7	1.181	6.881	2

Item	Likert Responses as Percentage							Selection Criteria				
	1	2	3	4	5	6	7	A	B	C	D	E
256. I know how to find the information I need on the Internet.	1.4	0	6.8	4.1	50.0	20.3	17.6	3	2.7	1.160	6.860	2
74. Memorization is the best way to learn.	2.9	11.7	26.2	30.1	23.3	2.9	2.9	5		1.255	6.255	2
23. Students should always do what their teacher says.	0	0	13.2	23.7	36.8	11.8	14.5	5		1.213	6.213	2
8. I am good at studying on my own.	0	0	7.8	11.7	39.0	23.4	18.2	5		1.141	6.141	2
2. One day I will stop learning.	0	1.3	9.3	13.3	26.7	24.0	25.3	4		1.566	5.566	2
33. Everybody can make progress if they try.	0	0	5.3	3.9	27.6	25.0	38.2	3	1.4	1.135	5.535	2
125. I change the way I write according to who will read it.	0	1.4	14.9	24.3	36.5	14.9	8.1	4		1.288	5.288	2
138. I know techniques to help me remember vocabulary.	0	0	15.1	23.3	43.8	5.5	12.3	4		1.279	5.279	2
117. Making mistakes is OK when speaking English.	0	0	7.9	12.9	46.5	19.8	12.9	4		1.278	5.278	2
234. It is my job to check my work for mistakes.	0	0	3.1	21.6	50.5	14.4	10.3	4		1.277	5.277	1
226. I need tests to motivate me.	6.9	22.5	35.3	20.6	12.7	0.0	2.0	4		1.263	5.263	1
38. If I find an English word that I don't know, I look it up in a dictionary first.	3.8	17.3	26.0	28.8	20.2	1.9	1.9	4		1.259	5.259	1
16. When I learn something new I feel good because I can stop learning it	0	1.3	4.0	12.0	38.7	28.0	16.0	4		1.252	5.252	1
36. If I find an English word that I don't know, I always ignore it and continue reading.	1.3	14.5	21.1	22.4	35.5	3.9	1.3	4		1.247	5.247	1
86. When I read an English text I need to understand every word in it.	2.9	2.9	12.6	21.4	44.7	10.7	4.9	4		1.243	5.243	1
244. I like myself.	0	0	1.4	11.3	32.4	26.8	28.2	4		1.240	5.240	1
114. I have looked for opportunities to speak English recently.	0	0	9.0	14.0	55.0	11.0	11.0	4		1.236	5.236	1
143. I only learn words that a teacher recommends.	1.0	0	7.8	18.6	43.1	18.6	10.8	4		1.232	5.232	1
136. I try to use new words outside class.	0	0	6.0	27.0	43.0	14.0	10.0	4		1.230	5.230	1
75. I am responsible for my learning.	0	0	1.3	0	21.3	36.0	41.3	3	1.3	0.855	5.155	1
12. I watch TV or videos in English in my own time	0	1.0	3.8	2.9	31.7	36.5	24.0	3	0.9	1.049	4.949	1
92. I can learn how to learn better.	8.9	6.9	9.9	5.0	31.7	21.8	15.8	3		1.868	4.868	1
162. I am able to decide which grammar is important to learn.	1.1	1.1	7.4	16.8	52.6	18.9	2.1	3		1.559	4.559	1

Item	Likert Responses as Percentage							Selection Criteria				
	1	2	3	4	5	6	7	A	B	C	D	E
229. I can choose the method of learning that suits me best.	0	0	3.2	21.1	50.5	15.8	9.5	3		1.509	4.509	1
163. I only learn grammar that a teacher recommends.	0	2.8	9.9	28.2	38.0	4.2	16.9	3		1.490	4.490	1
73. The teacher is responsible for my learning.	0	5.5	23.3	32.9	24.7	6.8	6.8	3		1.408	4.408	1
82. I like the teacher to make the choices about learning.	2.0	5.1	14.3	27.6	38.8	9.2	3.1	3		1.399	4.399	1
57. It is important to finish an exercise before my classmates.	4.1	6.8	21.9	35.6	21.9	4.1	5.5	3		1.398	4.398	1
154. I know my problem areas in grammar.	0	0	8.2	15.3	48.0	21.4	7.1	3		1.384	4.384	1
56. I have made my own plans for my learning next week.	0	4.0	24.0	26.0	32.0	6.0	8.0	3		1.380	4.380	1
144. I am able to decide myself which words are important to learn.	0	2.0	8.0	25.0	47.0	10.0	8.0	3		1.326	4.326	1
247. If I do badly in a test I know why.	0	0	5.6	9.7	54.2	16.7	13.9	3		1.305	4.305	1
66. Knowledge is something a teacher gives me.	13.3	9.3	30.7	34.7	8.0	2.7	1.3	3		1.300	4.300	1
119. I use different ways to practise speaking English.	0	1.0	9.2	35.7	35.7	9.2	9.2	2		1.345	3.345	1
240. If I am not sure about an answer I go to the next question.	0	0	5.0	13.0	55.0	14.0	13.0	4		1.213	5.213	0
178. I collect all necessary information before I start.	1.4	0	12.2	31.1	37.8	13.5	4.1	4		1.200	5.200	0
176. I prefer a structured plan when I study.	1.3	0	9.3	16.0	38.7	24.0	10.7	4		1.196	5.196	0
10. I give myself targets for studying.	0	0	5.3	17.1	40.8	25.0	11.8	4		1.189	5.189	0
11. I am good at planning my learning.	0	1.3	14.7	18.7	38.7	18.7	8.0	4		1.178	5.178	0
197. I am determined about learning English.	1.0	0	4.0	16.8	29.7	30.7	17.8	4		1.173	5.173	0
135. I choose the best ways for me to learn new words.	0	0	1.0	17.2	56.6	11.1	14.1	4		1.171	5.171	0
48. When I finish something I think about ways to do it differently in the future.	1.3	0	17.3	26.7	33.3	17.3	4.0	4		1.164	5.164	0
83. The work I do for my course is mine, not my teachers'.	0	1.3	3.9	13.2	43.4	15.8	22.4	4		1.163	5.163	0
218. I join in with classroom discussions.	0	0	5.0	19.8	40.6	22.8	11.9	4		1.154	5.154	0
18. I meet deadlines	0	1.3	6.6	19.7	43.4	14.5	14.5	4		1.147	5.147	0
185. I like class discussions.	0	1.0	8.7	23.3	39.8	15.5	11.7	4		1.141	5.141	0
68. I always trust the information I find on the Internet.	0	0	9.5	27.0	32.4	20.3	10.8	4		1.140	5.140	0
145. My vocabulary is better now than it was a year ago.	0	0	1.4	12.3	47.9	21.9	16.4	4		1.136	5.136	0

Item	Likert Responses as Percentage							Selection Criteria				
	1	2	3	4	5	6	7	A	B	C	D	E
110. Every word is important for understanding a listening text.	0	5.9	12.7	12.7	52.9	12.7	2.9	4		1.134	5.134	0
210. I try to relax when I am nervous about speaking English.	1.4	1.4	4.1	10.8	56.8	12.2	13.5	4		1.130	5.130	0
71. If I learn something well, it is because I studied well.	0	1.4	5.4	18.9	40.5	20.3	13.5	4		1.127	5.127	0
198. I hate to study with less than my best effort.	0	0	2.7	12.3	32.9	21.9	30.1	4		1.123	5.123	0
239. I like to work at my own pace.	0	0	5.9	16.8	34.7	25.7	16.8	4		1.120	5.120	0
235. Making mistakes is bad for language learning.	16.4	17.8	47.9	13.7	2.7	0.0	1.4	4		1.118	5.118	0
225. Praise from the teacher is important to me.	17.2	25.3	39.4	12.1	5.1	1.0	0.0	4		1.117	5.117	0
29. I know which is my best subject.	0	0	4.0	14.7	40.0	18.7	22.7	4		1.116	5.116	0
213. If someone is speaking English too fast I ask him/her to slow down or repeat.	1.0	0	5.9	17.6	47.1	15.7	12.7	4		1.114	5.114	0
129. If I try my writing will get better.	0	0	5.4	10.8	40.5	21.6	21.6	4		1.111	5.111	0
227. I need praise to motivate me.	19.4	22.3	36.9	18.4	1.9	1.0	0.0	4		1.110	5.110	0
95. To read you must proceed word by word.	0	0	12.7	15.7	52.0	7.8	11.8	4		1.104	5.104	0
121. I can help myself to improve my level of speaking.	0	0	10.7	17.5	52.4	5.8	13.6	4		1.101	5.101	0
183. I am happy to use different worksheets from the rest of the class.	0	1.4	15.1	23.3	42.5	12.3	5.5	4		1.096	5.096	0
115. I enjoy speaking English.	0	0	5.8	19.4	39.8	20.4	14.6	4		1.091	5.091	0
88. If I am not sure about something it bothers me.	14.9	27.0	39.2	12.2	6.8	0.0	0.0	4		1.084	5.084	0
200. I try to find out how to learn better.	0	0	3.0	17.0	48.0	21.0	11.0	4		1.081	5.081	0
47. When I finish something I think about the ways I worked.	0	1.3	16.0	28.0	37.3	13.3	4.0	4		1.080	5.080	0
7. When I study, I am an organised learner.	0	1.0	10.5	28.6	39.0	14.3	6.7	4		1.072	5.072	0
31. The teacher's job is to help me learn	0	0	9.2	15.8	46.1	18.4	10.5	4		1.070	5.070	0
202. I look for opportunities to practice English.	0	0	6.8	14.6	47.6	16.5	14.6	4		1.070	5.070	0
37. If I find an English word that I don't know, I always ask a teacher first.	0	0	1.0	10.1	49.5	23.2	16.2	4		1.062	5.062	0
212. I talk to others about how I feel about learning English.	0	0	6.8	11.0	53.4	16.4	12.3	4		1.014	5.014	0
133. I like learning new words.	0	0	4.9	17.5	50.5	15.5	11.7	4		0.993	4.993	0
182. I need time for personal reflection when I study.	0	0	5.0	11.9	51.5	18.8	12.9	4		0.989	4.989	0

Item	Likert Responses as Percentage							Selection Criteria				
	1	2	3	4	5	6	7	A	B	C	D	E
9. I know how to find the information I want.	0	0	3.9	15.8	53.9	13.2	13.2	4		0.981	4.981	0
188. I am motivated to learn English.	0	0	1.0	16.5	42.7	23.3	16.5	4		0.981	4.981	0
25. I know my own ability.	0	0	5.3	11.8	47.4	25.0	10.5	4		0.978	4.978	0
148. I can help myself to improve my level of vocabulary.	0	0	3.9	14.7	55.9	11.8	13.7	4		0.976	4.976	0
221. I want to learn in a more learner-centred way.	0	0	4.2	13.9	50.0	20.8	11.1	4		0.963	4.963	0
27. I need a teacher to help me.	0	0	3.0	15.0	47.0	23.0	12.0	4		0.960	4.960	0
65. I learn about all kinds of different things outside class.	0	0	1.3	14.5	50.0	18.4	15.8	4		0.958	4.958	0
181. When I study I only use the textbooks	0	0	4.1	10.8	55.4	17.6	12.2	4		0.944	4.944	0
172. I relate new things to my own personal experiences.	0	0	1.4	12.3	39.7	31.5	15.1	4		0.944	4.944	0
6. When I study I take breaks in order to maintain my concentration.	0	0	1.9	10.5	45.7	26.7	15.2	4		0.939	4.939	0
41. I choose my own ways of studying English.	0	0	0	15.5	41.7	28.2	14.6	4		0.924	4.924	0
214. I practice English with other students.	0	0	1.0	18.4	53.4	14.6	12.6	4		0.919	4.919	0
255. I keep a learning diary.	5.9	6.9	40.6	26.7	13.9	5.0	1.0	3		1.290	4.290	0
193. The other students know English better than me.	0	5.8	37.7	26.1	21.7	7.2	1.4	3		1.283	4.283	0
108. When I read I start at the beginning and read line by line to the end.	1.4	4.1	24.3	21.6	35.1	8.1	5.4	3		1.271	4.271	0
55. I know why I did well or did badly.	0	0	5.5	15.1	50.7	19.2	9.6	3		1.268	4.268	0
157. My grammar is better now than it was a year ago.	0	1.4	19.2	35.6	27.4	9.6	6.8	3		1.259	4.259	0
208. I know how much improvement I have made in the last six months.	0	4.2	5.6	20.8	47.2	13.9	8.3	3		1.258	4.258	0
24. The student's job is to develop as a person.	0	1.9	8.7	28.2	31.1	21.4	8.7	3		1.258	4.258	0
64. Science books contain only facts.	0	4.1	15.1	17.8	49.3	8.2	5.5	3		1.252	4.252	0
161. I choose the best ways for me to learn new grammar.	1.0	0	7.1	17.2	48.5	19.2	7.1	3		1.251	4.251	0
105. When I read in English I think about what the source of the text is.	0	1.0	27.7	28.7	29.7	5.9	6.9	3		1.246	4.246	0
49. I have changed the way I learn after thinking about it.	0	1.4	5.4	24.3	47.3	16.2	5.4	3		1.245	4.245	0
223. Errors must always be corrected.	21.8	28.7	34.7	6.9	5.0	3.0	0.0	3		1.241	4.241	0
42. I think about different ways of studying English.	0	0	3.0	24.0	50.0	16.0	7.0	3		1.224	4.224	0

Item	Likert Responses as Percentage							Selection Criteria				
	1	2	3	4	5	6	7	A	B	C	D	E
46. I can describe the learning strategies I use.	0	0	9.1	19.2	53.5	14.1	4.0	3		1.223	4.223	0
179. I have a general idea for studying, then organise the details later.	0	0	4.1	11.0	53.4	23.3	8.2	3		1.223	4.223	0
211. I avoid situations where there is a chance of making mistakes.	6.1	12.1	54.5	16.2	5.1	5.1	1.0	3		1.220	4.220	0
173. I approach a topic in a careful, step by step manner.	1.4	1.4	5.4	14.9	52.7	17.6	6.8	3		1.214	4.214	0
159. I like learning new grammar.	0	0	16.8	24.8	43.6	9.9	5.0	3		1.211	4.211	0
72. If I learn something well, it is because my teacher taught well.	2.7	1.4	16.4	46.6	27.4	4.1	1.4	3		1.202	4.202	0
241. I do not go on to the next question in an exercise until I am sure about the answer.	11.0	9.6	47.9	19.2	9.6	1.4	1.4	3		1.202	4.202	0
45. I think about how I study best.	0	0	4.1	9.5	37.8	31.1	17.6	3		1.198	4.198	0
20. All teachers are equally good	0	1.3	3.9	15.8	40.8	14.5	23.7	3		1.195	4.195	0
147. I worry if I don't understand all the words in a text.	0	8.1	23.0	23.0	36.5	6.8	2.7	3		1.190	4.190	0
242. I guess answers if I don't know them for sure.	1.0	0	5.9	17.8	53.5	16.8	5.0	3		1.181	4.181	0
106. I know the sources of the texts I read.	2.8	1.4	44.4	31.9	15.3	0.0	4.2	3		1.179	4.179	0
40. I usually need the teacher to help me with my English language learning.	1.0	8.9	30.7	28.7	24.8	3.0	3.0	3		1.177	4.177	0
5. I work hard to learn English.	1.0	1.0	3.8	5.7	37.1	27.6	23.8	3		1.177	4.177	0
203. I organise my time for studying.	0	2.7	6.8	21.6	51.4	10.8	6.8	3		1.175	4.175	0
199. I always notice my mistakes.	0	4.0	15.8	27.7	39.6	9.9	3.0	3		1.171	4.171	0
245. I need the teacher to check my answers.	2.0	6.1	31.3	34.3	19.2	4.0	3.0	3		1.166	4.166	0
215. I ask for help from English speakers.	0	0	5.9	14.9	56.4	13.9	8.9	3		1.166	4.166	0
142. I fix my problems in vocabulary.	0	1.4	8.3	29.2	41.7	13.9	5.6	3		1.165	4.165	0
201. I have clear goals for improving my English.	0	0	4.0	25.0	45.0	19.0	7.0	3		1.165	4.165	0
171. I use my background knowledge when I do something new.	0	0	2.0	7.9	52.5	23.8	13.9	3		1.160	4.160	0
158. I guess the meaning of new grammatical structures.	1.0	0	9.0	20.0	55.0	12.0	3.0	3		1.155	4.155	0
15. When I learn something new I don't feel good	0	0	1.4	2.7	39.2	27.0	29.7	3		1.155	4.155	0
233. I know how to check my own work for mistakes.	0	0	11.1	23.2	52.5	8.1	5.1	3		1.146	4.146	0

Item	Likert Responses as Percentage							Selection Criteria				
	1	2	3	4	5	6	7	A	B	C	D	E
167. I select learning techniques (i.e. taking notes, or drawing diagrams, or by listening, etc.) that suit my best way of learning.	0	0	1.4	9.7	54.2	15.3	19.4	3		1.145	4.145	0
186. I like working in pairs or small groups in class.	0	3.0	5.0	27.7	38.6	15.8	9.9	3		1.139	4.139	0
156. If I try my grammar will get better.	0	0	2.7	9.6	47.9	24.7	15.1	3		1.136	4.136	0
19. The student's job is to remember the content of all lessons	1.0	1.0	20.4	26.2	42.7	1.9	6.8	3		1.133	4.133	0
151. I try to find ways of practising grammar outside class.	0	0	10.9	27.7	48.5	8.9	4.0	3		1.125	4.125	0
160. I keep a record of new grammar.	2.0	1.0	6.9	18.6	49.0	16.7	5.9	3		1.120	4.120	0
109. I predict the content before I listen.	0	0	9.0	21.0	49.0	13.0	8.0	3		1.117	4.117	0
141. I know my problem areas in vocabulary.	0	1.4	6.8	16.4	52.1	19.2	4.1	3		1.114	4.114	0
128. I can help myself to improve my writing.	0	0	8.0	30.0	48.0	10.0	4.0	3		1.107	4.107	0
177. I follow textbooks as closely as possible.	5.4	8.1	50.0	23.0	10.8	1.4	1.4	3		1.091	4.091	0
168. I think about the context for something new I am learning.	0	0	2.7	9.6	52.1	24.7	11.0	3		1.083	4.083	0
89. There is no one correct way to write an essay.	1.3	0	1.3	6.7	46.7	22.7	21.3	3		1.083	4.083	0
248. I am motivated by making progress in learning.	0	0	1.0	2.0	29.7	31.7	35.6	3		1.083	4.083	0
63. My dictionary is always right with its definitions.	2.7	0	20.0	24.0	48.0	2.7	2.7	3		1.082	4.082	0
122. It is important to check one's writing.	0	0	8.0	8.0	42.7	26.7	14.7	3		1.080	4.080	0
76. I rely on the teacher when learning.	0	1.4	8.1	17.6	48.6	14.9	9.5	3		1.078	4.078	0
3. I want to learn something new every day.	0	0	2.6	6.5	22.1	26.0	42.9	3		1.076	4.076	0
59. When I like a learning activity, I know why I like it.	0	0	1.4	6.8	52.7	24.3	14.9	3		1.075	4.075	0
113. I look for opportunities to speak English outside class.	0	0	4.9	21.6	52.9	10.8	9.8	3		1.074	4.074	0
35. When I study English, if I am not sure what I have to do I ask somebody.	0	0	5.9	16.7	53.9	14.7	8.8	3		1.071	4.071	0
51. I know why I have problems learning.	0	0	6.8	17.6	58.1	10.8	6.8	3		1.070	4.070	0
170. I use my knowledge from other subjects when I study English.	0	0	3.0	8.9	46.5	28.7	12.9	3		1.067	4.067	0
101. Different types of text (novel, newspaper, web site etc.) are read in different ways.	0	0	1.0	9.8	40.2	32.4	16.7	3		1.065	4.065	0

Item	Likert Responses as Percentage							Selection Criteria				
	1	2	3	4	5	6	7	A	B	C	D	E
87. I need to be sure about the instructions for learning activities.	6.8	14.9	47.3	23.0	6.8	0.0	1.4	3		1.064	4.064	0
192. I do my English homework.	0	0	4.9	7.8	45.6	20.4	21.4	3		1.064	4.064	0
81. It is good to make your own choices about learning.	0	0	2.7	9.3	36.0	25.3	26.7	3		1.061	4.061	0
217. I make learning plans.	0	1.9	8.7	23.3	41.7	19.4	4.9	3		1.061	4.061	0
149. If I try my vocabulary will get better.	0	0	0	0	39.2	33.8	27.0	3		1.053	4.053	0
80. I make my own choices about learning.	0	0	3.9	9.2	43.4	22.4	21.1	3		1.052	4.052	0
232. My own needs are important to the way I learn.	0	0	0	2.7	46.6	31.5	19.2	3		1.046	4.046	0
44. I know my strong points and weak points related to learning English.	0	0	1.0	4.9	46.1	29.4	18.6	3		1.036	4.036	0
54. I choose the best way to learn something.	0	0	8.0	20.0	46.7	16.0	9.3	3		1.033	4.033	0
84. The work I do for my course is my teacher's, not mine.	0	0	1.3	8.0	34.7	25.3	30.7	3		1.025	4.025	0
196. I am confident I can learn English well.	0	0	2.0	7.9	34.7	28.7	26.7	3		1.015	4.015	0
43. When I have something to learn I try to think of different ways of doing it.	0	0	3.9	21.6	48.0	21.6	4.9	3		1.014	4.014	0
253. I know how to use English language reference books (encyclopedias, dictionaries, etc.).	0	1.4	4.1	9.5	54.1	18.9	12.2	3		1.010	4.010	0
103. It is best to read by starting at the beginning and reading line by line to the end.	0	1.9	13.6	34.0	36.9	11.7	1.9	3		0.999	3.999	0
107. When I read I think about the motives of the writer.	0	0	13.7	20.6	50.0	10.8	4.9	3		0.997	3.997	0
184. I like negotiating with other students in class.	0	0	7.0	22.0	50.0	12.0	9.0	3		0.993	3.993	0
104. I read in English outside class.	0	0	3.9	7.8	50.0	21.6	16.7	3		0.987	3.987	0
134. I keep a record of new words.	1.0	0	5.8	27.2	49.5	10.7	5.8	3		0.984	3.984	0
250. Students can help the teacher choose the subject of lessons.	0	0	3.0	22.0	52.0	18.0	5.0	3		0.984	3.984	0
62. Sara eats a lot of sweets so she must be fat.	0	0	6.8	14.9	58.1	9.5	10.8	3		0.979	3.979	0
131. I enjoy writing in English.	0	1.0	12.9	38.6	34.7	9.9	3.0	3		0.976	3.976	0
78. I enjoy making my own choices about learning.	0	0	1.3	7.9	39.5	27.6	23.7	3		0.976	3.976	0
58. I use the teacher's comments and corrections in my written work to improve my English.	0	0	2.9	9.7	45.6	26.2	15.5	3		0.965	3.965	0

Item	Likert Responses as Percentage							Selection Criteria				
	1	2	3	4	5	6	7	A	B	C	D	E
123. I check my writing.	0	0	4.0	8.0	49.3	25.3	13.3	3		0.954	3.954	0
99. I sometimes look up words on the Internet or in reference books.	0	0	5.4	2.7	60.8	16.2	14.9	3		0.952	3.952	0
190. I do extra work for learning English.	0	0	3.9	9.7	52.4	21.4	12.6	3		0.946	3.946	0
224. In learning it is important to work independently.	0	0	1.0	7.0	36.0	32.0	24.0	3		0.946	3.946	0
67. Knowledge is something I construct for myself.	0	0	2.6	2.6	57.9	15.8	21.1	3		0.945	3.945	0
222. Repetition is important for learning.	22.5	25.5	44.1	6.9	1.0	0.0	0.0	3		0.944	3.944	0
220. I am ready to learn in unfamiliar ways.	0	0	2.7	9.3	49.3	24.0	14.7	3		0.943	3.943	0
77. I rely on myself when learning.	0	0	0	8.0	45.3	22.7	24.0	3		0.941	3.941	0
94. Reading is an active activity.	0	0	1.0	6.9	42.2	28.4	21.6	3		0.933	3.933	0
139. I use techniques to help me remember vocabulary.	0	0	6.9	13.7	55.9	16.7	6.9	3		0.928	3.928	0
50. I have tried different ways of learning.	0	0	9.2	10.5	59.2	15.8	5.3	3		0.923	3.923	0
79. I want to make my own choices about learning.	0	0	0	3.9	38.2	26.3	31.6	3		0.919	3.919	0
249. It is necessary to practice using English outside the classroom.	0	0	0	2.7	26.7	22.7	48.0	3		0.916	3.916	0
216. I am aware of the feelings of others.	0	1.3	2.7	2.7	64.0	17.3	12.0	3		0.912	3.912	0
127. I know my problems in writing.	0	0	5.4	10.8	51.4	27.0	5.4	3		0.892	3.892	0
39. If I find an English word that I don't know, I try to guess it.	0	0	1.9	20.4	56.3	11.7	9.7	3		0.889	3.889	0
191. I think about what I have studied in class.	0	0	3.9	16.7	39.2	37.3	2.9	3		0.887	3.887	0
228. I motivate myself.	0	0	2.0	14.7	43.1	32.4	7.8	3		0.885	3.885	0
238. I choose the exercises I work on.	0	0	4.0	23.8	46.5	21.8	4.0	3		0.883	3.883	0
4. In general, learning continues all of a person's life.	0	0	0	1.3	23.4	15.6	59.7	3		0.883	3.883	0
206. I try to find the best environment for studying.	0	0	0	9.3	44.0	30.7	16.0	3		0.875	3.875	0
152. I know different ways of practising grammar.	0	0	12.5	30.6	50.0	4.2	2.8	3		0.871	3.871	0
146. Different types of text (magazine, letter, recipe etc.) have different vocabulary.	0	0	0	2.7	38.7	30.7	28.0	3		0.871	3.871	0
137. I try to use newly learned words in my essays.	0	0	2.9	19.6	57.8	11.8	7.8	3		0.867	3.867	0
204. I use my mistakes to help me do better.	0	0	1.0	5.8	63.1	14.6	15.5	3		0.853	3.853	0
22. I know how to study	0	0	4.0	17.3	57.3	16.0	5.3	3		0.846	3.846	0

Item	Likert Responses as Percentage							Selection Criteria				
	1	2	3	4	5	6	7	A	B	C	D	E
169. When I am learning something new I look for similarities with things I already know.	0	0	0	9.7	54.4	22.3	13.6	3		0.844	3.844	0
231. I can study independently.	0	0	1.0	5.9	46.1	33.3	13.7	3		0.841	3.841	0
126. There are different types of writing.	0	0	0	1.3	21.3	32.0	45.3	3		0.827	3.827	0
219. I reflect on my learning.	0	0	1.0	10.8	56.9	21.6	9.8	3		0.825	3.825	0
97. When I read in English, I predict the content of a text (using pictures, headings, the context, etc.).	0	0	1.0	11.8	57.8	20.6	8.8	3		0.814	3.814	0
174. I consider facts and come to objective conclusions.	0	1.4	0	18.9	59.5	17.6	2.7	3		0.776	3.776	0
52. I try to fix problems I have in learning.	0	0	0	5.3	50.7	33.3	10.7	3		0.760	3.760	0
13. When I learn something new I feel satisfaction in myself	0	0	0	0	14.7	34.7	50.7	3		0.729	3.729	0
165. Different people have different ways of learning.	0	0	0	0	13.3	17.3	69.3	3		0.721	3.721	0
85. I have a mature attitude to learning.	0	0	9.7	30.6	44.4	9.7	5.6	2		1.228	3.228	0
155. I can help myself to improve my level of grammar.	0	0	8.1	25.3	53.5	9.1	4.0	2		1.186	3.186	0
98. I read newspapers in a different way to books.	0	1.4	6.8	9.6	54.8	17.8	9.6	2		1.182	3.182	0
53. I know some different ways of learning.	0	0	4.1	5.5	72.6	11.0	6.8	2		1.127	3.127	0
28. I feel unlucky when I get bad marks.	1.3	3.9	9.2	23.7	48.7	9.2	3.9	2		1.123	3.123	0
207. I know the aim of the learning tasks I do.	0	0	1.0	8.1	55.6	27.3	8.1	2		1.076	3.076	0
209. I think about my progress in learning English.	0	0	2.0	7.1	65.7	16.2	9.1	2		1.074	3.074	0
118. I know different ways of practising speaking English.	0	1.4	9.6	28.8	46.6	6.8	6.8	2		1.026	3.026	0
69. In the last 4 months, I have thought that something a teacher told the class was wrong.	0	0	8.0	22.7	53.3	8.0	8.0	2		0.968	2.968	0
102. My general knowledge helps me to understand texts I read.	0	0	2.0	4.0	53.5	30.7	9.9	2		0.964	2.964	0
60. I always agree with what a teacher says.	0	0	9.2	30.3	48.7	9.2	2.6	2		0.873	2.873	0
132. I guess the meaning of new words.	0	0	1.0	7.8	68.9	18.4	3.9	2		0.658	2.658	0

A = Responses at 10+. B =Polarity (the smallest difference). C = Standard Deviation. D = Sum Score. E = No. of shaded.


 Indicates a top-50 item in that category (there are 75 questions with at least one shaded)

Table 10.4: Selection table

10.5 Movement of teacher estimates

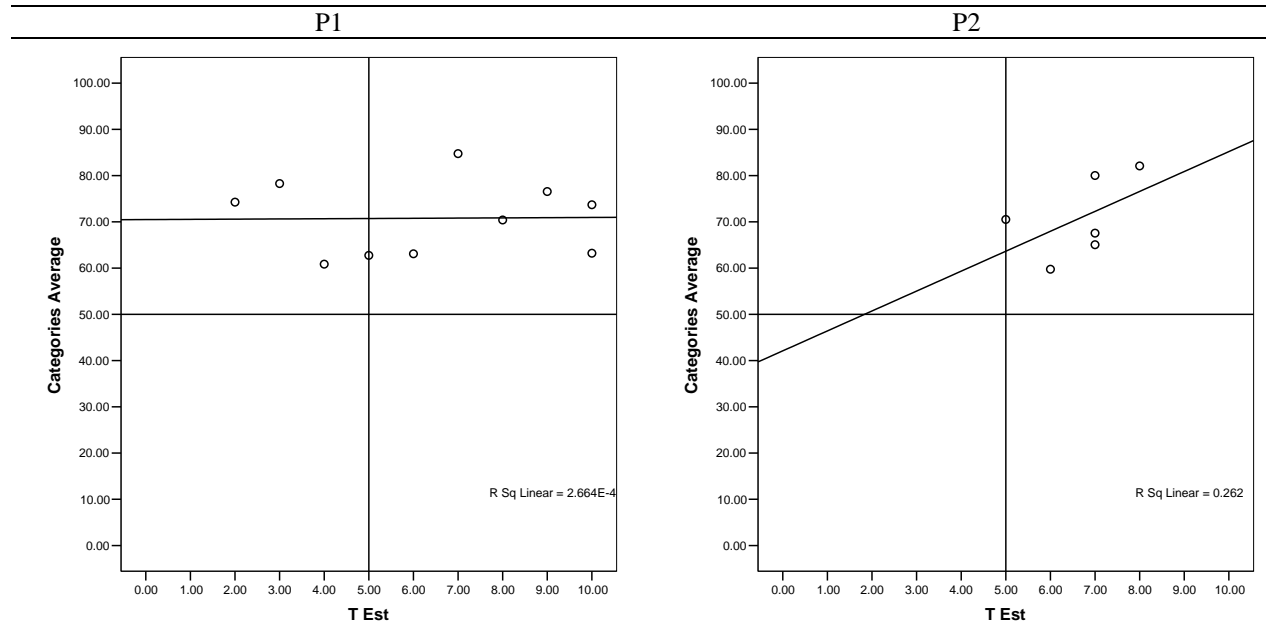


Figure 10.1: Categories Average scores plotted against Teacher Estimates

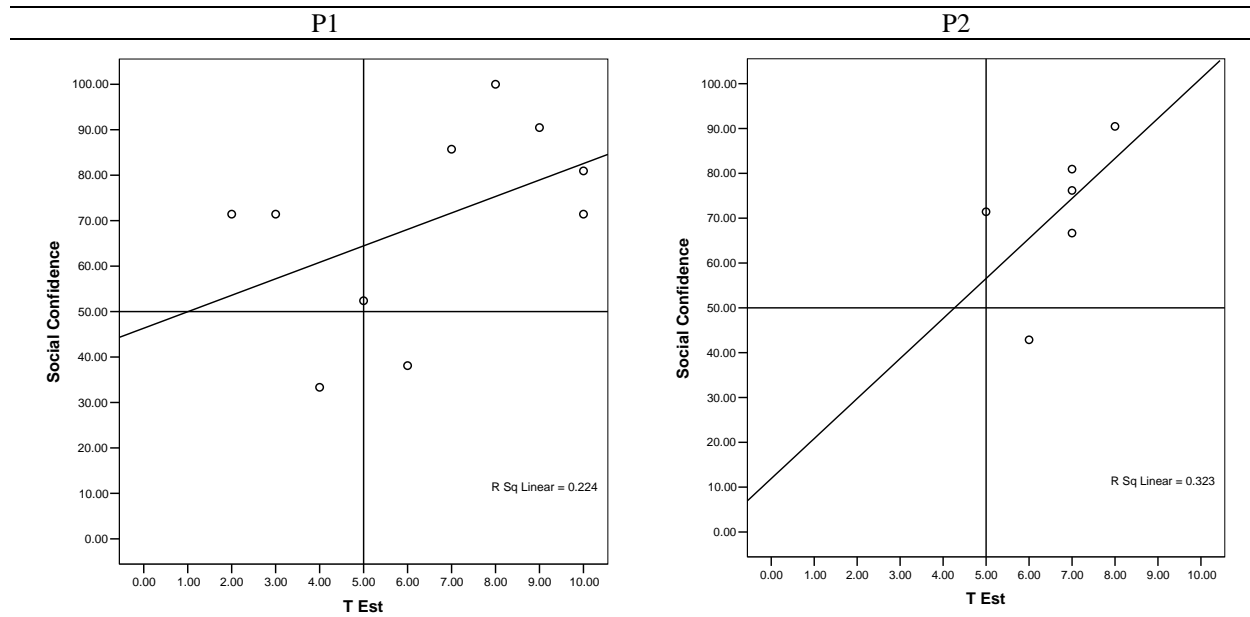


Figure 10.2: Social Comparison scores plotted against Teacher Estimates

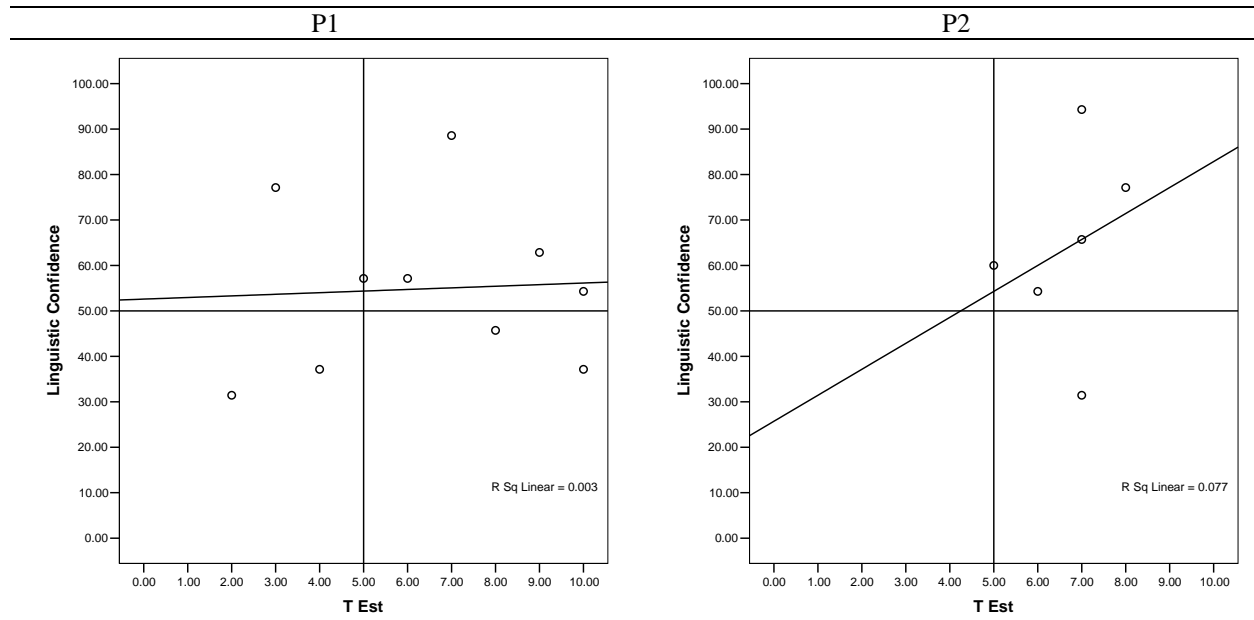


Figure 10.3: Linguistic Confidence scores plotted against Teacher Estimates

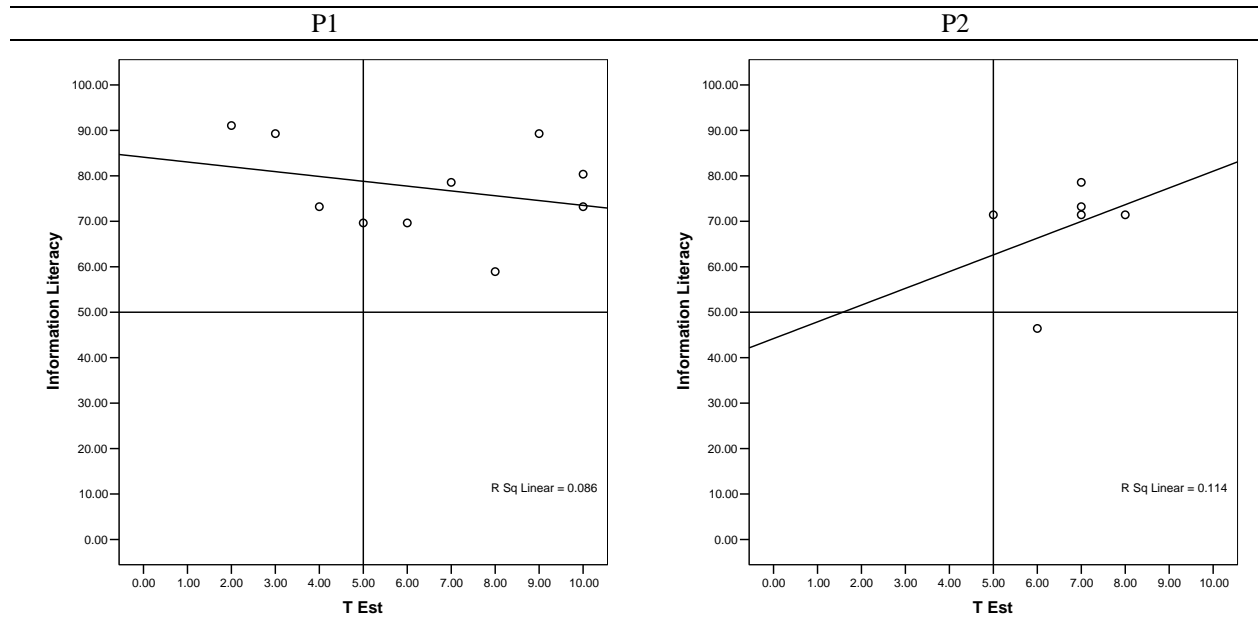


Figure 10.4: Information Literacy scores plotted against Teacher Estimates

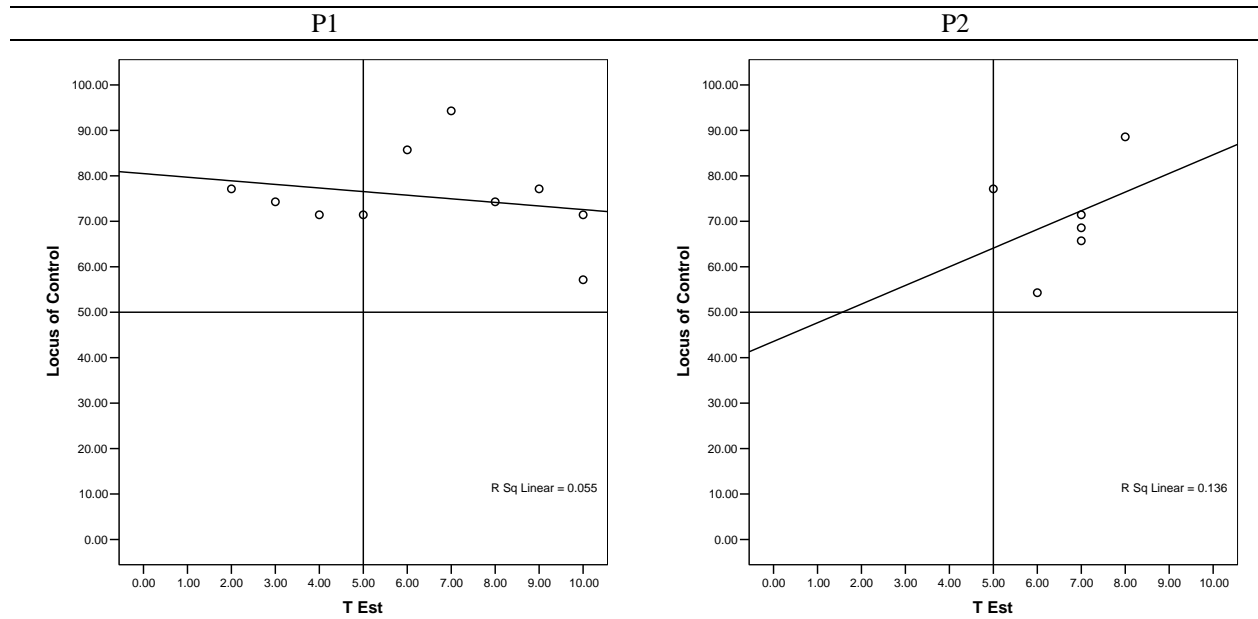


Figure 10.5: Locus of Control scores plotted against Teacher Estimates

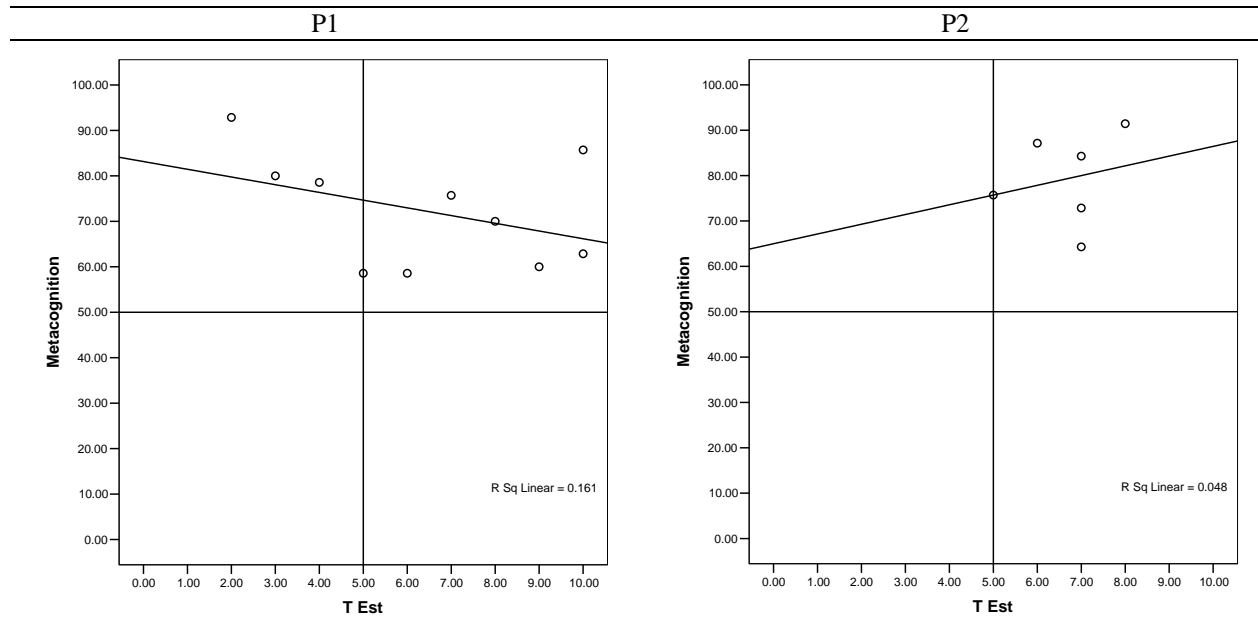


Figure 10.6: Metacognition scores plotted against Teacher Estimates

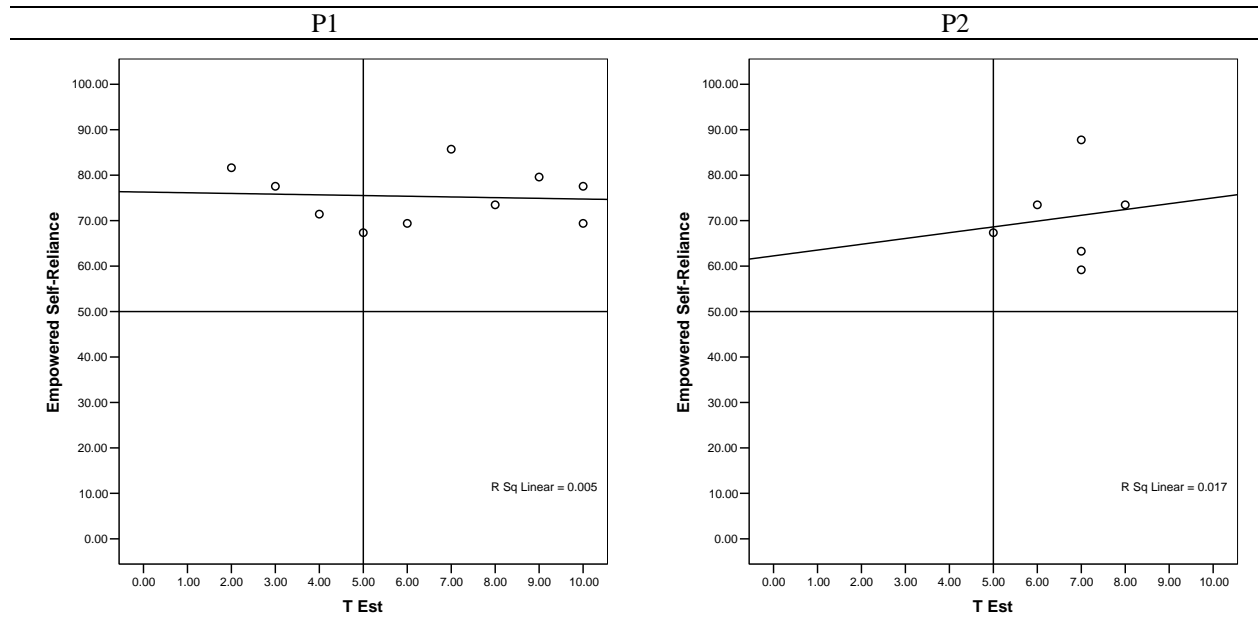


Figure 10.7: Self-Reliance scores plotted against Teacher Estimates