



Durham E-Theses

The Power of the Situation: Variability and Stability in Chinese University Students' Willingness to Communicate in English Classrooms

ZHANG, JIAYI

How to cite:

ZHANG, JIAYI (2018) *The Power of the Situation: Variability and Stability in Chinese University Students' Willingness to Communicate in English Classrooms*, Durham theses, Durham University. Available at Durham E-Theses Online: <http://etheses.dur.ac.uk/12781/>

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in Durham E-Theses
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full Durham E-Theses policy](#) for further details.

Academic Support Office, Durham University, University Office, Old Elvet, Durham DH1 3HP
e-mail: e-theses.admin@dur.ac.uk Tel: +44 0191 334 6107
<http://etheses.dur.ac.uk>

The Power of the Situation: Variability and Stability in Chinese University Students' Willingness to Communicate in English Classrooms

Jiayi Zhang

Abstract

Willingness to communicate (WTC) used to be studied as a relatively stable, trait-like predisposition; however, recent attention has been shifted to more dynamic, state-like components of WTC, i.e. possible fluctuations in state WTC over time. This research investigates variability and stability in both trait and state WTC. It particularly focuses on within-person variability in state WTC, which may lead to stable between-person differences, and situational antecedents that can either promote or hinder state WTC in L2 classrooms. Published research on state WTC was systematically reviewed, and frequently reported situational antecedents of state WTC were organised into a framework composed of three inter-linked layers: situation cues, situation characteristics, and underlying dimensions. Two high-density repeated measurement studies conducted in English classrooms in a Chinese university, investigating whether, how and why state WTC varied during a semester (Study 1) and during a lesson (Study 2), respectively. The two studies found that (a) state WTC varied both within a lesson and across different lessons during a semester; (b) within-person variability in state WTC was systematically related to the situational antecedents proposed in the framework (e.g. support, task-importance, task-interest, etc.); and (c) systematic within-person variability in state WTC predicted English learning performance, particularly in terms of communicative competence. This research provides novel insights into how within-person variability in state-like variables can be studied, and the proposed framework can be used as guidance for future research on state WTC and its situational antecedents. This research also suggests practical strategies for educators and language L2 teachers who would like to facilitate student state WTC by systematically shaping classroom learning situations.

**The Power of the Situation:
Variability and Stability in Chinese University Students'
Willingness to Communicate in English Classrooms**

By

Jiayi Zhang

A thesis submitted to
School of Education
in partial fulfilment of the requirements
for the degree of
Doctor of Philosophy

Durham University

2018

Table of Contents

Table of Contents	1
List of Figures	6
List of Tables	7
Declaration	8
Statement of Copyright	8
Acknowledgment	9
Introduction	10
1 WTC in Second Language Acquisition	13
1.1 Definition of WTC	13
1.2 WTC and Communication Behaviour	16
1.3 An Overview of Research on Trait WTC	17
1.3.1 Individual differences in trait WTC	17
1.3.2 Personal antecedents of trait WTC	18
1.3.2.1 L2 self-confidence	18
1.3.2.2 Motivation	19
1.3.2.3 Personality	21
1.3.2.4 Social context	22
1.3.2.5 Age and gender	23
1.4 Summary	24
2 Dynamic Approaches to Personality and Their Application to SLA	25
2.1 Introduction	25
2.2 Dispositional and Dynamic Approaches	25
2.3 Density Distribution Approach	27
2.4 Person-Situation Interaction	30
2.4.1 If ... then patterns	30
2.4.2 Situation-based contingencies	32
2.4.3 Psychological characteristics of situations	32
2.5 Application of Dynamic Approaches to SLA	34
2.6 Summary	35
2.7 Research Aim and Questions	35
3 A Systematic Review of Situational Antecedents of State WTC	37
3.1 Introduction	37
3.2 Inclusion Criteria	37
3.3 Methodological Considerations	38
3.4 Situation Cues	41

3.4.1	Interlocutors	41
3.4.1.1	Familiarity, participation and cooperation	41
3.4.1.2	Demographic features	42
3.4.2	Classroom atmosphere	42
3.4.2.1	Classmates.....	43
3.4.2.2	Class size.....	43
3.4.2.3	Teachers	43
3.4.3	Tasks	44
3.4.3.1	Topic	44
3.4.3.2	Type of activity	44
3.4.3.3	Preparation time	45
3.4.3.4	Assessment.....	45
3.5	Situation Characteristics.....	46
3.5.1	Task-confidence.....	46
3.5.2	Task-interest	47
3.5.3	Task-usefulness.....	47
3.6	Systematicity of Previous Studies.....	47
3.7	Proposed Framework of Situational Antecedents of State WTC.....	49
3.7.1	Major dimensions of situation characteristics.....	51
3.7.2	Specific situation characteristics and cues.....	51
3.8	Discussion	52
3.9	Summary	54
4	Philosophical Assumptions.....	55
4.1	Ontological and Epistemological Positions.....	55
4.2	Question-Led Research	56
4.3	Types of Research Questions	57
4.4	Summary	58
5	Information about the Research Context	59
5.1	Communicative Language Teaching in China	59
5.2	English Major Programmes and College English Courses	60
5.3	Chinese Students' WTC in English.....	62
5.4	Summary	65
6	Study One: Variability and Stability of L2 WTC during a Semester	66
6.1	Introduction.....	66
6.2	Methodology	67
6.2.1	Overview.....	67
6.2.2	Context.....	67
6.2.3	Participants	68

6.2.4	Ethical considerations	69
6.2.5	Research design	70
6.2.6	Procedure	70
6.2.7	Materials	72
6.2.7.1	Baseline measures	73
6.2.7.2	Momentary measures	76
6.2.7.3	Language exam performance	77
6.2.8	Data analysis	78
6.2.8.1	Analysis for research question 1	78
6.2.8.2	Analysis for research question 2	79
6.2.8.3	Analysis for research question 3	81
6.3	Results	83
6.3.1	Descriptive statistics	83
6.3.1.1	Trait WTC in English and in Mandarin	83
6.3.1.2	State WTC and self-reported communication behaviour	86
6.3.2	Variability in trait and state WTC	87
6.3.2.1	Within-person variability in state WTC	87
6.3.2.2	Between-person variability in trait and state WTC	89
6.3.2.3	Within- vs between-person variability comparison	90
6.3.2.4	Variability in self-reported communication behaviour	90
6.3.3	Systematicity of variations in trait and state WTC	91
6.3.3.1	WTC and personality	91
6.3.3.2	WTC and situation perceptions	92
6.3.4	WTC as a predictor of language learning performance	98
6.3.4.1	Trait WTC and exam performance	99
6.3.4.2	State WTC and self-assessed performance	100
6.3.4.3	Comparison between self-assessed and exam performance	103
6.3.4.4	Situation-contingent WTC and exam performance	103
6.3.4.5	Variability in state WTC and exam performance	106
6.4	Discussion	107
6.5	Summary	110
7	Study Two: Variability and Stability of L2 WTC during a Lesson	112
7.1	Introduction	112
7.2	Methodology	113
7.2.1	Overview	113
7.2.2	Context	113
7.2.3	Participants	115
7.2.4	Ethical considerations	115

7.2.5	Research design	116
7.2.6	Procedure	117
7.2.7	Materials	119
7.2.7.1	Baseline measures	119
7.2.7.2	Momentary measures	120
7.2.7.3	Language learning performance.....	121
7.2.8	Data analysis	122
7.2.8.1	Statistical analysis.....	122
7.2.8.2	Thematic analysis.....	122
7.3	Results and Discussion.....	125
7.3.1	Results and discussion of the statistical analysis	125
7.3.1.1	Descriptive statistics	125
7.3.1.2	WTC and communication behaviour	127
7.3.1.3	Variability in trait and state WTC.....	130
7.3.1.4	WTC as a predictor of language learning performance	132
7.3.1.5	Teacher- and self-assessed language performance	134
7.3.1.6	Discussion of the statistical analysis.....	135
7.3.2	Results and discussion of the thematic analysis	137
7.3.2.1	Positivity	137
7.3.2.2	Negativity.....	144
7.3.2.3	Duty.....	153
7.3.2.4	WTC and communication behaviour	156
7.3.2.5	Teacher and student criteria for language performance.....	159
7.4	Summary and Conclusion	164
8	Discussion.....	167
8.1	Introduction.....	167
8.2	Research Question 1: Variability	167
8.2.1	Between-person variability in trait WTC.....	167
8.2.2	Within-person variability in state WTC.....	169
8.2.3	Individual differences in within-person variability.....	170
8.3	Research Question 2: Systematicity.....	171
8.3.1	Personal antecedents of trait and state WTC	172
8.3.2	Situational antecedents of trait and state WTC.....	173
8.3.2.1	Teacher support.....	175
8.3.2.2	Classroom atmosphere	177
8.3.2.3	Peer cooperation.....	178
8.3.2.4	Task-interest.....	179
8.3.2.5	Task-usefulness.....	181

8.3.2.6 Task-confidence	182
8.3.2.7 Task-difficulty	183
8.3.3 Individual differences in within-person relationships.....	185
8.4 Research Question 3: Predictability	185
8.4.1 Trait WTC and language exam performance.....	186
8.4.2 State WTC and self-assessed language performance.....	187
8.4.3 Situation-contingent WTC and language exam performance	188
8.5 Unexpected Outcomes	188
8.5.1 Distinction between WTC and communication behaviour	188
8.5.1.1 WTC vs communication behaviour comparison.....	188
8.5.1.2 Chinese students' 'silence' revisited.....	191
8.5.2 Inconsistency between teacher and student assessments	192
8.6 Limitation and Direction for Future Research	194
8.7 Conclusion and Implication	197
8.7.1 Theoretical implications	197
8.7.2 Practical implications.....	199
References.....	201
Appendix A An Overview of Relevant Studies on State WTC	220
Appendix B Consent Form for Study 1	225
Appendix C Baseline Self-Report Questionnaire for Study 1	227
Appendix D Momentary Self-Report Questionnaire for Study 1	236
Appendix E Consent Form for Study 2.....	238
Appendix F Baseline Questionnaire for Teachers in Study 2	240
Appendix G Momentary Self-Report Questionnaire for Study 2	241
Appendix H Momentary Questionnaire for Teachers in Study 2.....	242
Appendix I Interview Schedule.....	243
Appendix J Observation Schedule	244

List of Figures

Figure 1.1. MacIntyre et al.'s (1998) heuristic model of variables influencing WTC (p. 547) ..	15
Figure 2.1. Fleeson's (2001) two extreme distributions of states over time (p. 1012)	27
Figure 2.2. Schmitz's (2006) comparison between aggregated and individual trajectories (p. 439) ...	29
Figure 3.1. The proposed framework of situational antecedents of state WTC	50
Figure 6.1. Distributions of non-English major students' trait WTC	84
Figure 6.2. Distributions of non-English major students' state WTC (aggregate level)	86
Figure 6.3. Aggregated and individual trajectories of state WTC over one semester	87
Figure 6.4. Variability in non-English major students' trait and state WTC	88
Figure 6.5. Three density distributions of state WTC	89
Figure 6.6. Variability in non-English major students' communication behaviour	91
Figure 6.7. Aggregated and individual interest-contingent WTC slopes	98
Figure 6.8. Variability in non-English major students' self-assessed performance	101
Figure 6.9. Distribution of individual WTC-performance slopes	102
Figure 6.10. State WTC trajectories for three individuals	107
Figure 7.1. Distributions of English major students' trait WTC	126
Figure 7.2. Distribution of English major students' state WTC (aggregate level)	127
Figure 7.3. Distribution of teacher-reported communication behaviour (trait)	128
Figure 7.4. Distributions of teacher- and self-reported communication behaviour (aggregate level) ...	130
Figure 7.5. Aggregated and individual trajectories of state WTC within one lesson	131
Figure 7.6. Variability in English major students' trait and state WTC	132
Figure 7.7. Distributions of English major students' end-of-term English scores	133
Figure 7.8. Distributions of teacher- and self-assessed baseline English scores	134
Figure 7.9. Distributions of teacher- and self-assessed performance in one lesson	135

List of Tables

Table 6.1. Descriptive statistics for trait variables in Study 1 ($N = 61-93$).....	85
Table 6.2. Between- and within-person variability in personality states estimated by HLM	92
Table 6.3. Between-person correlations between state WTC and personality states ($N = 102$)..	92
Table 6.4. Between- and within-person correlations between state WTC and situation perceptions...	94
Table 6.5. Descriptive statistics for within-person situation-WTC correlations	96
Table 6.6. Descriptive statistics for situation-contingent WTC slopes estimated by HLM	97
Table 6.7. Regression of performance on conscientiousness and trait WTC	100
Table 6.8. Regression of performance on trait and state WTC	100
Table 6.9. Regression of end-of-term performance on baseline and self-assessed performance....	103
Table 6.10. Correlations between performance and situation-WTC correlations ($N = 62-93$)....	104
Table 6.11. Regression of performance on trait WTC and interest-WTC correlations	105
Table 6.12. Correlations between performance and situation-contigent WTC ($N = 63-69$)	105
Table 6.13. Regression of performance on trait WTC and interest-contingent WTC.....	106
Table 7.1. Descriptive statistics for trait variables in Study 2 ($N = 30-31$).....	129
Table 7.2. Descriptive statistics for state variables in Study 2 ($N = 31$)	130

Declaration

I declare that this thesis is my own work. No material contained in this thesis has previously been submitted for a degree in this or any other institution.

Statement of Copyright

The copyright of this thesis rests with the author. No quotation from it should be published without the author's prior written consent and information derived from it should be acknowledged.

Acknowledgment

I would like to take this opportunity to express my gratitude to a number of people who contributed to the successful completion of my doctoral study.

Above all, I would like to express my deep, heartfelt gratitude to my supervisors, Dr Nadin Beckmann and Prof Jens F. Beckmann, for their patient, encouraging, and considerate guidance during the past three years. They not only shaped my work, but also provided me with care and support, which made this journey very enjoyable. It was their wonderful supervision that made me complete my doctoral study so quickly and successfully.

I am also very thankful to Dr Philip Nathan for his comments on my systematic review of the literature.

I would like to extend my thanks to all the participants in my research, including the two main studies as well as the pilot studies, for their full cooperation. The three English teachers participated in this research deserve special thanks because of their patience and support, though they have to remain anonymous here.

I also want to express my gratitude to the head of the English Department at the university where the research was carried out for granting me access to the classrooms.

My appreciation also goes to all the English teachers who facilitated my fieldwork. I particularly wish to thank Mrs Yingli Zhou for helping me to contact these English teachers.

Last but not least, words cannot express how much I feel indebted to my parents. Without their love and support, I would not have been able to do my doctoral study in Durham. I love you both!

Introduction

Since the emergence of communicative language teaching, more and more researchers and language teachers have been aware of the importance of communication in second language acquisition (SLA), believing that actively engaging in communication can contribute to language learning (Mitchell, 1988). To predict individuals' communication frequency, the concept of willingness to communicate (WTC) was introduced. WTC was originally seen as a stable personality trait that features individual differences, in that learners who are more willing to communicate tend to seize more opportunities to use the target language, and thus are likely to learn the language better than those who are less willing (e.g. McCroskey & Richmond, 1991). Recently, some researchers (e.g. Mystkowska-Wiertelak & Pawlak, 2017) have shifted their interests to the dynamic nature of WTC, exploring possible variability in WTC and situational variables related to this. As the goal of WTC research is to improve L2 learners' communication intention and language learning performance, the dynamic nature of WTC deserves more investigation, e.g. how it varies over time, and whether it is malleable. This shift runs parallel to a rekindled interest into dynamic, state related aspects of personality research (e.g. Beckmann & Wood, 2017; Fleeson & Jayawickreme, 2015; Minbashian, Beckmann & Wood, 2017; Rauthmann, Sherman & Funder, 2015; Rauthmann & Sherman, 2016a).

Illuminated by personality literature, this research adopts a dynamic perspective to study variability and stability in Chinese university students' WTC in English classrooms. It is concerned with both trait and state WTC, although its focus is on the state level. It not only aims at identifying stable individual differences (i.e. between-person variability) in trait WTC, but also at exploring within-person variability in state WTC and individual differences in this. It also aims to examine whether variability in WTC, particularly within-person variability in state WTC, is systematic and predicts language learning performance. To achieve these goals, a systematic review of research on state WTC and its situational antecedents was conducted, followed by two empirical studies investigating variability in trait and state WTC, and any underlying stable patterns.

Firstly, literature on within-person variability in state WTC and its situational antecedents reported by research in different L2 contexts was systematically reviewed. After searching major databases (Web of Science, ERIC and the British Education Index), 35 empirical studies' findings related to state WTC and its situational antecedents were

reviewed. They identified within-person variability in state WTC in different L2 learning contexts, and suggested different situational antecedents related to within-person variability. Frequently reported situational antecedents were categorised into situation cues, situation characteristics and underlying dimensions, and then organised into a multi-layered framework. This framework served as a basis for the variable selection of the empirical studies to investigate the systematicity of within-person variability in state WTC. This systematic review has been published in *System* (an international journal of educational technology and applied linguistics).

Two empirical studies were conducted in a Chinese university. Both employed high-density repeated measurement designs. Study 1 focused on within-person variability in state WTC across different lessons during a semester, and particularly aimed at examining the relationships between state WTC and its situational antecedents as summarised in the proposed framework in the literature review. The sample consisted of two classes of non-English major undergraduate students ($N = 103$). As well as reporting their trait WTC and other related traits on a set of baseline questionnaires at the beginning of the semester, these students completed a momentary questionnaire in each English lesson throughout the semester, indicating their state WTC and situation perceptions at specific points in time. Results showed that students' state WTC varied across different lessons during a semester, and this was systematically related to changes in situation characteristics and underlying dimensions as summarised in the framework. In addition, evidence was found indicating that students' language learning performance was predicted by their trait and state WTC.

Study 2 was a short-term study focused on within-person variability in state WTC across different communication activities during a lesson. It particularly aimed at further exploring the psychological process of communication generation. The sample consisted of two classes of English major undergraduate students ($N = 31$). As well as responding to baseline questionnaires, these students rated their state WTC, together with actual communication behaviour, in each communication activity during a specific lesson, and were interviewed in groups immediately after class to reflect on how their communication was generated or hindered. It was found that within-person variability in state WTC during a lesson was comparable in its amount to observed within-person variability during a semester. Results also indicated that trait WTC predicted language learning performance, particularly when focusing on communicative competence; however, compared to state

WTC, actual communication behaviour seemed to be a better predictor of language learning performance. This suggests that subjective communication intention (i.e. WTC) and actual communication behaviour are not the same, and more attention should be paid to actual behaviour to improve communicative competence.

This thesis consists of seven chapters. Chapter 1 will begin with an introduction to the concept of WTC and a review of the literature on trait WTC. Chapter 2 will review selected dynamic approaches as discussed in personality literature, followed by a dynamic perspective on second language acquisition. The research questions, emphasising within-person variability in state WTC and its underlying stable structure, will then be presented. Chapter 3 will be a systematic review of research on state WTC, providing a comprehensive grasp of situational antecedents that may explain within-person variability in state WTC. Chapter 4 will discuss the philosophical assumptions underlying the research. Chapter 5 will introduce the background information about the research context. The next two chapters will present the two empirical studies, respectively. Chapter 6 will be on Study 1, which will shed light on within-person variability in state WTC during a semester and statistical relationships between state WTC and its situational antecedents. Chapter 7 will be on Study 2, illustrating within-person variability in state WTC during a lesson and psychological processes of communication generation in different situations. In Chapter 8, main findings of the two studies will be summarised and discussed in relation to findings in the literature and theoretical and practical implications of this research will be proposed, and directions for future research will be suggested.

1 WTC in Second Language Acquisition

Communication plays an important role in education, as learning often occurs through effective communication (Powell & Powell, 2010). By interacting with teachers, peers, learning materials and other elements, learners can construct their own thoughts and perceptions of objects or issues, thus triggering further intellectual development (Arends, Winitzky & Tannenbaum, 2001). Even when engaging in non-linguistic tasks, language may improve learning performance (Vygotsky, 1978). Some researchers (e.g. Sprague, 1992) claim that classroom communication can facilitate learning in all subjects. Compared to students who eagerly communicate in class, students who stay silent are often seen as passive learners who are less likely to learn well (MacKinnon & Manathunga, 2003). As a result, students' active communication is always expected and encouraged in classroom settings.

In the field of second language acquisition, communication has received much attention. According to the Council of Europe, the aim of language learning is to overcome barriers to communication among people of different languages and cultural backgrounds, to enable them to better understand and closely co-operate with each other (Council of Europe, 2001). Similarly, the Japanese Ministry of Education, Culture, Sports, Science, and Technology (MEXT) defines foreign language proficiency as the capability of using a foreign language to communicate efficiently with people of different countries and cultures. To develop this communicative competence, student-centred communication activities such as giving speeches, presentations, and participating in debates and discussions are required (MEXT, 2011). Hence, modern language pedagogy emphasises communicative competence and spends much time on communication tasks (MacIntyre, Baker, Clément & Donovan, 2003). Researchers and language teachers believe that language learning gradually occurs during the process of interacting with others in communication tasks (Kang, 2005).

1.1 Definition of WTC

As an individual difference construct, WTC was originally introduced into L1 communication literature as a stable and trait-like predisposition explaining observed variability in communication behaviour (McCroskey & Richmond, 1990). Although McCroskey and Richmond (1990) noticed variability in WTC within individuals across

different contexts, they paid more attention to differences between individuals, i.e. some individuals tend to be talkative, while others seldom engage in interpersonal communication. They believed that one's L1 WTC is largely rooted in personality, and mainly influenced by personal variables such as extroversion, self-esteem, perceived communicative competence, communicative apprehension, and cultural background.

MacIntyre (1994) introduced the concept of WTC to SLA to predict L2 learners' use of target language. Instead of directly adopting the definition of WTC from L1 communication literature, L2 WTC is defined as "readiness to enter into discourse at a particular time with a specific person or persons, using a L2" (MacIntyre, Dörnyei, Clément, and Noels, 1998, p. 547). Compared to the original trait-like characteristic of WTC in L1 communication literature, the definition of WTC in the L2 learning context indicates the dynamic nature of L2 WTC and emphasises its potential variability across time and situations. Additionally, it is suggested that both oral communication (e.g. comprehension and speaking) and written communication (e.g. reading and writing) should be taken into consideration when studying L2 WTC (MacIntyre et al., 1998). This thesis is concerned with both oral and written communication, with a particular focus on speaking because it is more 'observable' than other forms of communication and is often assumed to be most effective in promoting language learning (MacIntyre et al., 2003). Hence, in this thesis, WTC is defined as L2 learners' intention to communicate with peers and teachers in L2 classrooms, particularly by ways of speaking.

MacIntyre et al. (1998) summarised the influential factors of L2 WTC and communication behaviour in a heuristic model, widely recognised as the dominant model of L2 WTC. The heuristic model is made up of six layers, each representing one or more variables relevant to WTC (see Figure 1.1). Depicted as a pyramid-shaped structure with L2 use at the top, the model not only presents multiple variables influencing L2 learners' WTC and communication behaviour, but also indicates "the immediacy of some factors and the relatively distal influence of others" (MacIntyre et al., 1998, p. 546). In this model, both enduring (e.g. personality traits) and situational variables (e.g. state communicative self-confidence) are presented as influential factors of L2 learners' WTC and communication behaviour. For example, a distinction has been made between state self-confidence (i.e. the momentary feeling of confidence to communicate in the target language in a specific situation) and trait self-confidence (i.e. one's general belief in his or her own competence to communicate appropriately in the target language), and it is

suggested that both variables affect L2 learners' WTC and communication behaviour.

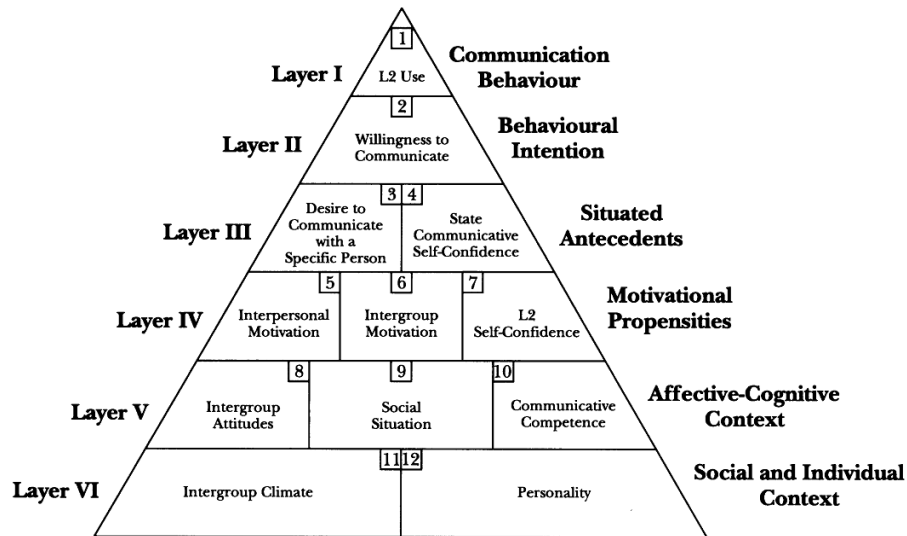


Figure 1.1. MacIntyre et al.'s (1998) heuristic model of variables influencing WTC (p. 547)

The variables listed in the top three layers are susceptible to situational differences. The apex of the pyramid is L2 use, the objective communication behaviour in the target language in a specific situation, and WTC (i.e. subjective L2 communication intention) is regarded as the only variable that directly predicts L2 use or communication behaviour. MacIntyre et al. (1998) situated WTC in specific situations, suggesting that it is directly influenced by situational antecedents. The top three layers indicate the dynamic nature of WTC, in accordance with MacIntyre et al.'s (1998) definition of WTC in L2 learning contexts. By comparison, the lower three layers are composed of enduring predictors of WTC. These are relatively stable and long-term properties within a person that apply to different temporal situations. These enduring variables indirectly influence L2 learners' WTC and L2 use through the mediation of the situational antecedents in the upper layers. The lower three layers indicate that one's WTC does not randomly vary across different situations, but systematically deviates from typical L2 communication tendency, and this can be explained by personal variables, such as L2 self-confidence and personality.

MacIntyre et al.'s (1998) definition and heuristic model – both indicate that L2 WTC is not completely fixed in personality, but changes across different communication situations – have been widely referenced by research in the field of SLA (e.g. Baker & MacIntyre, 2000; Clément, Baker, & MacIntyre, 2003; Ghonsooly, Khajavy, & Asadpour, 2012; Öz, Demirezen, & Pourfeiz, 2015; Peng, 2014; Peng & Woodrow, 2010; Yashima,

2002). Peng and Woodrow (2010) stated that WTC “is conceptualised to display dual characteristics at both trait and state levels” (p. 835). At the trait level, a general communication tendency is assumed that is rooted in an individual’s personality, whilst at the state level, an individual’s intention to communicate fluctuate across time and situations. The layout of distal and proximal variables in the heuristic model shows that, while the trait-like variables work as the foundation, the situational variables play more direct and decisive roles in predicting WTC and L2 use in specific situations.

1.2 WTC and Communication Behaviour

MacIntyre et al.’s (1998) pyramid model shows a distinction between WTC and actual communication behaviour, as WTC is located below L2 use and seen as the final psychological step before L2 use. However, some authors have confused WTC, particularly its state characteristics, with communication behaviour. For example, Cao’s (2013) longitudinal case study on the dynamic nature of WTC measured state WTC through classroom observation, such as counting participant turn-taking. However, WTC refers to people’s subjective *intention* to communicate in certain situations. What Cao (2013) observed was frequency of actual communication *behaviour* rather than subjective intention. Subjective intention is internal and cannot be easily observed. Hence, if the key to language learning is WTC rather than actual behaviour, then data would be better collected through self-reporting. However, it might be the case that what really matters is using the language to communicate rather than subjective intention, as not much evidence has been provided showing that WTC leads to language learning. Hence, it is necessary to distinguish WTC from communication behaviour, and provide evidence that WTC is indeed a meaningful construct that is associated with language learning performance and thus deserves investigation. Language learning performance here is operationalised by language exam scores, which not only include objective paper-based exam scores, but also oral exam scores or course teachers’ subjective judgments of students’ language performance in their classrooms (see section 6.2.2 for details).

As this research concerns both trait and state WTC, in the following section a brief overview of the literature on trait WTC and its personal antecedents will be provided. This will, to some extent, explain why WTC is seen as a meaningful individual difference variable in SLA.

1.3 An Overview of Research on Trait WTC

Early research on L2 WTC tended to focus primarily on its trait characteristics. These studies typically measured participants' trait WTC and other relevant predispositions through cross-sectional questionnaires with relatively large samples (e.g. MacIntyre & Doucette, 2010; Peng, 2014). One of the most frequently used questionnaires was a 20-item scale originally developed by McCroskey and Richmond (1990) to measure WTC in the L1 context. After completing this questionnaire, an overall score can be obtained indicating willingness to communicate in general, as well as a set of sub-scores indicating willingness to enter into discourse in different contexts (i.e. in public, meetings, groups, and dyads) and with different receivers (i.e. strangers, acquaintances, and friends). Although respondents indicate their WTC in different situations, the questionnaire does not measure thoughts and feelings at any particular point in time but captures learners' general predispositions towards communication (Mystkowska-Wiertelak & Pawlak, 2017). As explained by McCroskey and Richmond (1990), the questionnaire aims at showing that WTC in different contexts and with different receivers is strongly correlated, confirming their claim that WTC is a trait-like predisposition rooted in personality that remains stable across different situations. Studies using this type of questionnaire (e.g. Cetinkaya 2005; MacIntyre & Charos, 1996; Yashima, 2002) mainly aim at estimating the linear relationships between trait WTC and selected personal variables seen as possible predictors of trait WTC. It seems that, since the very start, researchers have been aware of possible fluctuations in state WTC across different situations and with different people, although early research tended to emphasise the stability of trait WTC.

1.3.1 Individual differences in trait WTC

Amongst students in the same context, some individuals may be more willing to communicate than others. As most previous studies on WTC regard it as a trait-like disposition, differences in trait WTC between individuals have been widely studied. Some researchers sorted their participants into groups of high, medium or low L2 WTC based on scores on trait WTC questionnaires (e.g. Oz, 2014). Oz (2014) conducted a survey with 168 EFL students in a Turkish university and reported that 20% of them described themselves as high in trait WTC, 14% as low, and the rest as moderate. Although the percentages changed slightly when comparing different receivers (i.e. strangers, acquaintance, and friends), students always differed in levels of trait WTC. A few studies

of this type have tested whether individual differences in trait WTC can predict language learning performance, although results were inconsistent. Some studies (e.g. Mahmoodi & Moazam, 2014; Naderifar & Esfandiari, 2016; Oz, 2014; Piechurska-Kuciel, 2018) reported that trait WTC was significantly associated with L2 learning performance, while others (e.g. Joe, Hiver & Al-Hoorie, 2017; Yashima, 2002) reported opposite findings.

In addition to these large-scale questionnaire studies, there are some very small scale observational studies that provide more comprehensive and in-depth illustrations of individual differences in trait WTC (e.g. Liu, 2002). In Liu's (2002) multiple case study of three Chinese graduate students in a US Midwestern university, a different label was used to capture each individual: a conditional participator (a student whose participation in classroom activities fluctuates remarkably with changes in situations, personality states and affect), a marginal interactor (a student who always listens attentively but rarely speaks up), and a silent observer (a student who completely withdraws from oral communication in the classroom). The different labels proposed by Liu (2002) show individual differences in the three students' trait WTC (e.g. the marginal interactor seemed to be more willing to communicate than the silent observer). Additionally, the labels indicate a possibility of within-person variability in state WTC (e.g. the conditional participator's state WTC varied across different situations), and individual differences in this variability (e.g. the conditional participator's state WTC fluctuated more than the silent observer's state WTC).

1.3.2 Personal antecedents of trait WTC

Research on trait WTC mainly aims at estimating relationships between trait WTC and different personal variables (e.g. Baker & MacIntyre, 2000; Clément et al., 2003; Yashima, 2002). Researchers are interested in questions such as: *What makes some learners actively initiate discourse, while others avoid entering the discourse?* The widely studied individual difference antecedents suggested by research on trait WTC include L2 self-confidence, personality, social context, motivation, age and gender. In the following section, I will briefly review previous findings concerning each of these antecedents of trait WTC.

1.3.2.1 L2 self-confidence

As proposed by MacIntyre et al.'s (1998) heuristic model, *L2 self-confidence* plays an

important role in predicting L2 learners' WTC and communication behaviour. Many studies (e.g. Fallah, 2013; Ghonsooly, Khajavy, & Asadpour, 2012; MacIntyre & Charos, 1996; Munezane, 2013; Peng & Woodrow, 2010; Yashima, 2002; Yashima, Zenuk-Nishide & Shimizu, 2004) have reported that L2 self-confidence is the most important predictor of trait WTC in L2 learning contexts. For example, by analysing data collected from a group of 579 students from eight universities in Eastern China through self-report questionnaires, Peng and Woodrow (2010) reported that L2 self-confidence was the strongest predictor of trait WTC.

Confident L2 learners refer to those who have sufficient levels of L2 competence and, are not extremely anxious when using the target language. According to MacIntyre et al. (1998), although actual competence (labelled as *communicative competence*) has an impact on L2 self-confidence and trait WTC, *perceived competence* would be a stronger and more direct predictor. The anxiety one experiences when using a L2 is called *language anxiety* (also labelled as *communication apprehension*, *foreign language anxiety*, and *communication anxiety*; see Mystkowska-Wiertelak & Pawlak, 2017 for details). It has been found that perceived competence and language anxiety correlate significantly, predicting L2 self-confidence (e.g. Fallah, 2013; Ghonsooly et al., 2012; Ortega, 2009).

Evidence has been provided showing that both perceived competence and language anxiety are strong predictors of trait WTC. Data was collected firstly with several groups of French as a second language (FSL) learners in Canada (e.g. Baker & MacIntyre, 2000; Clément et al., 2003; MacIntyre, Baker, Clément & Conrod, 2001; MacIntyre et al., 2003), and then with EFL learners in China (e.g. Liu & Jackson, 2008; Peng, 2014; Peng & Woodrow, 2010), Japan (e.g. Yashima, 2002), Iran (e.g. Ghonsooly et al., 2012), and Turkey (e.g. Cetinkaya, 2005; Öz et al., 2015). It has been suggested that language anxiety tends to be the best predictor of trait WTC in immersion L2 learning contexts, whilst perceived competence tends to be the best predictor in non-immersion contexts (MacIntyre et al., 2003).

1.3.2.2 Motivation

Motivation is frequently studied in trait WTC literature. In the heuristic model, MacIntyre et al. (1998) have suggested intergroup and interpersonal motivation as two variables of WTC. Intergroup motivation refers to attitude towards communicating with people from

another cultural background or using and learning another language, which is related to one's feeling of belonging to a particular group. By contrast, interpersonal motivation refers to attitudes towards communicating with specific people and is influenced by personal characteristics.

Evidence has been provided to show a direct or indirect relationship between motivation and trait WTC (or communication behaviour). For example, MacIntyre and Charos (1996) found a significant correlation between motivation and communication frequency with a group of adult FSL learners in Canada. MacIntyre et al. (2001) examined five L2 learning orientations (i.e. travel, job, friendship, personal knowledge, and school achievement) for FSL learners in Canada, and found that trait WTC positively correlated with all these motivational L2 learning goals. MacIntyre et al. (2001) also found that motivation was more effective in predicting students' trait WTC outside than inside classrooms. Similarly, MacIntyre et al. (2003) reported that significant correlations between motivation and trait WTC appeared in learners with immersion L2 learning experiences, but not in non-immersion L2 learners. These findings indicate that motivation may not be a key predictor of trait WTC in EFL contexts in China as well as other East Asian countries where students seldom communicate in English outside classrooms.

According to Yashima (2002, 2009), EFL learners' L2 learning motivation can be different from that of FSL learners. EFL learners are less likely to have clear intergroup attitude as they use English as a tool to contact foreigners across the world rather than people from a particular language community. Yashima (2002) introduced a new concept of *international posture*. This covers different L2 learning orientations as presented above, including "interest in foreign or international affairs, willingness to go overseas to stay or work, readiness to interact with intercultural partners, and, one hopes, openness or a non-ethnocentric attitude toward different cultures, among others" (Yashima, 2002, p. 57). Significant correlations between international posture and trait WTC have been found in EFL contexts such as Japan (Yashima, 2002, 2009) and China (e.g. Peng, 2015). Hence, international posture is believed to be an important motivational propensity influencing trait WTC in EFL contexts.

Another related personal antecedent of trait WTC is Dörnyei's (2005) *L2 Motivational Self System*, "which is a broad construct of L2 motivation, made up of three dimensions": ideal L2 self, ought-to L2 self, and L2 learning experience (p. 105). *Ideal L2 self* refers

to the person one would like to become in relation to L2 learning, *ought-to L2 self* refers to the characteristics one believes that ought to be achieved, and *L2 learning experience* refers to the L2 learning environment or relevant learning experiences (Dörnyei, 2005). Peng (2015) examined the relationships between L2 Motivational Self System and trait WTC among Chinese university students, and found that all the three dimensions, together with international posture, directly or indirectly predicted trait WTC inside English classrooms.

1.3.2.3 Personality

Another widely studied antecedent of trait WTC is *personality*. MacIntyre et al. (1998) conceptualised personality as a fundamental variable influencing L2 communication and used Big-Five personality traits to explain the relationship between personality and trait WTC. Big-Five personality traits comprise five major dimensions prevalent in the literature of personality science, namely extroversion, agreeableness, conscientiousness, emotional stability, and openness to experience (Goldberg, 1992). Some studies (e.g. MacIntyre & Charos, 1996; Ortega, 2009) have found that each of these traits directly or indirectly predicts L2 WTC. For example, MacIntyre and Charos (1996) found that agreeableness directly correlated with trait WTC, and the others indirectly correlated through language anxiety, attitudes towards the learning situation, integrativeness¹, and perceived communicative competence.

Among the five major personality traits, extroversion receives particular attention. It is believed that extroversion consistently plays a dominant and enduring role in predicting trait WTC. When studying WTC in its original L1 communication context, McCroskey and Richmond (1990) suggested that extroversion and L1 WTC are highly correlated. A number of questionnaire studies with relatively large sample sizes have provided evidence for the correlation between extroversion and L2 WTC. For example, with a group of 92 adult French learners in Ottawa, MacIntyre and Charos (1996) found that extroversion's correlation with trait WTC was the highest among the five personality traits. Similarly, Oz (2014) found that Turkish EFL learners' trait WTC significantly correlated with their extroversion, agreeableness, and openness to experience, with extroversion having the largest effect. Some researchers (e.g. Cetinkaya, 2005; Fu et al., 2012; Sun, 2008) suggest

¹ Integrativeness refers to one's desire to learn a L2 in order to communicate with people from the target language community.

that, compared to introvert learners, extroverts tend to be more confident in their L2 proficiency, making them more willing to communicate in the target language. It is pointed out that although extroversion may not have a strong relationship with written language production, it bears significant correlation with oral language production in both L1 and L2 (Dewaele & Furnham, 1999). By comparison, Chu (2008) reported slightly different results after testing the relationship between trait WTC and shyness (defined as being ‘low in extroversion’) among 364 English learners in a university in Taiwan. Chu (2008) found that, although extroversion correlated with both L1 and L2 WTC, extroversion’s relationship with L2 WTC was much weaker than its relationship with L1 WTC. It indicates that unlike L1 WTC, which mainly reflects one’s extroversion or talkativeness, willingness to communicate in L2 may be influenced by various personal and situational variables.

1.3.2.4 Social context

It has been suggested that L2 learners’ WTC and communication behaviour are also influenced by *social context*. Originally, social context referred to the socioeconomic power represented by the L2 community, as well as learners’ relatively stable attitudes towards the L2 community (see MacIntyre et al., 1998). However, when conceptualising WTC in a Chinese English as a second language (ESL) context, Wen and Clément (2003) argued that MacIntyre and colleagues’ interpretation of social context was from a “North American point of view” and might not fit the English learning context in China (p. 26).

According to Wen and Clément (2003), L2 communication in the Chinese ESL context is confined to classroom settings, and social context thus restricted to classroom atmosphere. Classroom atmosphere is co-constructed by the teacher and students in the classroom and is relatively flexible. Hence, both group cohesiveness and teacher support could possibly promote state WTC. Group cohesiveness refers to the extent to which students unite and wish to remain in the group. It is believed that strong group cohesiveness could help to build a supportive learning environment, thus fostering state WTC. Teachers also exert a crucial impact on classroom learning, especially in China where teachers tend to be seen as the authority. It is believed that a teacher’s attitude, involvement, immediacy and teaching style could also influence students’ state WTC in class (Wen & Clément, 2003).

1.3.2.5 *Age and gender*

Other widely studied predictors of trait WTC include demographic variables, such as *age* and *gender*. A widely cited study conducted by MacIntyre, Baker, Clément, and Donovan (2002) investigated the joint effect of age and gender on trait WTC of young FSL learners in a Canadian junior high school. They found a significant increase in L2 WTC from Grade 7 to Grade 8, which was then maintained to Grade 9. Moreover, they found gender differences: (a) Girls were generally more willing to communicate in French than boys; (b) Girls showed an increase in trait WTC in French from Grade 7 to Grade 9, while boys remained stable. A later study by the same group of researchers found that gender differences in L2 WTC appeared among junior high school students, but not high school or university students (Donovan & MacIntyre, 2004).

Other studies investigating the relationship between age and trait WTC recruited adult learners; however, they tend to find no significant age-related differences in L2 WTC (e.g. Aliakbari & Mahjoob, 2016; Ghanbarpour, 2016; Oz, 2014). Ghanbarpour (2016) did not find significant differences in trait WTC between three groups of Iranian EFL university students: below 24, from 25 to 29, and above 30 years old. Similarly, another study conducted in Iran found no significant difference between first-year and fourth-year English major students' trait WTC (Ahmadian & Shirvani, 2012). Hence, it seems that young learners may have a critical period to develop their L2 WTC, after which trait WTC does not develop. To enhance university students' willingness to communicate in L2 classes, interventions may be used, such as systematically managing the classroom atmosphere to trigger state WTC.

Gender differences have been widely discussed by previous studies; however, few have demonstrated these. It seems there is no significant gender difference in L2 trait WTC, although some studies claim that female learners are more willing to communicate. For instance, with a sample of ESL high school learners in Canada, Baker and MacIntyre (2000) found no significant gender difference in trait WTC (in either L1 or L2), although girls did show higher motivation and attitudes towards L2 learning and were more socially oriented than boys. Similarly, by studying a group of non-English major undergraduate students in China, Li (2012) found that female students' trait WTC in English was slightly higher, but the difference was not significant. These findings indicate that there might be a gender difference in L2 learners' willingness to communicate, but it should not be seen

as a major explanation for individual differences in trait WTC.

1.4 Summary

This chapter has introduced a definition of WTC in the L2 context and reviewed previous research on trait WTC. WTC here particularly refers to L2 learners' intention to talk with peers and teachers in L2 classrooms. The review of research on trait WTC suggests that trait WTC can be influenced by personal characteristics (e.g. L2 self-confidence, motivation, personality traits). When investigating individual differences in trait WTC, some studies noticed that state WTC in different situations may deviate from trait WTC. Such within-person variability is a relatively new concern for SLA research; however, it is in line with the current trend of incorporating within-person perspectives to study individual differences in the field of personality science. To provide insights into how within-person variability in state WTC can be studied, the next chapter will refer to some dynamic approaches in personality literature and demonstrate their implications for SLA research.

2 Dynamic Approaches to Personality and Their Application to SLA

2.1 Introduction

As reviewed in Chapter 1, many studies have examined WTC at the trait level, and both the definition and heuristic model of WTC indicate that state WTC in specific situations may vary over time. Some recent studies (e.g. MacIntyre, 2012; Mystkowska-Wiertelak & Pawlak, 2017; Peng, 2014) have suggested paying attention to the dynamic nature of state WTC as a complement to research on trait WTC. Although this inclusion seems innovative for SLA research, it is consistent with mainstream research in personality science (e.g. Beckmann & Wood, 2017; Fleenon & Jayawickreme, 2015; Minbashian, Beckmann & Wood, 2017; Rauthmann et al., 2015; Rauthmann & Sherman, 2016a). Hence, it is expected that a review of personality research with a dynamic perspective will productively inform research on state WTC. This chapter will review the use of dynamic approaches in personality literature and outline their application to SLA research.

This chapter will firstly introduce the dispositional and dynamic approaches to personality and the current trend of integrating these by maintaining both their principles. Different approaches (e.g. the density distribution approach and the cognitive-affective personality system approach) used to capture within-person variability in personality states and investigate the meaningfulness of such variability will be investigated. As within-person variability in personality states seems to be systematically related to changes in situations, objective situation features, subjective situation perceptions, and situational taxonomies will be introduced. Finally, recent applications of the dynamic approach to SLA literature will be outlined.

2.2 Dispositional and Dynamic Approaches

There are two prevailing research approaches to personality and individual difference in personality science. One is the dispositional or *trait approach*, emphasising relatively fixed personality traits. This is represented by Goldberg's (1992) Big-Five personality traits, which consistently differentiate one individual from another. This approach has provided a useful framework for personality psychologists to describe differences between individuals' *typical* thoughts, feelings and behaviours. Personality research is largely based on these individual differences, i.e. relatively stable personality traits that describe a person in relation to others. WTC research in the L1 communication context

belongs to this approach, as it regards WTC as a stable and trait-like predisposition that describes general communication tendencies.

However, within the trait approach, less attention has been paid to the fluctuations in thoughts, feelings and behaviours *within* a person (i.e. within-person variability over time and across situations) and these are of great interest (Beckmann, Minbashian, & Wood, 2011). Epstein (1994) criticises the trait approach as merely describing a general behavioural tendency, rather than the whole process of individual behaviour generation. For example, via the trait approach, whilst we may find that one is extrovert or talkative in general, we would not be able to know, for instance, what makes people decide to actively engage in verbal communication at a specific moment, and why they appear to change their state WTC and communication behaviour over time. The trait approach merely provides “a description of the surface attributes” (Epstein, 1994, p. 121).

To describe and explain within-person variability, the concept of a personality *state* has been introduced. Researchers should always specify whether they are referring to trait or state level when discussing personality-related constructs (Fleeson, 2017). Whilst a personality trait emphasises stable individual differences in personality, a personality state is one’s thoughts, feelings, and behaviours at a given moment in time, which might vary meaningfully over short periods. The *processing approach* in personality literature suggests studying within-person variability in personality states that underlie behaviours across different situations (Beckmann, Beckmann, Minbashian, & Birney, 2013; Beckmann, Wood & Minbashian, 2010). There has long been a debate on which approach is more appropriate for understanding human behaviours. Although research on personality is primarily based on between-person comparison, a purely trait approach lacks insight. This is because it pays little attention to how variables operate and interact within individuals, i.e. whether thoughts, feelings and behaviours vary over time and across situations (Epstein, 1994). Fleeson (2001) and others (e.g. Judge, Simon, Hurst & Kelley, 2014; Sherman, Rauthmann, Brown, Serfass & Jones, 2015) have shown that there is large within-person variability in trait-related personality states that is highly reliable and can capture individuals. Within-person variability in personality is too significant to be neglected as an error variance in analyses.

Although seemingly conflicting, the two approaches are not incompatible. Researchers have increasingly realised that both traits and states are crucial for a comprehensive

understanding of personality, and suggested integrating them by maintaining both of their principles simultaneously when studying personality and individual differences (Beckmann et al., 2010; Beckmann et al., 2013; Debusscher, Hofmans, & De Fruyt, 2016; Fleeson, 2001; Fleeson & Jayawickreme, 2015; Fleeson & Leicht, 2006; Mischel & Shoda, 1998). By incorporating a within-person perspective to investigate psychological processes and structures at the individual level, personality research goes beyond just describing how people differ to explaining why they differ by understanding underlying behaviour generation processes (Fleeson, 2017). The following section will elaborate on two approaches used for incorporating a within-person perspective into research of personality and individual differences.

2.3 Density Distribution Approach

Fleeson and colleagues (Fleeson, 2001; Fleeson & Leicht, 2006) integrated trait and processing approaches by suggesting density distributions as complete descriptions of individuals' day-to-day variability in trait-relevant personality states. According to Fleeson (2001, 2017), everyone's behaviour over time and across situations forms a distribution, and there are stable individual differences not only in central tendencies (means), but also in sizes (standard deviations) and shapes (skews and kurtoses) of distribution. Two extreme distributions are presented in Figure 2.1 to illustrate these.

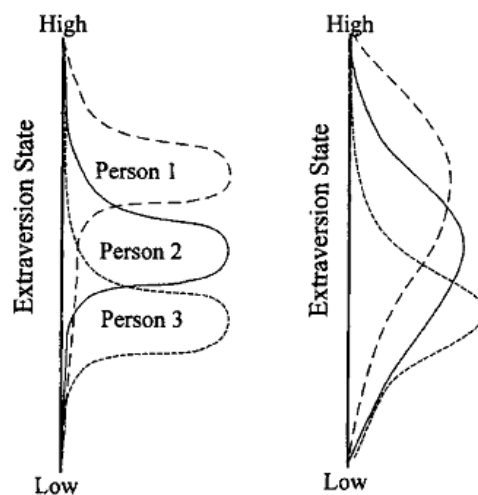


Figure 2.1. Fleeson's (2001) two extreme distributions of states over time (p. 1012)

The three narrow distributions on the left show low within-person variability and little overlap between individuals. Thus, the central tendency (or mean) could characterise each

individual. However, the three distributions on the right are wide and present considerable within-person variations, suggesting that everyone can manifest all levels of the trait. That is, individuals can behave similarly sometimes or in some situations, and thus the central tendency is a less appropriate description. As within-person variability in personality is realised in everyday thoughts, feelings, and behaviours, it is hypothesised that personality traits are manifest in relevant states (Fleeson, 2001, 2017; Fleeson & Leicht, 2006). Likewise, Beckmann et al. (2010) suggest that to study personality at the within-person level, individual different variables need to be conceptualised as state variables. As Big-Five personality traits are widely used to characterise central behavioural tendencies, the same scales have been adopted to characterise affective, cognitive and behavioural states.

To model density distributions of Big-Five-relevant states, Fleeson (2001) employed experience sampling methodology. Focusing on within-person variability in states over time and across different situations, this is an approach employing repeated measurements of the same group of participants in their daily life at different times of the day for several weeks (Barrett & Barrett, 2001; Fisher & To, 2012). Although collecting self-reported data, it does not rely solely on participants' memory, but assessing momentary thoughts, feelings, and behaviours when responding to the self-report measure and capturing situational features in a natural context (Barrett & Barrett, 2001). Consequently, it is "more ecologically valid and less prone to the biases in recall that affect retrospective reports of behaviour" (Beckmann et al., 2011, p. 3). In his original study Fleeson (2001) found that (a) within-person variability in trait-related personality states and affect is systematic, meaningful and substantive; (b) both average behavioural tendencies and behavioural variability show stable individual differences; and (c) within-person variance is not only caused by situational variation, but also by interaction between personality and the situation. More recent studies have been carried out using experience sampling methodology, confirming that there is systematic within-person variability in personality states (e.g. Church et al., 2013; Judge et al., 2014; Wilson, Thompson & Vazire, 2017).

Although originally tested with Big-Five-relevant states, the usage of the density distribution approach and experience sampling methodology has not been restricted to research on the five major facets of personality. Fleeson and Leicht (2006) indicate that the density distribution approach extends to variables beyond the five major personality traits. Similarly, it is suggested by Beckmann et al. (2011) that this approach could be adopted to evaluate other relevant personality constructs.

When the same group of individuals are repeatedly measured over time, data involves hierarchical structures as different occasions are nested within individuals (Raudenbush & Bryk, 2002). Hence, data collected through approaches such as experience sampling is better viewed as having two or three levels of hierarchy. The states observed over time at Level 1 are nested within individuals at Level 2, who are also nested within classes or schools at Level 3 (Raudenbush & Bryk, 2002). Failing to address these hierarchical structures in data may cause problems, such as aggregation bias. For example, a figure presented by Schmitz (2006) compares four individuals' learning trajectories to an aggregated trajectory. As shown in Figure 2.2, the aggregated trajectory shows four abrupt changes, two rises followed by two falls, but no single individual trajectory shows more than one change. Similarly, de Vaus (2001) has suggested studying changes found through longitudinal research at both the aggregated and individual levels. When aggregating different individuals, changes in some can be cancelled out by changes in others. That is, a low level of aggregate change cannot be easily interpreted as an absence of individual changes. Merely studying the aggregated trajectory would be misleading because changes at the individual level can be overlooked.

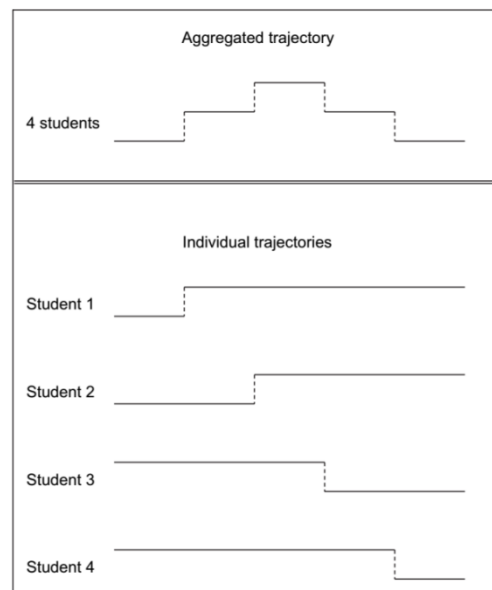


Figure 2.2. Schmitz's (2006) comparison between aggregated and individual trajectories (p. 439)

The aggregated level and the individual level should be clearly distinguished, and data should be analysed separately (see de Vaus, 2001). With a hierarchical linear model, the relationships between dependent and independent variables, and the variation occurring at each level can be formally depicted by equations in a sub-model, and the relations

between different levels within the model can be specified (Raudenbush & Bryk, 2002). Hierarchical linear models have been widely applied to study the influences of situational antecedents on trait-relevant personality states such as conscientiousness and neuroticism (e.g. Beckmann et al., 2010; Huang & Ryan, 2011; Minbashian, Wood, & Beckmann, 2010). Details of how to establish hierarchical linear models will be presented in section 6.2.8.

2.4 Person-Situation Interaction

As has been discussed earlier, some researchers in the field of personality science have suggested that both enduring personality characteristics of a learner and perceptions of learning situations impact on behaviour and learning performance (Nijhuis, Segers & Gijsselaers, 2007; Ntalianis, 2010). Nevertheless, previous research has mainly focused on the role of fixed personality traits in predicting students' learning, and scant empirical research has investigated the impact of the learning situation (Pawlowska, Westerman, Bergman & Huelsman, 2014) or the interaction between personality and the learning situation (Nijhuis et al., 2007). It is necessary to understand the latter because it is argued that learning would be most successful when certain personal characteristics match situations (Pawlowska et al., 2014).

2.4.1 If ... then patterns

Variation in personality states can to some degree be explained by the impact of situations (see Fleeson, 2001, 2007; Fleeson & Leicht, 2006). Being aware of individual differences in behavioural responses to the same situation, Fleeson (2001) attributed them to differences in individuals' sensitivity to given situational properties. However, Mischel and Shoda (1998) suggest psychological situations and an individual's personality system jointly influence the process of behaviour generation. As there are individual differences in personality, individuals may respond differently or even inversely to the same situational properties.

Mischel and Shoda (1998) established a link between trait and processing approaches by explaining how within-person processes develop into dispositions that can explain behaviour. They suggested that situation and personality jointly influence the process of behaviour generation within a person. In their *cognitive-affective personality system* (CAPS), individual differences are characterised by the available cognitive and affective

units. When an individual is experiencing a certain situation, a subset of interconnected thoughts and feelings will be activated, which will then lead to a corresponding behavioural response (Mischel & Shoda, 1998). For instance, when a L2 learning situation is pleasant, positive affect and goals within a student may be activated, which then may trigger state WTC and communication behaviour.

The dynamic nature of the behaviour generation process does not imply that the cognitive and affective structure within an individual is unpredictable. On the contrary, relatively stable *if...then patterns* do exist, which determine relationships between situations and behavioural responses (Mischel & Shoda, 1998). When *if* (the situation) changes, so does *then* (the behavioural response). For example, in a L2 classroom, when the teacher is nearby, a learner may feel extremely anxious and fearful of making errors, thus remaining silent. However, when the teacher is not present, the learner may feel more at ease and thus become more willing to communicate in the target language. Such *if...then patterns* are rooted in CAPS and work as the foundation of within-person variability across different situations, indicating the consistency of the same individual's behavioural responses within the same kind of situations over time.

As individuals' cognitive and affective structures can differ, what captures personality are underlying patterns of behavioural responses across situations, i.e. each individual's "characteristic behavioural signature" (Shoda, Mischel & Wright, 1994, p. 683). If one's personality system is unchanged, the person's *if... then patterns* remain stable. In other words, rather than negating stability in personality, CAPS is a meaningful construct that manifests personality coherence and consistency (Shoda et al., 1994). Unlike traits that are arguably rooted in biology, states reflect cognition and affect that are likely more modifiable (Minbashian et al., 2010). Compared to central behavioural tendencies, *if ... then patterns* are not only more accurate indicators of personality, but also of more practical importance for researchers and practitioners who aim to elicit behavioural changes to facilitate learning. Additionally, although the person-situation interaction is at the within-person level, it is not necessarily to be studied individually but can be investigated interpersonally by categorising individuals who share the same underlying patterns into types (Mischel & Shoda, 1998). To categorise *if ... then patterns* requires identifying the crucial situational features that can trigger within-person variability, and measuring them together with behaviour, so to understand the underlying cognitive-affective structure (Fleeson, 2001; Shoda et al., 1994).

2.4.2 Situation-based contingencies

Stable *if...then patterns* in CAPS suggest another way to show stability in personality, i.e. how an individual typically responds to certain kinds of situations. It is believed that within-person variability in personality states is a meaningful psychological process rather than random error variance, as it may be systematically explained by situational changes (Fleeson, 2007). To understand this systematicity, the construct of a *contingency* was introduced. A contingency is “a systematic relationship between a given state and a given situation characteristic” (Fleeson, 2007, p. 830). It is the coefficient that estimates the co-variation between a state and a situational variable within an individual rather than across different individuals.

It is suggested that contingencies can be studied at both the aggregated and individual levels. At the aggregated level, one can estimate the average contingencies for a *typical individual*, while at the individual level people may differ in contingencies because of the person-situation interaction (Fleeson, 2007). That is, some people are more likely than others to adapt their thoughts, feelings and behaviours to changes in situations, and thus have larger amounts of within-person variability in their personality states. Evidence has been provided by Fleeson (2007), showing that Big-Five personality states are systematically contingent on some situation characteristics, and individuals differ in their contingencies. Other researchers have conducted studies using similar approaches to support Fleeson’s (2007) findings, demonstrating that situation-based contingencies can be used to characterise personality (e.g. Huang & Ryan, 2011; Judge et al., 2014; Sherman et al., 2015; Wood, Beckmann, Birney, Beckmann, Minbashian & Chau, 2018). Details of how to estimate contingencies will be elaborated in section 6.2.8.2.

2.4.3 Psychological characteristics of situations

The findings presented above show that within-person variability in trait-related personality states is meaningful as it is associated with changes in situation characteristics. This section will define what situation characteristics refer to, and present two taxonomies.

Rauthmann et al. (2014, 2015) suggested that previous research on situation uses inconsistent language (such as using the term ‘features’ to indicate ‘cues’ or ‘characteristics’). The inconsistency in the literature and lack of distinction between different situation concepts may hinder communication and development in the field.

Hence, Rauthmann et al. (2014, 2015) categorised situational information into three levels (i.e. classes, cues, and characteristics), and suggests future research should discriminate between these. Situation *classes* describe types of situations (e.g. study or work situations), and situation *cues* refer to the physical elements that comprise a situation, including interlocutors, tasks, etc. Situation cues are objective features of situations, whereas the processing of these objective features depends on how they are interpreted by individuals, and thus creates subjective perceptions of situations, i.e. situation *characteristics* (e.g. task-confidence, task-interest, and task-usefulness). Situation characteristics refer to people's idiosyncratic perceptions and interpretations of situations, offering a psychological dimension to situations. It is believed that every external variable (here the situation) has to interact with personality to influence behaviour, and thus individual differences exist in perceptions of objectively similar situations (Fleeson, 2017).

While psychological characteristics of persons have been thoroughly analysed, with the Big-Five being regarded as the most widely accepted taxonomy, psychological characteristics of situations have been sparsely studied and no unified taxonomy has been found (Rauthmann et al., 2014). Rauthmann et al. (2014) therefore proposed a taxonomy of situation characteristics named 'Situational Eight DIAMONDS', in which situation characteristics were categorised into eight major dimensions: Duty (Does work have to be completed?), Intellect (Does the situation require deep thinking?), Adversity (Is someone being blamed?), Mating (Are there potential romantic encounters?), pOsitivity (Is it a pleasant situation?), Negativity (Does the situation cause negative feelings?), Deception (Is someone being deceived?), and Sociality (Are there opportunities for social interaction?) (Rauthmann & Sherman, 2016b). Research has been conducted to test the effects of 'Situational Eight DIAMONDS', as well as personality traits, on trait-related personality states. It was found that (a) personality states could be predicted by their corresponding traits; (b) there were between-person relationships between personality states and situation characteristics; (c) there were within-person relationships between personality states and situation characteristics (i.e. contingencies); and (d) individuals differed in their situation-based contingencies (Sherman et al., 2015).

Other than 'Situational Eight DIAMONDS', there is a more recent taxonomy of situation characteristics proposed by Parrigon, Woo, Tay and Wang (2017). This is referred to as 'CAPTION': Complexity (How complex is the situation?), Adversity (How depleting

and/or difficult is the situation?), Positive Valence (How positive is the situation?), Typicality (Is the situation common and straightforward?), Importance (Will the situation lead to the achievement of one's goals?), Humour (How humorous is the situation?), and Negative Valence (How negative is the situation?). Although using a different methodological approach, 'CAPTION' overlaps with 'DIAMONDS' in dimensions such as Positivity and Negativity (see Parrigon et al., 2017 for details). However, as the 'CAPTION' taxonomy was published when the current research was being conducted, Rauthmann et al.'s (2014, 2015) terminology will be used throughout this thesis for consistency.

2.5 Application of Dynamic Approaches to SLA

Similar to research in personality science literature, there is also a growing interest in dynamic phenomena in second language acquisition. As Larsen-Freeman and Cameron (2008) suggest, applied linguistics can be characterised as an interplay of dynamic systems. For example, when learning a language, a set of individual and situational variables interact, jointly affecting the learning process. Hence, language is a dynamic system, and the process of language learning is dynamic in nature (de Bot, Lowie & Verspoor, 2007). The dynamic systems perspective considers interactions among different influencing factors in a specific situation (e.g. a L2 classroom), instead of analysing linear relationships between variables in isolation as the dispositional approach tends to do.

Although the dynamic systems perspective is relatively new in SLA, it has received growing attention. Dörnyei, MacIntyre and Henry (2015) have adopted a dynamic systems perspective to conceptualise L2 motivation. They suggest that research interests have shifted from linear relationships between motivational dispositions to a more dynamic perspective, such as fluctuations in L2 motivation across different situations and the impact of these on L2 behaviour. More recently, Mystkowska-Wiertelak and Pawlak (2017) have taken this a step further, integrating a macro and micro-perspective to provide a comprehensive interpretation of both trait and state WTC. The macro-perspective focuses on the linear relationships between trait WTC and relevant dispositions, while the micro-perspective is more context sensitive exploring state WTC in specific situations and its fluctuation over time. Most previous empirical studies on WTC adopted a macro-perspective, but the dynamic nature of WTC (i.e. state WTC) cannot be studied without adopting a micro-perspective (Mystkowska-Wiertelak & Pawlak, 2017). Therefore, the

traditional dispositional approach needs to be supplemented by a dynamic systems perspective (Dörnyei et al., 2015).

2.6 Summary

This chapter has reviewed the integration of the dispositional and dynamic approaches in personality science literature, which corresponds to the current interest in the state level of WTC and its within-person variability. The density distribution approach that features repeated measurements of the same group of participants over time has been introduced as a useful approach to capture within-person variability. It has been accepted that both personal and situational variables jointly influence within-person variability in personality states. With the current focus on situational antecedents, the classification of situation information has been adopted in order to distinguish the subjective perceptions of situations (i.e. situation characteristics) from the objective features of situations (i.e. situation cues), and valid taxonomies of situation characteristics in personality science literature have been introduced. Moreover, the application of the dynamic approach to SLA has been reviewed. The literature has pleaded for a dynamic perspective to study constructs related to L2 learning, such as L2 motivation and WTC. As a result, this research not only studies WTC at its trait level, but also adopts a dynamic perspective to explore potential variability in state WTC.

2.7 Research Aim and Questions

This research concerns variability and stability in both trait and state WTC. However, the main focus is on within-person variability in state WTC and its underlying stable patterns (i.e. whether within-person variability in state WTC is systematically associated with changes in learning situations). Like research in personality literature, research on WTC has indicated that communication generation is a psychological process jointly influenced by personality and the situation (MacIntyre et al., 1998). However, compared to the impact of personality on trait WTC, the impact of situation or person-situation interaction on state WTC has seldom been studied (Peng & Woodrow, 2010). To fill this gap, this research pays attention to the underlying situation characteristics (subjective perceptions of situations) that may influence state WTC in L2 classrooms. If state WTC shows within-person variability systematically related to selected situation characteristics, then a concern is whether this has an impact on L2 learners' language learning performance. The

goal is to demonstrate whether adopting such a dynamic perspective is meaningful for research on WTC and L2 teaching in real classrooms.

The main questions addressed in this research are as follows:

1. Variability: How variable is learners' WTC in language learning classrooms in a Chinese EFL context?
2. Systematicity: How is the variability in WTC related to the learners' subjective perceptions of the language learning classroom situation in a Chinese EFL context (i.e. situation characteristics)?
3. Predictability: Does the systematic variability of WTC predict language learning performance?

3 A Systematic Review of Situational Antecedents of State WTC

3.1 Introduction

The literature reviewed in Chapter 2 suggests that personality states may vary over time, and this within-person variability can be systematically related to changes in situations. Hence, investigating whether, how, and why learners show more state WTC in some situations than in others could be informative for researchers as well as practitioners. A small group of recent studies have explored the potential within-person variability in state WTC and identified several situational variables (e.g. interlocutors, tasks, etc.) that could affect state WTC; however, findings are still fragmented and unsystematic. These studies focused on selected situational variables and seemed to ignore the distinction between different kinds of situational variables, i.e. the necessary differentiation into situation cues and characteristics (see section 2.4.3). In addition, they have used various labels to represent the same situational variable, or used the same label yet with reference to different variables. The inconsistency in terminology use may hinder future research on state WTC. Hence, it is believed that a systematic review of literature on state WTC and its situational antecedents will benefit the variable section of the empirical studies and contribute to communication and development in the field.

This chapter will systematically review published research that has paid attention to fluctuations in state WTC across different situations, particularly focusing on frequently reported situational variables that might cause within-person variability in state WTC. Different kinds of situational antecedents of WTC (e.g. situation cues and characteristics) will be organised into a multi-layered framework. This aims at (a) proposing a consistent terminology for future research on state WTC and its situational antecedents, (b) achieving more conceptual clarity regarding the different types of situational antecedents of WTC, and (c) providing guidance for future research into the dynamic processes underpinning state WTC. This systematic review has already been published (Zhang, Beckmann & Beckmann, 2018).

3.2 Inclusion Criteria

To identify the full breath of relevant studies, I searched databases through Web of Science and EBSCO (which includes ERIC and the British Education Index) up until July 2017 using the key words: willingness to communicate (WTC) and second language (and

its variations: L2, foreign language, English, EFL, ESL, FSL). The search was limited to publications in English since 1996 when WTC was first introduced to SLA by MacIntyre and Charos (1996).

Altogether, 219 studies on L2 WTC were found. Most of the studies were concerned with trait WTC. The authors tried to isolate different variables to determine their correlations with WTC at the trait level (e.g. MacIntyre et al., 2003; Peng, 2007a; Yashima et al., 2004). However, not all studies clearly distinguished between trait and state WTC. Some saw WTC as a dynamic phenomenon and investigated fluctuations in WTC across situations, though they did not specify that they were interested in state WTC or the dynamic nature of WTC. Hence, rather than using more specific search terms (such as state WTC), the studies on state WTC were manually selected from the complete list of WTC studies by reading the abstracts and checking the full articles when needed. It was found that 26 empirical studies discussed the dynamic nature of WTC and its situational antecedents. In addition, the references of the 26 studies were examined in order to identify other relevant studies that were not included in the above databases. Nine additional studies were found through such snowballing and included in the analysis.

All studies that investigated situational components of state WTC were included in this review. The 35 empirical studies were analysed in terms of the situational variables reported and the study's methodological quality. Findings and discussion concerning state WTC and its antecedents were extracted from each study. In these studies, a wide range of situational antecedents, both situation cues (objective features of situations) and situation characteristics (subjective perceptions of situations), were addressed. The situational antecedents were coded into different categories and sub-categories, and arranged into different levels. No study was excluded from this review; however, the methodological quality of the studies was taken into consideration in judging the strength of the evidence. The methodological quality of the studies was analysed in relation to the research designs and data collection methods. Findings as well as methodological information of the studies will be discussed in the following section. A list of the studies and the methodological approaches adopted is included in Appendix A.

3.3 Methodological Considerations

Participants in studies on state WTC are mainly university students, including undergraduates and adult language learners in university-based language courses. Four

exceptions include a study of primary school students aged between eight and nine (Buckingham & Alpaslan, 2017), two studies of adolescents in secondary schools (Joe et al., 2017; MacIntyre, Burns & Jessome, 2011), as well as a case study of a Korean physician in the US (Kang, 2006). Most studies have been conducted with participants from Asian countries (China, Korea, Japan, Iran and Turkey), and only a few studies (eight studies) included participants from other countries, such as Canada, Poland, and Australia.

Amongst the 35 studies considered, there are eight cross-sectional survey studies, with relatively large samples (ranging from 101 to 2,156). For example, Peng and Woodrow (2010) studied how WTC was affected by the Chinese EFL classroom environment, by employing a questionnaire with 579 university students from eight universities in eastern China. Although it was not explicitly stated that WTC was studied at the state level, these studies, to some extent, paid attention to the impact of situations upon WTC. However, as WTC was only measured once in these studies, it is difficult to determine how it might fluctuate over time and how such fluctuations might be causally linked to the changes in situations.

Nearly half of the 35 studies (i.e. 15) are small-scale studies, among which ten explicitly state that they are case studies. From a single case or a small number of cases (not more than twelve), a large amount of data was collected using various data collection methods (e.g. simulated recall interviews, observations, reflective journals), and the majority of these (i.e. 11) use longitudinal designs with durations ranging from a few weeks to several months. For instance, Zhong (2013) studied five Chinese students in a language school in New Zealand for 18 weeks using semi-structured interviews, classroom observations, stimulated recall and learning logs. The sample sizes as well as the particular settings in which these studies were conducted impose a challenge to the generalisability of the respective findings. Such studies, however, have the potential to provide some orientation in generating testable hypotheses with regard how state WTC might fluctuate across different situations and in terms of what learner and/or situation characteristics might trigger such fluctuations.

Only three relatively large-scale longitudinal studies were found (MacIntyre et al., 2011; de Saint Léger & Storch, 2009; Zarrinabadi, 2014). In Zarrinabadi's (2014) study, 50 English major undergraduates in Iran were asked to write focused essays over a six-week

period, describing the situations in which they communicated with their teacher in English. This study sought to establish how students' state WTC was influenced by the teacher. Another example is MacIntyre et al.'s (2011) study of 100 Canadian adolescents in a French immersion program. MacIntyre et al. (2011) showed that students' state WTC fluctuated across contexts, and gave numerous examples of different learning contexts; however, they did not clearly identify the underlying situational antecedents for these instances.

Interestingly, five recent studies, three small-scale studies (MacIntyre & Legatto, 2011; Mystkowska-Wiertelak, 2016; Pawlak & Mystkowska-Wiertelak, 2015) and two relatively large-scale studies (Mystkowska-Wiertelak & Pawlak, 2014; Pawlak et al., 2016), used high-density repeated measurement approaches to investigate moment-by-moment variability in state WTC. They repeatedly measured the same groups of students' state WTC (and other relevant variables) during very short periods, such as a task or a lesson. Among these studies, only MacIntyre and Legatto (2011) measured state WTC after communication, using video recordings to remind the participants of what was happening while they were communicating. They asked six female FSL university students to talk about eight topics in a few minutes and then watch video recordings of their performance, and rate moment-to-moment WTC during the task. All other studies measured state WTC at the same time when students were communicating. For example, Pawlak et al. (2016) asked a sample of 60 Polish undergraduates separated into four groups to report state WTC 13 times over a 60-minute period in class. While the students were engaging in a lesson, they rated state WTC every 5 minutes when hearing a computer-generated beep. Pawlak et al. (2016) found not only differences between the four groups, but also fluctuations over time both within one of the groups and within the individual members of that group ($n_{group1} = 12$). Although such high-density repeated measurement designs are relatively novel for research in SLA, such designs have been employed as a prime method in personality science literature (as outlined in detail in section 2.3). Such studies point to the dynamic nature of WTC and provide new and interesting insights for future research on state WTC and its situational antecedents.

Strictly speaking, to establish causal links between different situational antecedents and state WTC, the adoption of an experimental research design would be necessary. This is because the inclusion of at least one randomised control group allows controlling for effects of potentially confounding variables (i.e. those unrelated to the experimental

manipulation, yet with a potential impact on the outcome). This allows the researcher to more confidently attribute an effect to a particular cause (e.g. a situational factor) (de Vaus, 2001). However, in the course of the current research only four experimental studies with rather small samples were found in the literature, each focusing on a certain element of the classroom setting. For example, using a sample of 18 students, Yu (2015) investigated the effect of interlocutors' participation and cooperation on L2 learners' state WTC in dyadic interactions, and found that L2 learners' state WTC changed across interlocutors. The lack of experimental studies of sufficient size is one major challenge to validly establishing whether causal links exist between presumed situational antecedents and state WTC.

To offer a comprehensive review of the possible situational antecedents of state WTC, the following sections will first present the situation cues (according to Rauthmann's terminology) reported in the literature. The situation cues include interlocutors, classroom atmosphere, topic, and activity, which have been studied relatively widely. It will then present the situation characteristics, which are relatively under-explored, such as task-confidence, task-interest, and task-usefulness.

3.4 Situation Cues

3.4.1 Interlocutors

A situational variable proposed in the original model of L2 WTC is the specific person one is communicating with (MacIntyre et al., 1998). This person is commonly termed as the 'interlocutor'. Previous studies have shown that L2 learners' WTC is influenced by some characteristics of their interlocutors, such as familiarity with the interlocutors (e.g. Kang, 2005; Riasati, 2012), the interlocutors' participation and cooperation (e.g. Pawlak & Mystkowska-Wiertelak, 2015; Riasati, 2012), as well as other demographic features of the interlocutors (e.g. Cao, 2011; Eddy-U, 2015; Kang, 2005; O'Sullivan, 2002; Pawlak et al., 2016; Riasati, 2012).

3.4.1.1 *Familiarity, participation and cooperation*

Findings on interlocutor familiarity, participation and cooperation are relatively unambiguous. It has been found that students prefer talking with friends in comparison to strangers or acquaintances (e.g. Kang, 2005), and enjoy communicating with cooperative and participatory interlocutors (e.g. Cao & Philp, 2006; Kang, 2005; Pawlak &

Mystkowska-Wiertelak, 2015; Riasati, 2012). Familiar and cooperative interlocutors reduce learners' fear of speaking a L2; while interlocutors who actively participate in and contribute to communication make learners feel excited in and responsible for delivering information (Kang, 2005). The findings indicate that state WTC is not only influenced by one's relationships with the interlocutors and interlocutors' behaviour, but more importantly, affected by the person's subjective perceptions of the interlocutors' cooperation and contribution.

3.4.1.2 Demographic features

Some demographic features of the interlocutors, such as ethnicity (e.g. Cao, 2011; Kang, 2005, 2006), L2 proficiency (e.g. Eddy-U, 2015; Kang, 2005; Yu, 2015), gender (e.g. Eddy-U, 2015; Riasati, 2012), age (e.g. Riasati, 2012), and appearance (e.g. Kang, 2005), have been found to affect L2 learner's state WTC; however, relevant findings are still tentative and inconsistent. For instance, both Eddy-U (2015) and Yu (2015) found that L2 learners preferred interlocutors who had higher L2 proficiency, as they believed that more proficient speakers could help them with language learning. To note, both Eddy-U (2015) and Yu (2015) recruited non-immersion English learners who had few opportunities to communicate with native speakers. By comparison, findings reported by Kang (2005) in an immersion English learning context were less straightforward. On the one hand, students tended to be willing to communicate with native English speakers who could offer language support, which seemed to agree with Eddy-U's (2015) and Yu's (2015) findings. On the other hand, students noted that they were relatively unwilling to communicate with non-native peers who had higher English proficiency because of the feeling of insecurity and fear of 'losing face', which contradicted Eddy-U's (2015) and Yu's (2015) findings. As findings about the impact of interlocutors' demographic features on state WTC are still limited, conclusions cannot be easily drawn; however, it seems that what directly influences state WTC is not interlocutors' L2 proficiency as such, but whether L2 learners feel supported when communicating with these interlocutors.

3.4.2 Classroom atmosphere

In classroom settings, L2 learners' state WTC seems to be influenced by the classroom atmosphere. Several researchers (e.g. Eddy-U, 2015; Lee, 2009; Riasati, 2012) have shown that a positive and stress-free classroom atmosphere – conceptualised as being co-created by classmates who cooperate with each other, as well as teachers who support

their students – is likely to facilitate students' state WTC.

3.4.2.1 Classmates

It has been argued that, when students find that their classmates are actively engaged in class, their state WTC can be boosted (Peng, 2012). Nevertheless, it is noted that if a few students dominate the interaction in class, other students' state WTC and opportunities to participate will be dramatically reduced (de Saint Léger & Storch, 2009). Class cohesiveness has been suggested to contribute to higher state WTC and better performance in class (e.g. Dörnyei & Kormos, 2000; Khajavy, Ghonsooly, HosseiniFatemi & Choi, 2014; Peng, 2007b; Wen & Clément, 2003).

3.4.2.2 Class size

A link between class cohesiveness and class size has been suggested. As it would be harder to achieve close contact and cohesiveness among a larger group of students, a bigger class might reduce students' state WTC (Wen & Clément, 2003). To investigate the effect of class size on state WTC, Khazaei, Zadeh and Ketabi (2012) compared state WTC between three classes of 5, 10, or 15 adult EFL learners. They found that students in bigger classes felt more anxious and thus avoided communicating, whereas smaller classes offer more opportunities for interaction and built up students' confidence, thus facilitating state WTC.

3.4.2.3 Teachers

Some students perceive the teacher as a more influential factor in contributing to a supportive classroom atmosphere (Lee, 2009). Research (e.g. Cao, 2011; Fallah, 2014; Peng, Zhang & Chen, 2017) suggests that teacher support is mainly manifest in *teacher immediacy*, which refers to a teacher's verbal and non-verbal behaviour that reduce the distance and enhance close relationships with the students (e.g. encouragement, confirmation, and smile). Fallah (2014), for example, found that teacher immediacy indirectly affected state WTC through confidence and motivation. State WTC is also influenced by a teacher's *teaching style* and *classroom management* (e.g. Cameron, 2013; Peng, 2012; Riasati, 2012; Zarrinabadi, 2014). For instance, Zarrinabadi (2014) suggested that students' willingness to participate in a class communication activity is influenced by their teachers' topic selection, error correction and time given for task preparation.

However, different students may interpret the same behaviour of a teacher differently, especially in regard to non-verbal expressions. Hence, it should be noted that what directly influences state WTC might not be teacher immediacy as such, but students' subjective perceptions of teacher support based on teachers' behaviour.

3.4.3 Tasks

Task is considered as an overarching label to include all situation cues related to the work students are asked to do. Hence, situation cues related to either the content being discussed (i.e. the topic) or the design of the task (i.e. the activity) will be included in this section.

3.4.3.1 Topic

The thematic categories of topics have been found to influence L2 learners' state WTC. Students prefer topics that they are familiar with and interested in, which reduces the difficulty of communication and increases confidence and state WTC accordingly (e.g. Cao, 2011; Kang, 2005; MacIntyre & Legatto, 2011; Mystkowska-Wiertelak, 2016; Riasati 2012; Wolf, 2013). The attractiveness and familiarity of a topic is linked to a student's *topic relevant background knowledge* as well as the extent of *L2 vocabulary* that the student possesses (MacIntyre & Legatto, 2011; Mystkowska-Wiertelak, 2016; Pawlak & Mystkowska-Wiertelak, 2015).

3.4.3.2 Type of activity

A range of research has found that L2 learners' state WTC fluctuates across different types of activities (e.g. Cao, 2011; Eddy-U, 2015; Ghasemi, Kermanshahi & Moharami, 2015; Pawlak et al., 2016; Peng, 2012; de Saint Léger & Storch, 2009). Pawlak et al. (2016) and Eddy-U (2015) reported that L2 learners enjoyed game-like communication activities most, while Cao (2011) showed that learners preferred group projects. Instead of emphasising any particular type of activity, communication activities in classrooms have been roughly categorised into dyadic, group, and whole-class activities. Findings about which type of activity is preferred by students are not conclusive, but it seems that students prefer group activities with three or four interlocutors (e.g. Cao, 2011; Cao & Philp, 2006; Riasati, 2012). With a small number of peers, a group activity causes potentially less anxiety and offers more opportunities for students to communicate and

generate multiple perspectives (Cao, 2011). However, some students, especially those with lower language competence, tend to prefer dyadic activities (e.g. Cao, 2013; Mystkowska-Wiertelak, 2016) because the turn-taking in dyads is less competitive and makes students feel more obliged and less fearful; while others prefer whole-class activities because they believe that they can learn more through teacher-led activities than cooperative activities (e.g. Lee, 2009; Riasati, 2012; Zhong, 2013). Other than comparing between dyadic, group, and whole-class activities, Mystkowska-Wiertelak and Pawlak (2014) went further by comparing monologue and dialogue tasks. They found that L2 learners preferred monologues to dialogues, although the initially high state WTC in monologues tended to decrease during the task, whereas the initially low state WTC in dialogues tended to increase. To sum up, students seem to be more willing to communicate in activities that are less competitive and more effective in facilitating language learning, and state WTC within an activity may change over time.

3.4.3.3 Preparation time

The time given for task preparation has been suggested to be another contributor to activity participation (e.g. Freiermuth & Jarrell, 2006; Riasati, 2012; Zarrinabadi, 2014; Zhong, 2013). For example, students regard simultaneous conversations (e.g. face-to-face talking) as more demanding and thus show lower levels of state WTC than in written communications (e.g. online chatting) because they do not have enough time to formulate opinions, search for appropriate vocabulary, and check the grammar (Freiermuth & Jarrell, 2006; Zhong, 2013). Providing students with sufficient preparation time is likely to raise their confidence and state WTC in communication activities (Riasati, 2012).

3.4.3.4 Assessment

Whether the performance in an activity is assessed or not is another factor that might influence state WTC. Some researchers (e.g. Riasati, 2012) suggest that when students are being assessed, they would be more anxious and thus reluctant to communicate. However, others (e.g. Eddy-U, 2015) argue that assessment is the only antecedent that contributes to pressure, but at the same time prompts state WTC, because students might see the grades as requirements or short-term goals that motivate them to overcome negative feelings (e.g. anxiety). However, as the relevant research is rather limited and the sample sizes of the existing research tend to be rather small (not more than 25 participants), there is a need for future studies to clarify the impact of assessment on state

WTC in class communication activities.

The antecedents previously discussed are the situation cues, and there are many studies available investigating them. It should be noted that these objective features are effective only as they are subjectively perceived by learners. For example, individual learners might interpret a teacher's behaviour differently. However, the evidence base for understanding situation characteristics and their effects on state WTC is still limited.

3.5 Situation Characteristics

3.5.1 Task-confidence

In MacIntyre et al.'s (1998) heuristic model, *state communicative self-confidence*, which is defined as "a momentary feeling of confidence, which may be transient within a given situation" (p. 549), is one of the immediate precursors of WTC. Previous studies have suggested that lacking confidence in task performance has a detrimental effect on state WTC (Cao & Philp, 2006; Riasati, 2012). A lack of confidence is often underpinned by a fear of making errors and being negatively evaluated by others, preventing students from speaking a L2 (e.g. Eddy-U, 2015; Kang, 2005; Riasati, 2012).

However, the terminology used in relation to confidence varies across studies. Examples of confidence-related terminology used in various studies are: *security* (e.g. Kang, 2005), *ease* (e.g. Liu & Littlewood, 1997), *anxiety* (e.g. Liu, 2002), and *embarrassment* (e.g. Liu, 2002). Most of these studies are based on relatively small samples (ranging from three to 25 participants). For example, after interviewing a group of 25 Chinese residents in Macau, Eddy-U (2015) found that most of the confidence-related responses were linked to a fear of making errors, though participants expressed it variously as anxiety, embarrassment, unease, etc. To my knowledge, the only large-scale survey study undertaken was conducted by Liu and Littlewood (1997). Collecting data from 437 university lecturers and 2,156 English learners in a university in Hong Kong, Liu and Littlewood (1997) were interested in why East Asian students were often seen as passive learners who tend to keep silent in class. They concluded that East Asian students were willing to communicate but experienced unease when speaking English, and this was closely related to their lack of confidence in English competence (Liu & Littlewood, 1997).

3.5.2 Task-interest

Another situation characteristic that might influence state WTC is task-interest, which is defined as the curiosity in and engagement with a specific task (Dörnyei, 2009). Some authors (e.g. Kang, 2005) refer to the feeling of elation when engaging in L2 communication as *excitement*, which might subsequently be related to task-interest. It could be argued that being interested in a task is a necessary (yet not sufficient) precondition for excitement; at the same time, previously experienced excitement might trigger interest in engaging in a similar task in future. Compared to findings related to negative affect (e.g. lack of confidence or fear), less is known about positive affect relevant to L2 communication, such as excitement and joy. In order to better facilitate language learning, research could focus more on identifying what affectively prompts – in addition to what hinders – L2 learners' state WTC. Strategies to prompt state WTC in L2 classrooms will be suggested at the end of this thesis (see section 8.7.2).

3.5.3 Task-usefulness

In some situations, even if students are not interested in a task, they may feel motivated by their perceptions of task-usefulness, which has been variously labelled as *task effectiveness* (e.g. Zhong, 2013), and *task orientation* (e.g. Khajavy et al., 2014; Peng & Woodrow, 2010). Kang (2005) conceptualised such perceptions as *responsibility*, which is particularly related to the purposes of maintaining some kind of interpersonal relationships and/or gaining personal benefits. Similarly, Bernales (2016) reported that students' L2 use was influenced by both self-imposed goals of becoming proficient in the L2 and teacher expectation. The motivation to use the target language in order to achieve personal goals and meet teachers' expectations suggested by Bernales (2016) seems to resonate with Kang's (2005) concept of 'feeling responsible to talk'.

3.6 Systematicity of Previous Studies

Although a number of situational antecedents of state WTC have been discussed in the literature, such discussions tend to ignore the necessary differentiation into situation cues and characteristics. Situation characteristics (subjective perceptions of situations) have neither received enough attention nor been clearly distinguished from situation cues (objective features of situations). Most of the situational antecedents that have been studied are situation cues, such as how many people are present in a given situation;

whereas situation characteristics, such as whether students feel supported when engaging in a task, have only been occasionally mentioned. As the objective situation cues have to be subjectively perceived by individuals, the subjective situation characteristics seem to be more important and direct predictors of state WTC and deserve more investigation. The conceptual distinction between situation cues and characteristics is best accommodated by a multi-layered framework.

However, to my knowledge, only one study (Kang, 2005) distinguished clearly between situation cues and characteristics. Kang (2005) observed and recorded four Korean students' participation in an ESL module at a North American university over eight weeks. After each class, participants were asked to watch the recordings of that class and to retrospectively reflect on how their state WTC was affected at different points in time. Kang (2005) concluded that the underlying situation characteristics affecting L2 learners' state WTC are security, excitement, and responsibility. Each of the situation characteristics is influenced by various situation cues related to the topic, interlocutors, and conversational context.

Security is defined as feeling safe from the fear of making errors or losing face when communicating in L2. Kang (2005) found that security is mainly perceived based on the features of the interlocutors, such as familiarity with the interlocutors, support offered by the interlocutors, the number of interlocutors present, as well as the interlocutors' ethnicity and L2 proficiency. Learners' topic-related prior knowledge and the stage (e.g. at the beginning) and process of a conversation (e.g. after making errors) also influence the feeling of security.

Excitement refers to the feeling of elation about participating in communication, which is partly perceived based on the topic, such as the attractiveness of the topic and one's familiarity with the topic. The self-perceived level of accomplishment of the task also plays a role. Kang (2005) found that excitement is also perceived in response to situation cues related to the interlocutors, including the interlocutors' ethnicity, appearance, cooperation, and participation.

Another situation characteristic suggested by Kang (2005) is *responsibility*, the sense of duty to deliver or understand a message during communication. Kang (2005) suggested that responsibility is affected by the perceived usefulness, importance and sensitivity of the topic being discussed, together with one's prior topic knowledge. The number of

interlocutors present and the interlocutors' participation and cooperation also influence the perception of responsibility.

However, due to the small sample, Kang's (2005) findings might provide a limited basis for generalisation, and other antecedents that might influence L2 learners' state WTC in other contexts might remain unidentified. Nevertheless, Kang's (2005) attempt to systematically organise situation cues and situation characteristics in relation to state WTC provides a useful basis for future research.

Additionally, previous research has used various labels to represent the same situational antecedent or used the same label yet with reference to different antecedents. For instance, Peng (2007b) reported that L2 learners' state WTC was influenced by group cohesiveness and classroom climate. In her later article, Peng combined the two antecedents into one called classroom atmosphere, "the mood, emotions, or climate sensed and shared by the class group" (Peng, 2012, p. 208). However, for Riasati (2012), classroom atmosphere is co-created by the class group and the teacher. This inconsistency of terminology use is another challenge to research into state WTC and its situational antecedents.

3.7 Proposed Framework of Situational Antecedents of State WTC

To systematically organise major situational antecedents of state WTC as suggested by previous research and to provide a consistent terminology for future research, a multi-layered framework of situational variables is proposed. In the proposed framework, situational antecedents of state WTC are systemically combined into three interlinked layers, i.e. situation cues, situation characteristics and the underlying dimensions of situation characteristics (see Figure 3.1). The proposed framework emphasises the role of situation characteristics (subjective perceptions of situations) in influencing state WTC.

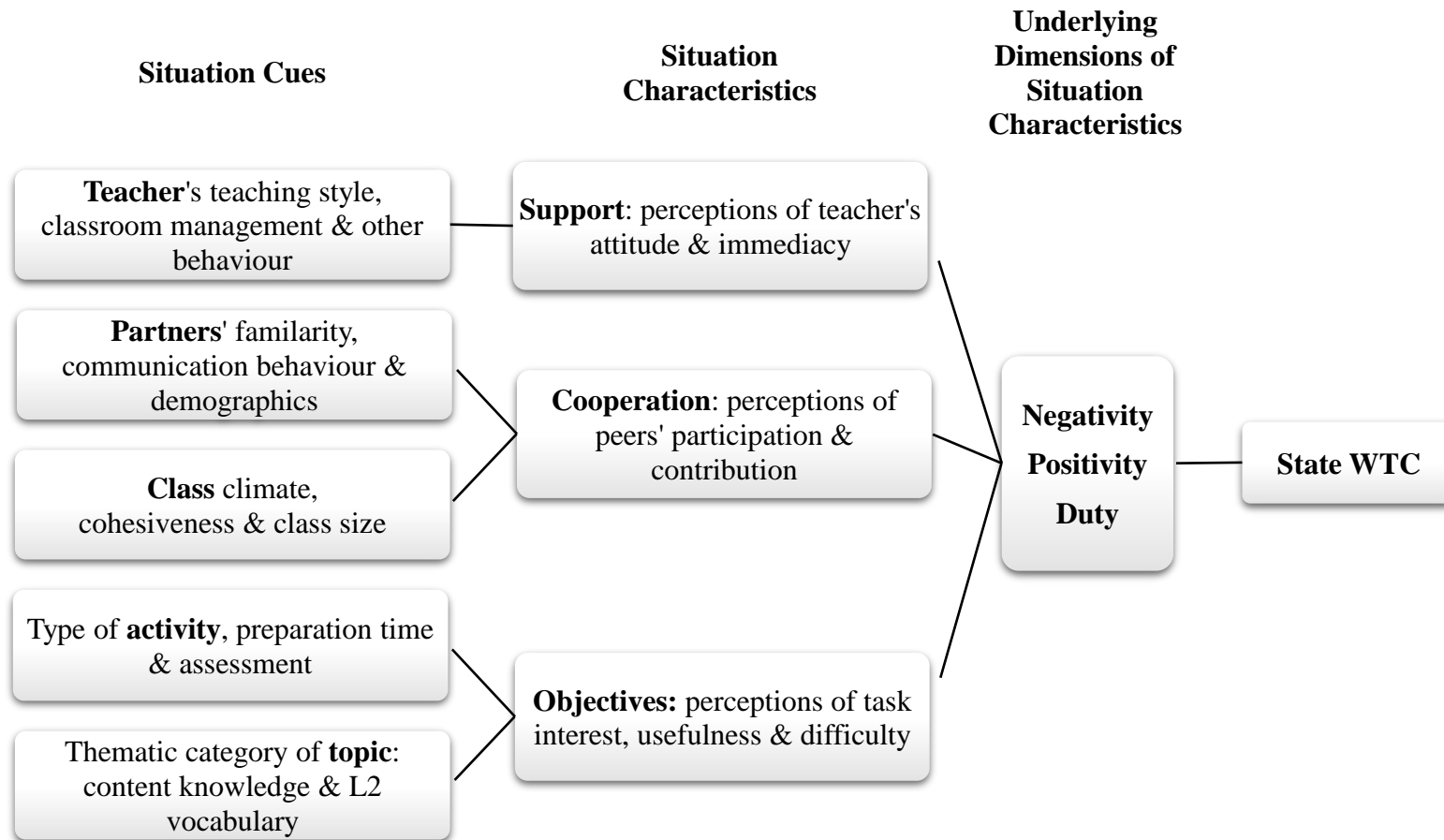


Figure 3.1. The proposed framework of situational antecedents of state WTC

3.7.1 Major dimensions of situation characteristics

Based on the review of prior empirical evidence, three out of Rauthmann et al.'s (2014) eight major dimensions of situation characteristics were deemed to be relevant to L2 learning situations and therefore selected, namely negativity, positivity, and duty. *Negativity* refers to “any sort of negative feeling (e.g. frustration, anxiety, tension, guilt, anger, etc.)” that may be elicited by a situation (Rauthmann et al., 2014, p. 708). It covers the lack of confidence or fear of making errors in using a L2 regularly mentioned in the literature (e.g. Cao & Philp, 2006; Eddy-U, 2015; Kang, 2005, 2006; Liu, 2002; Liu & Littlewood, 1997; Riasati, 2012). In contrast to negativity, *positivity* is suggested to represent the elation elicited by a situation, including “pleasant, fun, enjoyable, playful, simple, clear, and easy-to-navigate” (Rauthmann et al., 2014, p. 708). Because of the breadth of this concept, i.e. it captures any positive feeling elicited by the situation, both interest (Dörnyei, 2009; Eddy-U, 2015) and excitement (Kang, 2005), reported as situational antecedents of state WTC, can be seen as part of the positivity dimension. *Duty* refers to the extent to which students perceive a class to “contain work, fulfilling duties, attending to tasks, helping people with things, solving problems, and making decisions” (Rauthmann et al., 2014, p. 702). Rauthmann et al.'s (2014) concept of duty is parallel to that of Kang's (2005) responsibility.

3.7.2 Specific situation characteristics and cues

As the proposed framework focuses on the L2 learning situation, a number of situation characteristics directly relevant to class settings are specified as underlying the major dimensions. Based on the literature review, the specified situation characteristics are summarised as support, cooperation, and objectives. These situation characteristics are subjective perceptions of various situation cues, which are categorised into five themes (i.e. teacher, class, partners, activity, and topic). Kang's (2005) conversational context is excluded from this framework. Although similar situation cues, such as the stage of a task or class session, are also reported by Pawlak, Mystkowska-Wiertelak, and their colleagues (Mystkowska-Wiertelak, 2016; Mystkowska-Wiertelak & Pawlak, 2017; Pawlak & Mystkowska-Wiertelak, 2015; Pawlak et al., 2016), they are not commonly reported situation cues and existing findings are markedly inconsistent. For example, Mystkowska-Wiertelak (2016) reported that students' state WTC increased from the beginning to the middle of a class and then declined towards the end; while Mystkowska-

Wiertelak and Pawlak (2017) reported different tendencies with three groups of students: a consistently high level of state WTC throughout a class in Group 1, a gradually increasing trend in Group 2 and a slightly decreasing trend in Group 3. It may be argued that what actually influences state WTC is not the stage of a class, but one's perception of the learning situation at that particular point in time. As discussed by Mystkowska-Wiertelak (2016), students' relatively low state WTC at the start might be because they were waiting for the teacher to outline the lesson and present something interesting, while the downwards trend towards the end might be explained by fatigue after engaging in the learning/communication activities during the middle of the class. That is, students' momentary thoughts and feelings (e.g. task-interest) can be more important than the actual stage of a class in influencing state WTC.

Support refers to one's perception of teacher attitude and immediacy, which is perceived based on situation cues relevant to the teacher, such as the teacher's teaching style and classroom management (e.g. time for task preparation, topic selection, and error correction), as well as other verbal and non-verbal behaviour (e.g. smile, nod, and feedback).

Cooperation refers to one's perception of peer participation and cooperation. In whole-class activities, class cohesiveness, classroom climate, and class size might be influencing factors; while in dyadic or group activities, situation cues relevant to the specific interlocutors one is talking with might make a difference, including one's relationship with the respective partners, partners' communication behaviour and their demographic features (e.g. gender, ethnicity, and L2 proficiency).

Objectives refer to how one perceives the task according to the dimensions of task-interest, task-usefulness, and task-difficulty. Objectives are mainly perceived based on the type of the activity (e.g. dyadic, group, or whole-class activity), task preparation time, as well as assessment. The thematic category of the topic functions as another situation cue affecting one's perception of a task, as both content knowledge and topic-related L2 vocabulary might also be relevant.

3.8 Discussion

As previously discussed, L2 learners' state WTC and communication behaviour are interactively impacted by their general personal characteristics or traits (i.e. personality)

and perceptions of specific situations. However, compared to the great body of research on the personal antecedents that predict trait WTC, research on the situational antecedents that affect state WTC is more limited and still relatively fragmented. This chapter has suggested making clear distinction between trait and state WTC and reviewed the literature relevant to state WTC and its situational antecedents. To better facilitate future research, the underlying patterns common to previous findings have been identified, and the main situational variables have been integrated into a coherent multi-layered framework. This is the first attempt to arrange previously suggested situational antecedents of state WTC together.

The proposed framework systemically categorises situational antecedents of state WTC into three interlinked layers: situation cues (i.e. teacher, class, partners, activity, and topic), situation characteristics (i.e. support, cooperation, and objectives), and major psychological dimensions of situation characteristics (i.e. negativity, positivity, and duty). When summarising these, this study follows Rauthmann et al.'s (2014, 2015) situational taxonomy and applies their terminology to the literature of state WTC. The framework makes clear distinction between situation cues and characteristics, which is in line with Kang's (2005) study that differentiates the psychological antecedents of state WTC from physical situational variables. However, to my knowledge, no other research has made this meaningful distinction between subjective perceptions and objective features when studying the situational antecedents of state WTC. By suggesting a multi-layered framework, I hope to contribute to a better understanding of the different types of situational antecedents of state WTC and provide a consistent terminology for future research.

However, it is worth noticing that most of the reviewed studies were very small-scale studies (i.e. with less than twelve participants), and only a couple of them used high-density repeated measurement designs to track state WTC trajectories over time. Hence, one has to be very cautious when making use of the findings reported by a number of studies reviewed in this work. Situational variables that seemed to affect state WTC in some studies might not apply to other L2 learning contexts. The validity of the proposed framework of situational antecedents of state WTC has yet to be tested. Hence, the next step is to test whether the proposed framework is applicable to the specific context concerned in this research, i.e. English learning classrooms in Chinese higher education.

3.9 Summary

This chapter has presented a comprehensive review of published research that has paid attention to fluctuations in state WTC across different situations. Findings of 35 studies were analysed. These studies have observed within-person variability in state WTC and identified different types of situational variables related to this. These results have shed light on the first two research questions in a generic sense: (a) L2 learners' state WTC varies over time, and (b) this variability is systematically related to some situational characteristics. Situational antecedents of state WTC suggested in these studies were categorised into either situation cues or characteristics. These situation cues and characteristics were then systematically organised into a multi-layered framework, which may serve as guidance for future research on state WTC and its situational antecedents.

4 Philosophical Assumptions

Some researchers take the view that methodological choices are not only based on the research questions, but are also influenced by the philosophical assumptions that researchers bring to their work (e.g. Crotty, 1998; Hughes & Sharrock, 1997). Different philosophical positions reflect different world views and different ways of how knowledge is created to understand the world, which will lead to different research designs and methods employed to collect and analyse information. Hence, these authors claim that researchers should be aware of the assumptions they hold prior to conducting their research (Creswell, 2014; Creswell & Clark, 2011).

4.1 Ontological and Epistemological Positions

Ontology is concerned with ‘what is’ questions: what is believed to constitute reality or what is the nature of the world to be investigated (Blaikie, 2000; Crotty, 1998; Grix, 2001; Hay, 2002). In the social sciences, ontology is mainly concerned with whether the social world is independent of the individuals involved in it (Grix, 2001). Researchers hold different views on this question. A major ontological position is objectivism, which asserts the existence of facts or truths that are independent from individuals’ acknowledgements (Bryman, 2016). Hence, objectivists believe that the researcher’s job is to discover the truth. Constructivism, as an alternative ontological position, however, asserts that the social world is constructed by individuals’ social interactions and their subjective perceptions of them. This includes researchers who also bring their own meanings and understandings to their work (Matthews & Ross, 2010; Mertens, 2010).

These two ontological positions reflect different views on what we can know or what exists in the social world, whilst epistemological positions are concerned with the ways we gain knowledge about the social world (Grix, 2001; Hughes & Sharrock, 1997). Epistemology refers to a justification of what can be regarded as knowledge and how we acquire it (Matthews & Ross, 2010). In the social sciences, the major epistemological issue concerns whether the social world can (and should) be studied using the same principles, procedures, and ethos as the natural sciences (Bryman, 2016). Positivism is an epistemological position that develops from the objectivist ontological position, advocating that the approaches of the natural sciences can (and should) be adopted to study social phenomena. Research in a positivist tradition often tends to employ

quantitative methods for data collection. Although, it must be said that the use of either qualitative or quantitative information in research is not a useful differentiator for the underpinning epistemology (for instance, qualitative, e.g. interview-based data are also used in experiments, which arguably represent a research design following a positivist tradition). However, interpretivism (or constructivism) opposes the application of the approaches of the natural sciences to the social sciences because social phenomena are different from the objects of the natural sciences and may not be observable (Bryman, 2016; Grix, 2001; Matthews & Ross, 2010). Interpretivism prioritises individuals' subjective interpretations and understandings of social phenomena, which is often associated with the use of qualitative information as data in respective research (Creswell & Clark, 2011).

4.2 Question-Led Research

However, some authors challenge the necessity of identifying ontological and epistemological positions before conducting research, arguing that decisions about research designs and methods of data collection and analysis are primarily of a technical than a philosophical nature (e.g. Bryman, 1988; Robson, 2011; Gorard & Taylor, 2004; White, 2017). They believe that research questions should always be the starting point and different questions require different approaches (i.e. research designs). As particular research designs can only answer certain types of questions, positioning oneself as either an 'objective observer' or a 'subjective interpreter' may restrict the kinds of questions one can answer or lead to an unsuitable approach to answer the questions one attempts to answer. Hence, it is unnecessary and may even be counterproductive to identify oneself with a particular philosophical position before starting research (White, 2017). In short, ruling out certain forms of research designs (or data collection methods) based on one's ontological and epistemological position could render it impossible to answer (research) questions that one has. Therefore, researchers are advised to think in a more pragmatic² way rather than sticking to positivist or interpretivist epistemology (Denscombe, 2017). As criticised by White (2017), the positivism versus interpretivism division "is far from watertight" and "provides an over-simplified model" (p. 8). There are no 'good' or 'bad' approaches, but approaches that are more or less suitable or appropriate for achieving

² This refers to pragmatism in the philosophical tradition of William James and John Dewey (e.g. Dewey, 1957; James, 1975).

particular research purposes (Denscombe, 2017). Decisions regarding research design, methods of data collection, and analysis should be determined by research questions, i.e. these decisions should be ‘question-led’ rather than ‘methods-led’ (White, 2017). When designing research, researchers need to bear their research questions in mind, and think about what type of evidence is needed to answer these questions (de Vaus, 2001). Pragmatists make use of all possible approaches they deem appropriate to answer their research questions (Creswell, 2014; Mertens, 2010). That is, for a better understanding of research problems, quantitative and qualitative approaches to data collection may be combined to capitalise on the strengths and compensate for the weaknesses of either approach (Johnson & Onwuegbuzie, 2004). Authors such as Denscombe (2008), Gorard and Taylor (2004), and Johnson and Onwuegbuzie (2004) suggest moving beyond the quantitative versus qualitative research argument as both are important methods that can be combined to provide a more complete picture. However, this is not to diminish the value of ‘mono-method’ research nor to claim that every piece of research should combine quantitative and qualitative data. A justification is needed when choosing a particular methodological approach or when combining different approaches and the justification needs to be based on the kind of research questions posed (Creswell & Clark, 2011).

4.3 Types of Research Questions

Different types of research questions call for different research designs and subsequently different data collection methods, as different types of data are required to answer the questions. According to de Vaus (2001), research questions can be categorised into two fundamental types: descriptive and explanatory. Descriptive questions are concerned with ‘what is going on?’, while explanatory questions are concerned with ‘how or why is it going on?’. Answers to the latter kind of questions often involve causal explanations (Gorard & Taylor, 2004; de Vaus, 2001; White, 2017). De Vaus (2001) notes that explanatory questions build on good descriptions. In other words, research needs to aim for a proper description first to describe what is happening before venturing into attempts to explain how or why this is happening (de Vaus, 2001).

Research reported in this thesis focuses on variability in both trait and state WTC and its relationships with learning situations and language learning performance. Such research focus calls for a longitudinal design with high-density repeated measurements. Data collected across different occasions will be used to describe (and potentially explain)

fluctuations in WTC and to identify patterns of such fluctuations. By not being able to employ an experimental design, this research can only reveal co-variation rather than inferring causation. Hence, answers to explanatory questions provided by this research may be tentative. For the three research questions, the first is a descriptive question, aiming at identifying to what extent WTC varies. The other two questions are more explanatory, as they mainly aim at explaining why WTC might vary and whether WTC is a predictor of language learning performance. However, prior to this, it needs to be established whether WTC does vary and whether this variation is associated with variations in learning situations and performance. Therefore, two studies were carried out using different methods to answer the research questions step by step. In terms of data collection methods, study 1 used questionnaires to collect information to describe variations in WTC and determine the existence of relationships between WTC and related variables, followed by study 2, using semi-structured interviews to acquire rich and detailed descriptions from the participants, which form the basis for proposed explanations for how learning situations may influence a student's WTC and how a student's WTC may affect their language learning performance from their points of view.

4.4 Summary

This chapter has located the research within a philosophical position. The research is 'question-led', i.e. all methodological choices were determined by the research questions raised in section 2.7. As both descriptive and explanatory questions were asked, two studies were carried out using different methods to answer the research questions step by step. Details about the research design and methods of data collection and analysis will be presented in chapter 6 and 7.

5 Information about the Research Context

5.1 Communicative Language Teaching in China

Language teaching approaches that regard communication as a primary goal are generally referred to as communicative language teaching or CLT (Lee & VanPatten, 2003). CLT was first introduced to China in the late 1970s and challenged the effectiveness of the more teacher-centred traditional grammar-translation approach used in English classrooms. Since the 1990s, CLT has been approved by the Ministry of Education of the People's Republic of China (MOE) as the principal approach for teaching English (Yu, 2001). Curricula, textbooks, assessments and teacher training in both basic and higher education have been reformed to adopt the principles and practices advocated by CLT (Gil, 2016; Han & Yin, 2016).

English is an important subject at every stage of Chinese education. Chinese primary schools are required to start offering English classes from Grade 3 (age 8 to 9), aiming to develop young learners' interests and competence in using English to conduct day-to-day communication (MOE, 2001). For some economically more developed regions (e.g. Shanghai, Zhejiang, etc.), English language teaching may start as early as kindergarten or Grade 1 primary (Bolton & Graddol, 2012; Hu, 2005). In secondary schools, English, Chinese and mathematics are regarded as the three most important subjects (Bolton & Graddol, 2012; Song, 2000). During the past few decades, national policy and curricula have promoted more student-centred and communication-oriented pedagogy for English classes in secondary schools (Adamson & Morris, 1997; Hu, 2005). For example, the MOE suggests that the goal of English language teaching in senior high schools is to develop students' intercultural communication competence, particularly in terms of listening and speaking (MOE, 2003).

However, it seems that CLT has not been widely accepted and adopted by secondary school teachers, who tend to see English reading skills as more important. Some claim that only a few people may get the chance to communicate with foreigners, whereas most people read English in their daily life (Song, 2000). There is supporting evidence for this claim. For example, a national survey reported by Bolton and Graddol (2012) shows that only a few people use English often in their daily life, and that the clear majority seldom use English for purposes other than education. Moreover, Gaokao (i.e. the national college entrance examination in China), which determines access to higher education,

may have a washback effect on secondary school teachers' beliefs and pedagogy (Bolton & Graddol, 2012). The examination concentrates on reading rather than listening and speaking, indicating a need for accurate mastery of grammar rules and a wide vocabulary. Thus, holding class communication activities could be seen as a waste of time (Song, 2000). Consequently, students accustomed to teacher-centred language classes may feel uncomfortable when facing student-centred classes. This issue will be explored further in section 5.3.

At the higher education stage, a small number of students may choose English as their major, and all others, regardless of their chosen major, study compulsory College English. The differences between English education for English major and non-English major students in China will be outlined in the following section. A major goal of English teaching in Chinese higher education institutes, regardless of subject area, is to develop comprehensive competence, particularly listening and speaking skills, by using authentic English appropriate for a globalised society (Dai, 2008). Hence, English classes in colleges and universities may be more student-centred and communication-oriented, and thus challenging for recent secondary school graduates (Fang, 2010).

5.2 English Major Programmes and College English Courses

In China, university students are allocated departments and majors according to personal preferences and overall scores on the College Entrance Examination. It should be noted there is no additional English proficiency requirement for most English major programmes in China (see e.g. Liaoning Examination and Administration Service Centre, 2018 for details). Thus, English major students are not necessarily more proficient than their peers on other programmes at entrance. However, a study comparing different English as a foreign language (EFL) programmes in Japan found that students who studied on a programme with more L2 exposure were significantly more eager to communicate in English and have higher English proficiency, and these differences existed when students chose their programmes (Yashima & Zenuk-Nishide, 2008). This indicates that English major students may surpass their peers in motivation for and proficiency in English study upon starting their programmes, although this is not a university requirement.

All non-English major undergraduates in China must take College English courses. College English courses account for ten percent of total credits for a non-English major

undergraduate (MOE, 2007). College English courses mainly aim at preparing non-English major students to meet the growing needs of international communication, as English is seen as not just a language for interpersonal interaction, but also as a way of keeping in touch with the rest of the world. Non-English major undergraduates in China attend College English courses on a weekly basis for about three consecutive terms (one and a half academic years).

English major programmes differ from College English courses in many aspects, such as learning objectives, pedagogy, and degree of English exposure. The goal of English major education in China is developing English professionals who can use English proficiently for translation, teaching, management, and research purposes (ELT Advisory Board under MOE, 2000). Hence, English major programmes attempt to provide students with opportunities to use English daily. Almost all courses for English major students offered as part of the four-year undergraduate programme are delivered in English, focusing on listening, speaking, reading and writing skills, together with English-medium classes focusing on English culture, literature, and linguistics.

In the field of SLA, some research (e.g. Baker & MacIntyre, 2000; Llanes & Muñoz, 2013; MacIntyre et al., 2003; Yashima & Zenuk-Nishide, 2008) has compared students with immersion and other intensive learning experiences to non-immersion students. It is believed that immersion and other intensive learning experiences are likely to boost L2 learners' communication in the target language, because they offer more L2 exposure and interaction opportunities (Baker & MacIntyre, 2000). By comparing between French learners on intensive programmes, programmes containing different lengths of immersion experiences, and non-immersion programmes in a Canadian university, MacIntyre et al. (2003) found that immersion and other intensive learning experiences tended to encourage motivation, willingness to communicate and more frequent L2 communication. In a similar context, Baker and MacIntyre (2000) found that the differences between immersion and non-immersion students' L2 communication did not appear in L1 communication situations. Admittedly, the English learning context in China is considerably different from the French learning context in Canada; however, English major programmes in China might be seen as intensive language programmes likely to have advantages over College English courses with regard to English exposure and communication.

5.3 Chinese Students' WTC in English

Although tremendous efforts have been made by policy makers and language teachers in China to foster students' communication and communicative competence in English, CLT has failed to make the expected impact on Chinese students (Hu, 2002; Huang, 2007; Xiao, 2011). Chinese students are often seen as passive learners because of their apparent unwillingness to speak English and conduct interpersonal communication in either home country classroom settings or immersion situations in English-speaking countries (e.g. Liu, 2002; Liu & Littlewood, 1997; Tsai, 2017). To support this claim, evidence has been provided to show that Chinese students' levels of trait WTC in English are often below scale mid-point and tend to be lower than WTC levels in Mandarin. For example, by investigating a sample of 547 first-year non-English major undergraduates in a top university in Beijing, Liu and Jackson (2008) found that these students were slightly unwilling to communicate in English classes, although they were relatively willing to conduct interpersonal communication in Mandarin. Similarly, with a sample of 364 non-English major undergraduates in Taiwan, Chu (2008) found that the sample's mean trait WTC in English was below mid-point of the scale (2.46 out of 5) and was significantly lower than mean trait WTC in Mandarin (3.27 out of 5).

It is widely accepted that Chinese students' apparent silence can be 'explained' by traditional Chinese culture and communication norms, which are largely influenced by Confucianism (e.g. Dai, 2009; Hui, 2005; Li, 2012; Liu, 2002; Peng, 2007b; Wen & Clément, 2003). Chinese culture values reserved and implicit communication, and modesty (Liu, 2002; Peng, 2007b), so being talkative may be seen as being impolite and uneducated. Some researchers (e.g. Fu, Wang & Wang, 2012) claim that the fear of being regarded as 'showing off' in public is a major factor preventing Chinese students from actively communicating in both Mandarin and English. Some researchers (e.g. Peng, 2007b; Wen & Clément, 2003) note that the 'other-oriented' Chinese society consolidates Chinese students' unwillingness to communicate in English, particularly for those who are not highly proficient. This concern of being negatively evaluated can make Chinese students reluctant to take the risk of making mistakes and 'losing face' in front of others. Hence, keeping silent suggests a face-saving strategy for Chinese students to protect their positive social roles in English classrooms (Liu, 2002). Moreover, Chinese culture highlights collectivism, which brings about an 'insider effect' (Peng, 2007b; Wen & Clément, 2003). In the classroom context, this effect is manifest in a feeling of belonging

to a small group of friends, resulting in an unwillingness to interact with relatively unfamiliar classmates outside existing social groups. In a sociocultural context, the insider effect is a sense of belonging to a cultural community and an unwillingness to learn other languages or communicate with people outside that community.

In addition to the influence of Chinese culture and communication norms, Chinese learning traditions are regarded as obstacles to the effective implementation of CLT, as some important CLT content clashes with deeply rooted Chinese teaching and learning beliefs (Hu, 2002). Because of the extremely large Chinese population, schools always have large classes. To ensure a harmonious classroom environment for carrying out class activities effectively and efficiently, students are usually not allowed to talk freely in class, but rather required to be submissive and respectful. Under these circumstances, classes tend to be highly teacher-centred, and teachers always seen as an authority (Eng & Kumar, 2009; Hui, 2005; Peng, 2007b; Tien & Barnard, 2009; Wen & Clément, 2003). Keeping silent in classrooms, often interpreted in Western culture as a marker of negative learning attitudes and a lack of ideas or knowledge, shows politeness and respect to teachers in Chinese culture (Liu, 2002). Hence, Chinese students are used to passively receive knowledge from their teachers, instead of learning through actively questioning or interacting with their teachers and peers. Even when students can participate in classroom activities, they may be extremely cautious because they tend to believe that only correct answers and meaningful questions are welcomed (Hui, 2005; Peng, 2007b). Students seldom voluntarily participate in classroom activities when they are not highly confident about what they are going to say.

However, other researchers (e.g. Liu & Littlewood, 1997; Shi, 2006) argue that these interpretations are stereotypes of Chinese students. They criticise the tendency to use Confucianism as a convenient explanation for Chinese students' characteristics and behaviour. It is believed that the influence of Confucianism has been decreasing, so the stereotype of regarding teachers as the authority and banning students' talk in class may not apply to contemporary classrooms, particularly language classrooms (Shi, 2006). As no international survey has been conducted to compare Chinese students with their counterparts from other cultural backgrounds, this is no evidence showing that Chinese students are less willing to communicate. Rather, evidence has shown that Chinese learners prefer active learning styles and are eager to question their teachers and engage in some communication activities (e.g. Cheng, 2002; Liu & Littlewood, 1997; Shi, 2006).

For example, Liu and Littlewood's (1997) large-scale surveys conducted in Hong Kong showed that Chinese students were interested in participating in L2 communication activities, and preferred group discussions as L2 learning activities.

In addition, evidence has shown that the relatively unwillingness to communicate when using a L2 is not a unique feature of Chinese students but a common problem for L2 learners from different cultural backgrounds. For example, Lee (2009) conducted a 27-hour observation of L2 learners' participation in class discussions with a sample of six Korean postgraduate students in the US. The results showed that all participants were attentive listeners, although they rarely spoke during class discussions and none initiated conversation (Lee, 2009). In line with the above studies, Asmalı, Bilki and Duban (2015) compared 65 English major university students in Turkey to their counterparts in Romania. They found that, while their Romanian participants' average level of trait WTC was slightly above mid-point (6.52 out of 10), their Turkish participants' average level of trait WTC was considerably lower (3.55 out of 10).

Indeed, there are more similarities than differences between Chinese and Western classrooms, and Chinese students have many personal characteristics in common with their Western counterparts. Although we cannot deny the cultural impact on students' communication, it only partially explains this issue (Marlina, 2009). Seeing culture as the dominant factor in shaping students' classroom participation and communication overlooks individual differences. Individual differences in trait WTC within a cultural background may be more pronounced than differences across different cultural backgrounds. That is, not all Chinese students are passive learners unwilling to participate in communication activities and always silent in classrooms. If we find a student does not actively participate in a communication activity, it is not wise to blindly attribute it to Chinese communication norms and/or teacher-centred learning traditions, as it may not mean that the student is always unwilling to communicate, but merely indicates that the student temporarily decides not to enter into discourse in a specific situation. Willingness to communicate is not necessarily fixed but may vary across different activities with critical reflection on situations (Marlina, 2009). Hence, students' communication in classrooms can be less a result of cultural impact than one of the interaction between person and situation.

5.4 Summary

This chapter has introduced the Chinese EFL context where the research took place and identified an issue with Chinese students' communication and participation in English classrooms. It seems that schools and higher education institutes in China have widely adopted communicative language teaching to guide their English pedagogy; however, Chinese students are still stereotyped as passive learners who are unwilling to communicate in English classrooms both at home and abroad. However, no evidence has been found showing that Chinese students' L2 WTC levels are lower than those of L2 learners from other countries. Rather, both between- and within-person variations in WTC have been observed within and between countries. Hence, attention should be paid to individual differences and within-person fluctuations in WTC rather than looking for potential explanations in specific cultural backgrounds or ethnicity.

6 Study One: Variability and Stability of L2 WTC during a Semester

6.1 Introduction

The review of research on state WTC shows that within-person variability in state WTC has been observed in different L2 learning contexts, and a number of situation cues and characteristics have been suggested as situational antecedents. To address the first two research questions, this study mainly aims at examining whether such systematic within-person variability exists in an English classroom in a Chinese university. It also aims at testing whether it can predict English learning performance, to answer the third research question. As WTC displays both trait and state characteristics, this study also considers variability and stability in WTC at the trait level. That is, how much trait WTC differs from one individual to another (i.e. between-person variability in trait WTC), and whether these individual differences are associated with selected personality traits (i.e. systematicity of between-person variability in trait WTC).

As previously discussed, a high-density repeated measurement design might be the most appropriate approach to capture states in specific situations and monitor their fluctuations over time. However, only a handful of recent studies (e.g. MacIntyre & Legatto, 2011; Mystkowska-Wiertelak & Pawlak, 2014; Pawlak & Mystkowska-Wiertelak, 2015; Pawlak et al., 2016) have used it to investigate within-person variability in state WTC, and most of these tended to be short-term studies repeatedly measuring state WTC within very short periods of time (e.g. every 30 seconds within a few minutes). These studies were highly insightful, in that they provided individual state WTC trajectories over time, which could be evidence for the moment-to-moment variability in state WTC both across and within different tasks. However, they tended to measure state WTC several times while students were engaging in a communication task, which could arguably interfere students' performance, and scant attention has been paid to the relatively long-term within-person variability in state WTC. To fill this gap, this study was conducted to measure state WTC once in each lesson, to explore within-person variability across different lessons during a semester. Details about the methodology and findings of this study will be presented in this chapter.

6.2 Methodology

6.2.1 Overview

As a complement to previous research, this study was designed as an investigation of within-person variability in state WTC across different lessons during a four-month semester, making use of questionnaires to collect data. As well as reporting trait WTC and related personality traits through these baseline questionnaires, participants repeatedly reported their state WTC, together with related personality states and situation perceptions, during the semester through a momentary questionnaire. Their English exam scores were also collected through self-report questionnaires. Details about the data collection procedure will be presented in section 6.2.6.

6.2.2 Context

The study was conducted in the context of a College English course for first year undergraduate students in a key Chinese university, which is located in Beijing but enrolls students from all mainland Chinese provinces. With more than a hundred years of tradition, this university is among the top one hundred universities in China and specialises in science and engineering. All its non-English major undergraduates are required to take College English courses twice a week during the first three four-month semesters.

Based on their English exam scores on entrance, non-English major undergraduates in this university are usually allocated into two kinds of classes for College English courses: advanced for top students and regular for others. There are only two advanced classes for each grade, and all others (more than thirty classes) are regular classes. There are about fifty students in each class (class sizes of the advanced classes may be slightly smaller). Advanced classes are usually taught by more experienced teachers (e.g. associate professors and senior lecturers in the English Department), but curriculum, textbooks, and assessments for advanced and regular classes are the same.

The College English course under study consists of two types of lessons: reading and writing lessons, and oral lessons. The oral lessons for advanced classes take place in computer rooms, in which each student sits behind a computer screen interacting with the teacher and peers through headphones. This enables the teacher to talk with any individual student and monitor group discussions without walking around. For regular classes, all

lessons take place in traditional classrooms, where students sit in rows facing a blackboard and screen. In the traditional classroom, the teacher usually lectures from the front of the class and walks around during communication activities. The reading and writing lessons for both advanced and regular classes are relatively teacher-centred, emphasising grammar and vocabulary.

The national curriculum requires that College English assessments should be both formative and summative, assessing not only reading, writing, and translating skills, but also listening and speaking (MOE, 2007). In this university, the summative assessment is a paper-based exam at the end of each semester, assessing listening and reading comprehension, and composition writing. The exam papers for students in advanced and regular classes are the same. The formative assessment refers to course teachers' subjective judgment of their students' performance during a semester, mainly depending on class attendance and participation in communication activities.

This study focuses only on the oral lessons of the College English course, as they provide students with more opportunities to communicate in English. These lessons take place twice every other week. In each lesson, there are normally three major communication activities for students to participate in, such as presentations, group discussions, word games, etc. Each lesson lasts for a hundred minutes, with a five-minute break in the middle. Sometimes, teachers hold one or more communication activities during the first fifty minutes, leaving time for some listening exercises after the break; whilst in other classes, they began by commenting on student assignments, and all communication activities occur after the break.

6.2.3 Participants

Participants were selected based on convenience sampling (i.e. non-probability sample), as it was not possible to randomly select individuals. As a result, the sample needs to be seen as a specific context that may not be able to represent a wider population (see Cohen et al., 2011).

The sample was a group of 103 first-year undergraduates from two regular classes taught by the same teacher. The teacher was a Chinese female teaching fellow, who recently obtained her MA degree in ELT from a high-ranking Chinese university. At the time data collection started, the teacher had just started teaching the College English course for the

second semester. She was responsible for the students in four equivalent classes, and randomly selected two for this study.

As all participants were majoring in science (e.g. electrical engineering), there were far more males than females, with 85 males, 17 females, and one participant did not report his or her gender. Their age ranged from 17 to 21, with an average of 19 ($SD = 0.85$). Data reported by one participant was excluded from analysis, due to the consistency of her responses across all questionnaires³.

6.2.4 Ethical considerations

Ethics is an important issue when conducting empirical research, as any research may be an intrusion into participants' life. Researchers must consider the effects of research on participants, and be responsible to their participants (Cohen et al., 2011). The current research presented a low risk of harm to the participants, as it collected data relevant to L2 learning, without asking sensitive questions. However, participants invested time and effort to participate in completing questionnaires, which reduced the time they could have spent on English learning.

This study was reviewed and approved by the School of Education Ethics Sub-Committee at Durham University (15th December 2015). In advance to data collection, permission was also received from the head of the English Department at the target university.

To protect the privacy and confidentiality of participants throughout the research process, ethical guidelines issued by the British Educational Research Association (BERA) were followed. Prospective participants signed informed consent forms before they made decisions on participation. The consent form (see Appendix B) was written in English, with simple language that could be easily understood. It provided participants with the aims and procedures of the data collection, and explicitly stated that the participation in this research was completely voluntary and participants had the right to withdraw at any time. To ensure participants could understand the form before signing, the information was also presented to students in PowerPoint presentation in Mandarin (all participants' L1), and questions raised by participants were addressed in the introductory session.

³ This student responded 7 to all odd items and 1 to all even items across all measurement occasions.

Anonymity and confidentiality were protected throughout the research process. All questionnaires used were made anonymous. To guarantee anonymity and link responses together, confidential codes were created to replace participant names with a set of letters and numbers hiding their identities (details about the confidential codes will be introduced in section 6.2.6.1). The questionnaires were stored securely and password protected.

6.2.5 Research design

To describe individual students' momentary thoughts, feelings, and behaviours across a wide range of situations, which may form density distributions over time, the present study takes advantage of a high-density repeated measurement design. Thirteen measurement occasions were collected during the four-month period. However, this study may not be seen as a typical case of experience sampling methodology (see section 2.3), in that (a) the participants were measured only once a day during oral English classes rather than several times a day for several consecutive days; and (b) for fear of disturbing normal learning, they were asked to report their momentary thoughts, feelings, and behaviours in L2 communication activities immediately after finishing them, instead of reporting during communication.

6.2.6 Procedure

Data collection was carried out during the spring term of 2016. At the beginning of collection, the participants were just starting their second semester. In the first week of the semester, as part of an English lesson, I did a ten-minute presentation in Mandarin to both classes, introducing myself and the study to students. Students who agreed to participate signed the consent form, and completed a set of baseline questionnaires, including Big-Five personality traits, trait WTC in Mandarin, and demographic information. During the introductory section, the students indicated that they were going to have a meeting outside of class that could be used to complete the remaining baseline questionnaires⁴. Hence, questionnaires on trait WTC in English, and general perceptions of support received in the English classroom were completed several hours later when the English teacher was not present. Because the students had finished their first semester before participating in this study, their scores of the College English course for this

⁴ The university requires first-year undergraduate students to meet every weekday from 7 pm to 9 pm to study together. Each class has its fixed classroom. Students must come to this room, sign in and do their assignments. Students and their teachers can make use of this time to carry out academic and extracurricular activities.

semester were self-reported as baseline English scores.

From the following week onwards, the oral lessons of the College English course for both classes were observed twice every fortnight. The participants had oral lessons as usual without any intervention and I, as a non-participant observer, sat in a rear corner of the classroom, observing their communication behaviour. As the momentary questionnaire aimed to measure students' state WTC and situation perceptions in relation to a specific communication activity, it was better for it to be answered as soon as possible after the activity. To obtain immediate responses without interrupting normal learning, the questionnaire was distributed either during the break in the middle, or at the end of a lesson. In lessons when one or more communication activities had been held in the first fifty minutes, the momentary questionnaire was distributed during the break, and participants responded in relation to the communication activity they had just finished before the break. In lessons when all communication activities were held after the break, the momentary questionnaire was distributed at the end, and the students responded in relation to the last communication activity they engaged in. While participants were responding to the questionnaire, I was there to offer support for those who had concerns or questions. Participants were asked to reflect on their momentary thoughts, feelings, and behaviours during the specific activity they had just finished, and to complete the questionnaire as soon as possible based on their reflections. The participants were then asked to hand in their responses to me within five minutes.

For both classes, 13 oral lessons of the College English course were observed over 15 weeks. The momentary questionnaire was distributed 13 times during the semester; however, due to class attendance and other reasons, not all participants responded to it 13 times. Although individual total responses ranged from 1 to 13, almost all participants reported their state WTC more than once. About 45% of participants completed all 13 measurement occasions, and more than 80% responded ten times or more. Altogether, 1118 responses were received, corresponding to an average of 11 responses per person ($SD = 3.09$; response rate = 84%).

At the end of the semester, participants took the final exam of the College English course and then self-reported their end-of-term scores. To note, the score that a student could access was a weighted mean of the result of the paper-based exam and teacher subjective judgment of performance during the semester (see section 6.2.2). The raw scores on the

paper-based exam or teacher judgement were not available for this study (see the following section for details).

6.2.7 Materials

All materials sent to participants were paper-and-pencil questionnaires. These were based on validated scales commonly employed in the literature, although, adaptations were made to adjust these to this specific context. When a scale uses a set of items to measure the same construct, Cronbach's alpha can be used to show internal consistency. High internal consistency indicates that people respond similarly to items measuring the same construct. For example, students who have high levels of trait WTC should score highly on different items on the trait WTC scale. If responses to items on the same scale are greatly different, then the scale must have problems (Coolican, 2014). An alpha coefficient higher than 0.70 indicates relatively high internal consistency.

To ensure that participants could understand questionnaire items and respond to them properly, they were translated into Mandarin by myself, a native Mandarin speaker. The Mandarin version of the questionnaires was then back-translated into English by a professional translator, a native Mandarin speaker who spoke English as a second language.

Before data collection, feedback was sought from the English teacher of the participants, to ensure that the content of the items on the questionnaires was appropriate for this context. In addition, the questionnaires were piloted with a group of 30 non-English major undergraduate students from the same university. During piloting, focus was on the questionnaire, clarity and wording of the items as well as instructions and the layout. The questionnaire was adjusted based on the results of the pilot. One major change was that an item asking for students' matriculation numbers was taken out. By asking for matriculation numbers, I aimed to link individual responses to English scores reported by the teacher. However, some students in the pilot felt uncomfortable reporting their matriculation numbers, given that they had been asked to self-report their English scores. As a result, participant matriculation numbers were not collected. Hence, it was impossible to know raw scores on the paper-based exam or teacher judgment, and the self-report English score was the only way to indicate language exam performance. Another adjustment occurred in the instructions on creating confidential codes. The instructions asked students to write down the first two letters of their parents' first names;

however, results of the pilot study showed that students might confuse first names with last names⁵, which could result in a high rate of duplication between codes and/or inconsistency in codes reported by the same individual on different occasions. Hence, the phrase ‘first name’ was highlighted in bold, and examples were provided to clarify it.

6.2.7.1 *Baseline measures*

The baseline measures used in this study consisted of questionnaires on demographic information, Big-Five personality traits, trait WTC in Mandarin and in English, and general perceptions of support received in the English classroom (see Appendix C). Each scale or subscale consisted of a set of items measuring different facets of the construct. Other than items on demographic information such as major subject and age, all other items were statements (e.g. ‘I am willing to participate in group discussions.’), to which the participants were asked to provide a rating on a 7-point scale ranging from very inaccurate (scored as 1) to very accurate (scored as 7). Rating scales are widely used in research and are particularly useful for researchers who are interested in participants’ attitudes, perceptions and opinions (Cohen et al., 2011). By providing a range of responses for participants to select from, researchers can not only know participants’ attitudes towards each statement, but also be aware of the degree to which participants agree or disagree with each statement.

Demographics. Demographic information, including programme of study, year of study on the programme, age, and level of class, was collected as part of the baseline measure. Each participant was asked to create a confidential code based on given instructions (see Appendix C), write down the code on the questionnaire, and remember it. During the semester, the participants continued reporting these codes every time they responded to a momentary questionnaire, so reports from the same participant could be linked together without sacrificing the anonymity. Gender was indicated in the confidential code.

Personality traits. As part of the baseline measure, participants reported their Big-Five personality traits by completing a 50-item IPIP version of Big-Five Factor Inventory (see

⁵ This is an issue related to translation and Chinese culture. The first two letters of first names (rather than last names) are asked to reduce duplication. Many Chinese last names have the same first two letters, such as ‘LI’ or ‘ZH’, whereas the first two letters of Chinese first names can be various. However, the phrases ‘first name’ and ‘full name’ can be used interchangeably in Chinese without further explanation. Hence, some students might, on some occasions, write down the first two letters of full names instead, which are the first two letters of last names because last names come first in Chinese.

Goldberg, 1992; available at <http://ipip.ori.org/>). The IPIP scale is a ready-to-use measure of personality traits that assesses the five major dimensions of personality traits, i.e. extroversion, agreeableness, conscientiousness, emotional stability, and openness to experience, with ten items for each dimension. Participants were asked to honestly describe themselves as they were when they responded to the questionnaire. Cronbach's alpha of the items on extroversion, agreeableness, conscientiousness, and emotional stability were all above .70 (see Table 6.1), indicating high internal consistency in these scales. Cronbach's alpha of items on openness to experience in the current sample was slightly lower but still very close to .70.

Trait WTC in Mandarin. To indicate trait willingness to communicate in Mandarin, Cao and Philp's (2006) 25-item WTC questionnaire was used. This was originally adapted from McCroskey and Richmond's (1990) 20-item scale (the most frequently used scale for trait WTC mentioned in section 1.3) by adding five items specifically focusing on students' trait WTC in class activities. Participants were asked to report their intention to enter into discourse in situations where they had free choice. Communication situations considered not be applicable to this specific context were modified or deleted. For example, the situation of 'talking with a stranger on the bus' was replaced by 'talking with a stranger on campus', and the situation of 'talking with a garbage collector' was deleted. Three other communication situations that the participants in the current context might commonly come across were added to this questionnaire (e.g. 'talk with a fellow student when engaging in extracurricular activities'). Altogether, 26 items were sorted into four broad dimensions, with seven items concerning communication with strangers, seven items concerning communication with acquaintances, seven concerning communication with friends, and five concerning communication in class activities. The internal consistency of the trait WTC in Mandarin scale was high (alpha = .89).

Trait WTC in English. As the College English class is the most common, if not the only, situation for non-English major students in China to communicate in English, I decided to use MacIntyre et al.'s (2001) Willingness to Communicate in the Classroom scale rather than the more widely used McCroskey and Richmond's (1990) scale to measure students' trait WTC in English. However, MacIntyre et al.'s (2001) items were originally developed for FSL learners in Canada, a context significantly different from the one in this study. Additionally, some of the communication situations described in MacIntyre et al.'s (2001) questionnaire were somewhat out of date, such as 'reading letters from a pen

pal'. To adapt the questionnaire to the current study, items were amended considerably, by excluding items not suitable for participants (e.g. 'A stranger enters the room you are in, how willing would you be to have a conversation if he talked to you first?') and including alternatives. These new questions described communication situations that were more up-to-date (e.g. 'write a piece of status or a comment in English on SNS') and were more likely to occur in Chinese EFL learning contexts (e.g. 'do a role-play standing in front of the class in English'). A 36-item questionnaire was developed with items grouped into three dimensions, 15 items concerning speaking, 13 concerning writing, and eight concerning reading. The internal consistency of the 36-item scale of trait WTC in English was high ($\alpha = .96$).

Classroom support. Peng and Woodrow's (2010) items on teacher support, student cohesiveness and task orientation were used to measure perceptions of support received in the classroom. These three dimensions of classroom support correspond to situation characteristics (i.e. support, cooperation, and objectives) highlighted in the framework proposed. Items on teacher support and student cohesiveness measure the extent to which a student feels supported by the teacher and classmates. Peng and Woodrow's (2010) items on task orientation were used to measure student perceived quality and appropriateness (in terms of usefulness, interest, familiarity, difficulty, etc.) of classroom tasks. A high overall score on this scale indicates perception of the classroom as a place that facilitates language learning. As this study focused on thoughts, feelings, and behaviours in communication activities inside classrooms, items relevant to assignments in Peng and Woodrow's (2010) questionnaire (e.g. 'Class assignments are clear, so everyone knows what to do.') were modified (e.g. 'Instructions for activities are clear so everyone knows what to do.'). Based on the results of the systematic review, a few additional items relevant to classroom support were added. For example, in Peng and Woodrow's (2010) questionnaire, items relevant to classmate concern student cohesiveness (e.g. 'I am friendly to members of this class.' and 'I make friends among students in this class.'), emphasising relationships with other students. However, several studies (e.g. Cao & Philp, 2006; Kang, 2005; Pawlak & Mystkowska-Wiertelak, 2015; Riasati, 2012) show that, apart from relationships with classmates, their participation and cooperation also influenced perception of support. Therefore, items such as 'the classroom climate is active' and 'my classmates are supportive' were added, and the scope of this dimension was extended from student cohesiveness to any support that was offered by peers. Seven items were included to assess each of the three broad dimensions of

classroom support (i.e. teacher, peer and task). Reliability of the 7-item teacher support in the current sample was rather low ($\alpha = .59$); however, when the item ‘The teacher interrupts students to correct their errors when students are speaking or writing in English.’ was excluded from the scale, Cronbach’s alpha rose to .82. Thus, this item was not included in the analysis. The internal consistency of the items on task support, student support and the overall scale (after the above item was deleted) was high (see Table 6.1).

6.2.7.2 *Momentary measures*

The questionnaire that the participants responded to throughout the semester was composed of four parts, situational cues and characteristics, major dimensions of situation characteristics, state WTC and communication behaviour, and personality states and affect (see Appendix D). It directed the participants to reflect on experiences during the communication activity they had just finished and report their thoughts, feelings, and behaviours at that particular point in time. All the items were on a 7-point scale, from not at all (scored as 1) to extremely (scored as 7). Cronbach’s alphas for individual subscales are reported in Table 6.3.

Situation cues and characteristics. The first ten items were relevant to situation cues in the English classroom context, such as task, teacher, class, and respective partners. However, this questionnaire emphasises subjective perceptions of situation cues (i.e. situation characteristics), in terms of task-usefulness, task-interest, task-difficulty, teacher support and peer support. The situation cues and characteristics measured in this questionnaire were selected based on the framework proposed.

Dimensions of situation characteristics. This section assessed the major dimensions of situation characteristics that might be relevant to the current language learning situation. Rauthmann and Sherman’s (2016b) ultra-brief 8-item assessment of situation characteristics was used to assess the classroom situation on a more generic level than situation cues and characteristics in a specific communication activity. One item (‘Somebody is being deceived.’) was excluded from the questionnaire, as deception seemed unlikely to occur in English learning classrooms. The remaining seven dimensions were all measured. Nevertheless, focus was on the three dimensions suggested in the framework proposed: positivity, negativity and duty.

State WTC and communication behaviour. Two items were included to tap into

communication during the specific activity. One measured *intention* to communicate during the activity ('I was willing to communicate in English in the activity.'). i.e. state WTC, while the other measured *actual* communication behaviour ('I did communicate in English in the activity.'). Each response to the momentary questionnaire received a score representing state WTC and a score representing communication behaviour.

Self-assessed performance. One item measured self-assessed language learning performance during the specific communication activity under study (i.e. 'I performed well during the activity.').

Personality states. In this part, Big-Five personality was measured at the state level. The items were derived from Fleeson's (2001) adjective-based Big-Five scales. Due to the constraint of the number of items feasible per report in a high-density repeated measurement study, each of the five major personality dimensions was represented by three adjectives (Extroversion: talkative, energetic, assertive; Agreeableness: cooperative, trustful, warm; Conscientiousness: organised, hardworking, responsible; Neuroticism: insecure, optimistic, vulnerable; Openness to experience: intelligent, inquisitive, creative). Internal consistency of items on each subscale is presented in Table 6.3.

Positive and negative affect. Items on participant emotion were selected from Watson, Clark and Tellegen's (1988) Positive and Negative Affect Schedule (PANAS). Both positive and negative affect were measured with five adjectives (Positive affect: excited, inspired, proud, attentive, interested; Negative affect: afraid, upset, nervous, ashamed, irritable). Positive and negative affect adjectives were presented in alternating order. Please see Table 6.3 for Cronbach's alphas.

Task-confidence. The last item on this questionnaire measured task-related confidence in a specific communication activity, as the state facet of confidence has been suggested by MacIntyre et al. (1998) and others (e.g. Cao & Philp, 2006; Riasati, 2012) as an important situational antecedent affecting state WTC.

6.2.7.3 Language exam performance

Participants were asked to self-report their end-of-term scores of the College English course. These were on a hundred-point scale: a score below 60 was a fail, between 60 and 70 was a pass, between 70 and 80 was a merit, and above 80 was a distinction. Each score

was a combination of the end-of-term English exam score and teacher judgement. The end-of-term exam was a paper-based language exam emphasising grammar and vocabulary (i.e. the summative assessment), while the teacher judgment was more subjective focusing on class participation and performance during the semester (i.e. the formative assessment).

6.2.8 Data analysis

6.2.8.1 Analysis for research question 1

Data analysis began by summarising raw data on each scale and subscale. When necessary, items were reverse scored, so a higher score indicated a higher level on that variable. Afterwards, scores were averaged across all items belonging to the same scales. For example, everyone who completed the baseline questionnaires received a mean score across all the L2 trait WTC items, representing their central tendency to communicate in English (i.e. trait WTC in English). For the state data collected through the momentary questionnaire, scores were averaged across occasions. For example, scores for state extroversion were firstly averaged across different items (i.e. talkative, energetic, and assertive) within each occasion and then averaged across different occasions within individuals, so each person received an overall scale representing how extrovert the person was during the semester. Descriptive data of trait WTC in Mandarin, trait and state WTC in English, as well as communication behaviour in English were compared and analysed in detail.

State WTC data were then compared across different measurement occasions. Group aggregated state WTC trajectory was compared to individual trajectory, and standard deviations were used to quantify individuals' within-person variability in state WTC, and then compared to between-person variability in both trait and state WTC.

Another way to quantify between and within-person variability in state WTC is using an unconditional two-level hierarchical linear model, with state WTC being the dependent variable. In this model, Level 1 estimates state WTC at the individual level, i.e. the variation of an individual's state WTC around his or her mean, whilst Level 2 estimates state WTC at the aggregate level, i.e. the deviation of individuals' means around the grand mean. The equations of this model are as follows:

$$\text{Level 1: state WTC} = \pi_0 + e$$

Level 2: $\pi_0 = \beta_{00} + r_0$

Where π_0 is an individual's mean level of state WTC, e is the variation of the individual's state WTC around his or her mean, β_{00} is the grand mean of state WTC across individuals and occasions, and r_0 is the deviation of individuals' mean levels of state WTC around the grand mean.

In this model, everyone's state WTC is modelled as a result of their mean level of WTC (π_0) plus random variation (e), and each individual's mean level of WTC is modelled as a result of the grand mean of WTC (β_{00}) plus random variation (r_0). The value of e captures the amount of within-person variation in state WTC, the value of r_0 captures the amount of between-person variation in state WTC, and the sum of within- and between-person variation ($e + r_0$) is the amount of total variation in state WTC over the semester. The results of the unconditional HLM were compared to the variations calculated through the descriptive analysis (i.e. standard deviations), as they both aimed to quantify and compare between- and within-person variability in state WTC.

6.2.8.2 Analysis for research question 2

To investigate the systematicity of between- and within-person variability, Pearson's correlation coefficients were used to estimate the extent to which trait and state WTC correlated with selected personal and situational variables. Pearson's correlation is the most frequently used coefficient to analyse interval level data. The coefficient ranges from -1 (perfect negative correlation) to $+1$ (perfect positive correlation), indicating the extent to which a pair of variables co-vary. In this study, correlation coefficients were calculated both between and within individuals. Correlational analyses at the between-person level were conducted between (a) trait WTC and personality traits, (b) trait WTC and general perceptions on classroom support, (c) individual means of state WTC and situation characteristics, as well as (d) individual means of state WTC and personality states and emotions. These analyses aimed at explaining between-person variability (i.e. individual differences) in trait and state WTC.

Analyses at the within-person level aimed to investigate whether within-person variability in state WTC could be explained by changes in the learning situation. Correlation coefficients between state WTC and situation characteristics were calculated within individuals, indicating their contingencies. These were then compared and averaged

across individuals, to show individual differences in contingencies and a typical individual's contingencies. However, as some participants missed one or more measurement occasions (i.e. College English lessons during the semester), the number of responses varied. One could argue that, when investigating within-person variability, a larger number of responses collected per person would be better. In other words, data provided by individuals observed on all or most measurement occasions tend to be more reliable than data provided by individuals observed on only a few occasions. Hence, an average of all individuals' contingencies may not be the best estimation of the typical individual's contingency.

By comparison, HLM can take individuals' reliabilities into consideration when analysing data, which makes it a more appropriate approach to analyse contingencies (Raudenbush & Bryk, 2002). Several conditional two-level hierarchical linear models were carried out, with state WTC being the dependent variable and each situational antecedent being the independent variable, respectively. At Level 1 of each model, there was only one predictor, which was person-centred (i.e. calculating individuals' means across different occasions). Hence, Level 1 of each conditional model estimated the relationship between state WTC and each situational antecedent at the individual level. An example of the Level 1 equation is as follows:

$$\text{State WTC} = \pi_0 + \pi_1 (\text{task-importance}) + e$$

Where π_0 was an individual's mean level of state WTC, π_1 was an individual's slope for predicting state WTC by task-importance, and e was the variation of the individual's state WTC around the intercept.

A sample of the Level 2 equations is as follows:

$$\pi_0 = \beta_{00} + r_0$$

$$\pi_1 = \beta_{10} + r_1$$

Where β_{00} was the grand mean of state WTC across individuals and occasions, r_0 was the deviation of individuals' mean levels of state WTC from the grand mean, β_{10} was the grand mean of slope for predicting state WTC by task-importance, and r_1 was the deviation of individuals' slopes from the mean slope.

In multilevel modelling literature, a slope refers to the regression coefficient representing the relationship between the dependent and independent variables. In this model, state WTC is modelled as a function of task-importance, and π_1 represents the contingency between state WTC and task-importance for each individual. The value of β_{10} shows a typical individual's contingency between state WTC and task-importance, and the value of r_1 shows between-person differences in contingency between state WTC and task-importance. HLM can calculate individuals' contingencies (Level 1) and estimate the typical individual's contingencies and individual differences in contingencies (Level 2). However, coefficients provided by HLM are unstandardised. As a result, correlation coefficients, which can give estimations of effect sizes, were also calculated to answer the same question through two complementary approaches of data analysis.

6.2.8.3 Analysis for research question 3

To examine whether L2 learners' intention to communicate can predict their language learning performance, trait and state WTC's relationships with end-of-term English scores and self-assessed English class performance were estimated using correlation and regression coefficients. Analyses were carried out both between and within individuals.

To show how trait WTC predicted language learning performance, correlation coefficients between trait WTC and end-of-term English scores were calculated. As language learning performance might be affected by a set of interrelated variables (e.g. Big-Five personality traits), regression coefficients were also calculated. Multiple regression analysis allows modelling the relationship between each independent variable and the dependent variable after relationships between the independent variables are removed (Howitt & Cramer, 2014). Whether trait WTC predicted English learning performance when controlling for other trait variables, such as personality traits and baseline English scores, was tested.

To investigate whether state WTC could predict language learning performance, whether individuals' means of state WTC correlated with end-of-term English scores and/or individuals' means of self-assessed English class performance was tested. Regression analyses were also conducted to investigate whether state WTC predicted language performance when controlling for trait WTC. These analyses of relationships between state WTC and performance were at the between-person level.

As the focus of the third research question was on how within-person variability in state WTC predicts language learning performance, the within-person relationships between state WTC and performance were analysed. Correlation coefficients between state WTC and self-assessed English class performance were calculated within individuals to show whether the typical individual's self-assessed language learning performance co-varied with state WTC during the semester and whether there were individual differences in this co-variance. Additionally, a conditional two-level hierarchical linear model was carried out (see the previous section for details about conditional HLM), with self-assessed performance being the dependent variable and state WTC being the independent. This aimed at confirming the results of the within-person correlational analysis.

Another way to investigate how within-person variability in state WTC predicts language learning performance is analysing how situation-contingent WTC (i.e. the contingencies between state WTC and selected situation perceptions) correlates with end-of-term English scores. Investigating state WTC aims to improve language learners' state WTC and learning performance by systematically shaping classroom situations, which indicates that language learners are expected to respond to changes in learning situations to have better performance. Hence, I assume that if students could actively adjust state WTC with changes in learning situations, then they would probably perform well in language classes and exams. For example, individuals whose state WTC increases with an increase in perceived task-importance are more likely to perform well than individuals whose state WTC is relatively fixed or decreases with an increase in perceived task-importance.

The relationships between situation-contingent WTC and end-of-term English scores were analysed. As presented earlier, there are two approaches to estimating individuals' contingencies between state WTC and situation perceptions: (a) calculating correlation coefficients within individuals and (b) making use of the π_1 slopes provided in HLM outputs. Hence, relationships between situation-contingent WTC and end-of-term English scores were analysed in two ways. Firstly, individuals' correlation coefficients between state WTC and situation perceptions were used to represent situation-contingent WTC, and the correlations between these and end-of-term English scores were calculated. Secondly, individuals' π_1 slopes provided in HLM outputs were used as situation-contingent WTC, and their correlations with end-of-term English scores were estimated. Whether situation-contingent WTC predicted language learning performance when controlling for baseline English scores and trait WTC was also tested.

Finally, whether differences in language learning performance could be explained by the individual differences in state WTC and the individual differences in within-person variability in state WTC was investigated. Participants were divided into three groups according to their state WTC: Group 1 has a relatively variable and moderate level of state WTC, Group 2 has a relatively stable and high level of state WTC, and Group 3 has a relatively stable but low level of state WTC. The participants were sorted in descending order of within-person variability in state WTC. The first third of the sample (i.e. students whose state WTC varied most) were selected as Group 1. The remaining students, those whose state WTC were relatively stable compared to students in Group 1, were then divided into two groups according to individual mean of state WTC, i.e. the half with higher levels of state WTC being Group 2 and the half with lower levels of state WTC being Group 3. One-way ANOVA was conducted to investigate whether these three groups differed in their end-of-term English scores.

6.3 Results

The following section will present the key findings. Firstly, I will present descriptive statistics on trait and state WTC in English, compared to other relevant variables such as trait WTC in Mandarin and communication behaviour. To answer the first research question, I will quantify between- and within-person variability in trait and state WTC. I will then report trait and state WTC's relationships with selected personal and situational variables, particularly within-person relationships between state WTC and situational characteristics, to show the systematicity of between- and within-person variability in WTC. To answer the third research question, the relationships between WTC (including trait, state and situation-contingent WTC) and language learning performance (including end-of-term English scores and self-assessed English class performance) will be reported. Additionally, I will compare the exam scores of participants who reported different levels of state WTC and different amounts of variability in state WTC.

6.3.1 Descriptive statistics

6.3.1.1 *Trait WTC in English and in Mandarin*

As shown in Figure 6.1 (panel 1), the sample's average level of trait WTC in English (mean = 4.39) was slightly above the mid-point of the 7-point scale, indicating that participants were neither very willing nor unwilling to communicate in English on average. For comparison purposes, the distribution of trait WTC in Mandarin is also

presented in Figure 6.1 (panel 2). The average level of trait WTC in Mandarin (mean = 4.60) was slightly higher than the average level of trait WTC in English ($d = 0.23$). Moreover, participants differed in their trait WTC in English ($SD = 1.07$) more than they differed in trait WTC in Mandarin ($SD = 0.73$). Almost all participants' trait WTC in Mandarin was above 3 (min = 2.96) and below 6 (max = 6.04), indicating that they were relatively willing or just slightly unwilling to communicate in Mandarin; however, there were a few participants whose trait WTC in English was much lower than 3 (min = 1.50) or higher than 6 (max = 6.88), indicating that some participants might be extremely unwilling or willing to communicate in English. Despite the above differences, trait WTC in English positively and moderately correlated with trait WTC in Mandarin, $r(86) = .44$, $p < .01$ (two-tailed) suggesting that those who were relatively willing to conduct L1 communication were also relatively willing to conduct L2 communication.

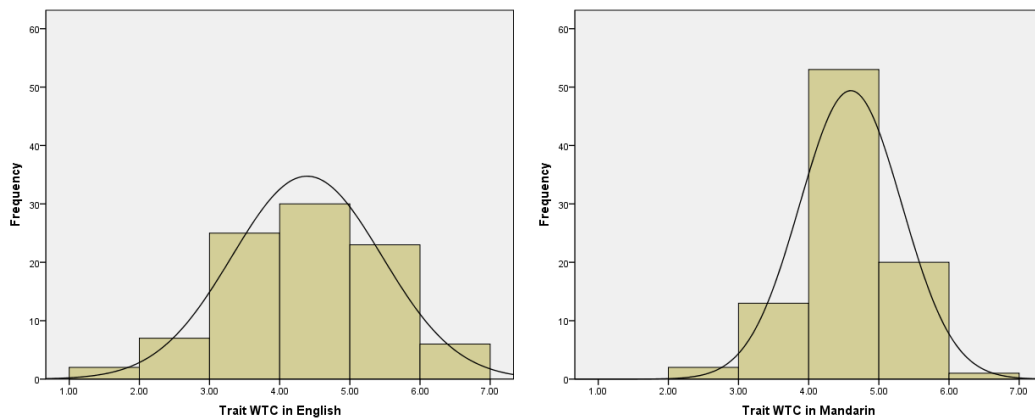


Figure 6.1. Distributions of non-English major students' trait WTC

To further explore the similarities and differences between trait WTC in English and Mandarin, correlation coefficients were calculated to investigate their relationships with Big-Five personality traits. As shown in the Table 6.1, trait WTC in Mandarin mainly correlated with extroversion ($r = .55$); however, extroversion did not significantly correlate with trait WTC in English ($r = .19$). Trait WTC in English was associated with agreeableness ($r = .21$), conscientiousness ($r = .23$), and openness to experience ($r = .30$). However, among these traits only agreeableness was also significantly associated with trait WTC in Mandarin ($r = .22$). Hence, it seems that L1 communication intention mainly depends on extroversion (and, to a lesser extent, on agreeableness); however, L2 trait WTC is a more complicated predisposition, which is related to a different set of personality traits other than extroversion or talkativeness.

Table 6.1. Descriptive statistics for trait variables in Study 1 ($N = 61-93$)

	Mean	<i>SD</i>	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.
1. Trait extroversion	3.88	0.95	(.80)											
2. Trait agreeableness	5.36	0.68	.18	(.71)										
3. Trait conscientiousness	4.79	0.92	-.02	.33**	(.80)									
4. Trait emotional stability	4.20	1.06	.14	.04	.17	(.84)								
5. Trait openness to experience	4.67	0.68	.31**	.22*	.27**	.01	(.69)							
6. Teacher support	5.91	0.80	-.04	.36**	.19	-.17	.05	(.82)						
7. Peer support	5.60	0.81	.16	.44**	.27**	-.06	.12	.51**	(.84)					
8. Task support	5.26	0.86	.10	.39**	.28**	-.07	.12	.64**	.51**	(.79)				
9. Classroom support	5.58	0.69	.09	.48**	.29**	-.11	.12	.85**	.80**	.86**	(.90)			
10. Trait WTC in Mandarin	4.60	0.73	.55**	.22*	.18	.05	.16	.20	.47**	.40**	.43**	(.89)		
11. Trait WTC in English	4.39	1.07	.19	.21*	.23*	-.17	.30**	.30**	.43**	.39**	.45**	.44**	(.96)	
12. Baseline English score	72.01	8.98	.07	.21*	.33**	-.15	.12	.13	.21	.05	.15	.18	.42**	
13. End-of-term English score	71.97	10.13	-.02	.21	.29*	-.15	.16	.23	.15	.12	.18	.13	.49**	.72**

Coefficients in brackets represent Cronbach's α for the respective scales.

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

6.3.1.2 State WTC and self-reported communication behaviour

As discussed in section 2.3, state data collected through repeated measurements could be analysed at two levels: the aggregate level and the individual level. That is, a mean could indicate either the typical individual's mean (i.e. the grand mean) or a specific individual's mean, and a standard deviation could be used to quantify either the deviation from the grand mean (between-person variability) or the deviation from individual mean (within-person variability).

To calculate everyone's mean state WTC during the semester, state WTC scores were averaged within each individual across all measurement occasions. The distribution of individuals' mean scores for state WTC is shown in Figure 6.2 (panel 1). Mean state WTC was significantly correlated with trait WTC, $r(91) = .53, p < .01$ (two-tailed). The grand mean was then calculated by averaging across all individuals' means. The grand mean of state WTC (mean = 4.95) was higher than the mean of trait WTC ($d = 0.50$).

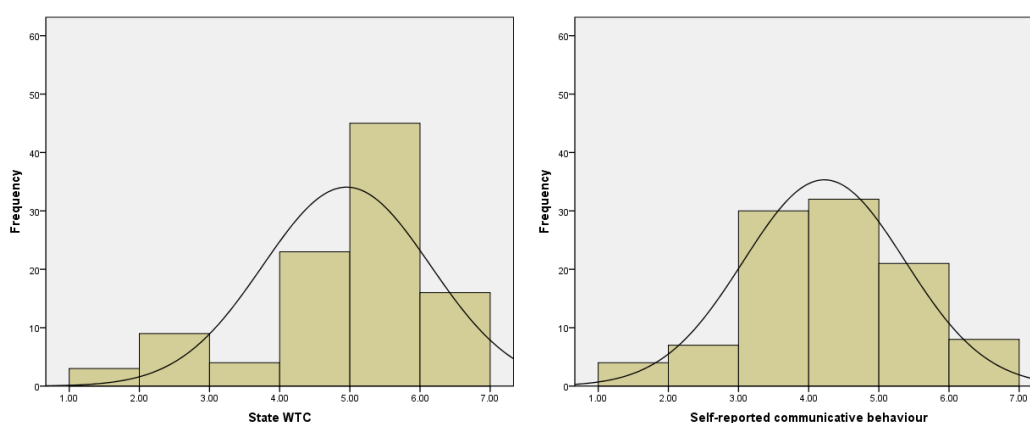


Figure 6.2. Distributions of non-English major students' state WTC (aggregate level)

To compare subjective communication intention (as reflected in state WTC: 'I was willing to communicate in English in the activity.') with actual communication behaviour, an item was included assessing self-reported communication behaviour in a specific class or activity (i.e. 'I did communicate in English in this activity.'). The distribution of individuals' mean scores for self-reported communication behaviour across the semester is shown in Figure 6.2 (panel 2). At the aggregate level, the grand mean of self-reported communication behaviour (mean = 4.22) was lower than the grand mean of state WTC ($d = 0.63$). At the individual level, self-reported communication behaviour positively and moderately correlated with both state WTC, $r(91) = .43, p < .01$ (two-tailed) and trait

WTC, $r(100) = .44, p < .01$ (two-tailed); however, for 85% of the participants, mean self-reported communication behaviour was lower than mean state WTC. This indicates that communication intention and behaviour could be related, but they are not the same construct.

6.3.2 Variability in trait and state WTC

To show how state WTC fluctuated at the aggregate level during the semester, state WTC on each measurement occasion was averaged across all individuals. As depicted in Figure 6.3 (panel 1), the aggregated trajectory fluctuates little, although showing a roughly decreasing trend. It indicates that state WTC did not significantly vary across different measurement occasions. However, when looking at each participant individually (panel 2), not a single individual's trajectory was the same as the aggregated trajectory (i.e. fluctuating around 5 on a scale from 1 to 7). Panel 2 shows that individuals' state WTC changed dramatically across different measurement occasions, and they differed from each other on each measurement occasion. If focus was only on aggregate level, then both between- and within-person variability could be overlooked.

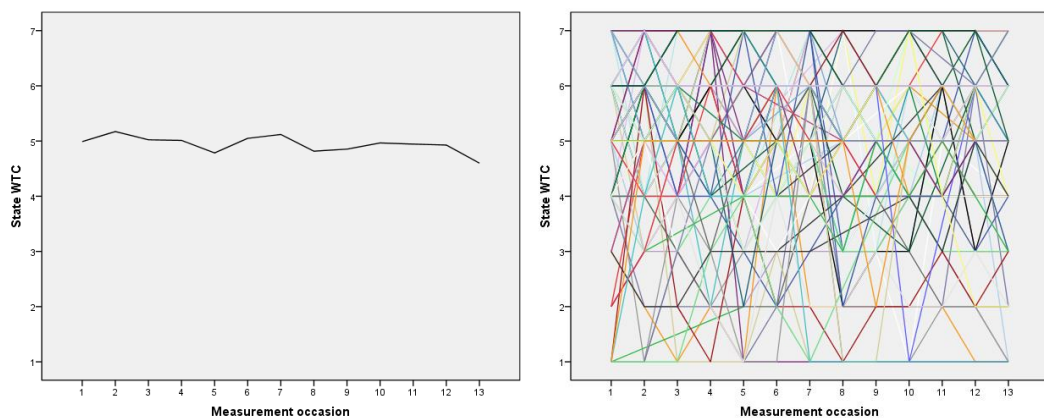


Figure 6.3. Aggregated and individual trajectories of state WTC over one semester

To quantify variability in trait and state WTC and to compare between different variability quantities, in the following sections I will follow Fleeson's approach of using standard deviations to estimate the amounts of between- and within-person variability (see Fleeson, 2001).

6.3.2.1 Within-person variability in state WTC

The *total* amount of variability (total variation in Figure 6.4) was quantified by calculating

the standard deviation of scores for state WTC across all momentary responses collected over the semester ($N = 1118$), regardless of whether the responses were from the same individual. By doing this, it is assumed that individuals overlapped completely in their distributions of state WTC over time, and there were few between-person differences. This figure indicates the maximum amount of within-person variability.

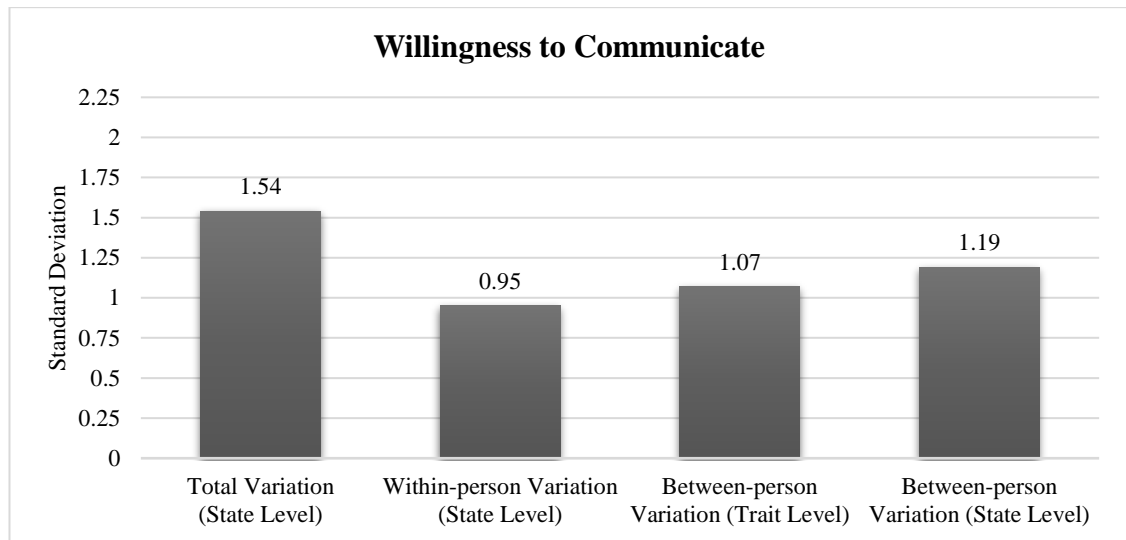


Figure 6.4. Variability in non-English major students' trait and state WTC

The amount of *within-person* variability in WTC was quantified by calculating the standard deviation of scores for state WTC across all measurement occasions for each individual separately, each deviation representing an individual's amount of within-person variability across the semester (the participants who had only one report were excluded from this analysis). To show the typical amount of within-person variability in WTC across the semester, a mean score was then calculated. As shown in the second bar of Figure 6.4, the average amount of within-person variability in state WTC accounts for two thirds of total variability. This indicates that individuals' density distributions of state WTC were wide, and there might be a large degree of overlap between individuals.

Figure 6.5 provides three density distributions of state WTC as examples. The distribution represented by the blue line is like the typical individual's distribution of state WTC during one semester, indicating relatively high state WTC and moderate within-person variability (mean = 5.00; $SD = 0.82$). Compared to the typical individual, the individual represented by the yellow line has a similar level of state WTC but a much higher level of within-person variability (mean = 5.23; $SD = 2.09$), whereas the individual represented by the green line has higher but less variable state WTC (mean = 6.00; $SD = 0.41$).

However, according to Figure 6.5, these three individuals' distributions overlap considerably: (a) like the one whose state WTC is less flexible (the green line), the individual who varies greatly (the yellow line) also tends to be very willing to communicate on most occasions; and (b) for the one whose state WTC is relatively fixed, state WTC may also vary on some occasions. In this circumstance, the means (i.e. central tendencies) are not adequate descriptions of individual distributions, and standard deviations (i.e. within-person variability) need to be considered.

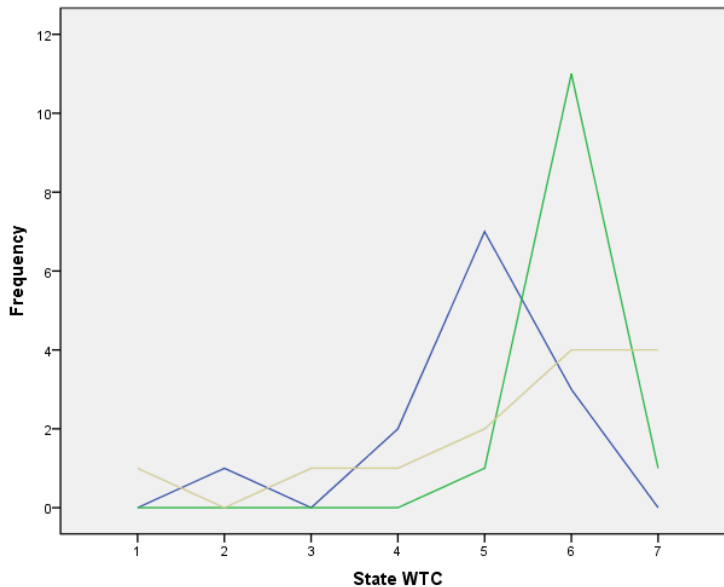


Figure 6.5. Three density distributions of state WTC

6.3.2.2 *Between-person variability in trait and state WTC*

The amount of variability between different individuals was quantified by two methods of analysis. Firstly, the amount of between-person variability was calculated at the *trait* level. The standard deviation of scores for trait WTC was calculated across individuals, to show how much individuals differed between each other in WTC at the trait level. Another way to show how much people differed in WTC was to quantify the amount of between-person variability at the *state* level. By averaging each participant's scores for state WTC across all measurement occasions, everyone received a mean score representing average level of state WTC across the semester. Then, the standard deviation of all individuals' mean scores for state WTC was calculated, representing the amount of between-person variability in WTC at the state level. The right two bars in Figure 6.5 show the amounts of between-person variability in WTC at trait and state levels,

respectively. The amounts of between-person variability in trait and state WTC were similar.

6.3.2.3 *Within- vs between-person variability comparison*

Results of the descriptive analysis show that the amount of within-person variability in state WTC was about the same as the amount of between-person variability in trait and state WTC. To further compare these amounts, they were estimated by using an unconditional model of HLM (see section 6.2.8.1). Results show that nearly half (46%) of the total amount of variability in state WTC occurred within individuals ($e = 1.12$)⁶. This is in line with the earlier findings using Fleeson's approach based on the calculation of standard deviations, suggesting that students' state WTC varied across different lessons during the semester, and the amount of this within-person variability was nearly as much as the amount of observed between-person variability in state WTC.

6.3.2.4 *Variability in self-reported communication behaviour*

For comparison purposes, amounts of between- and within-person variability in self-reported communication behaviour were estimated. Because the participants had not been asked to report their communication behaviour at the trait level on the baseline measurement, the amount of between-person variability in communication behaviour was only quantified at the state level, followed by the same procedure for state WTC. It was found that the amounts of within-person variability in communication behaviour were higher than the amounts of within-person variability observed in state WTC (see Figure 6.6). Standard deviations show that the average amount of within-person variability in communication behaviour was nearly 20% higher than the average amount of between-person variability. The result of HLM analysis was similar, in that nearly 70% of the total variability in communication behaviour was accounted by within-person variability ($e = 2.31$).

⁶ Level 1 model: State WTC = $\pi_0 + e$, where e refers to within-person variability in state WTC. Level 2 model: $\pi_0 = \beta_{00} + r_0$, where r_0 refers to between-person variability in state WTC. The total amount of variability in state WTC is $e + r_0$.

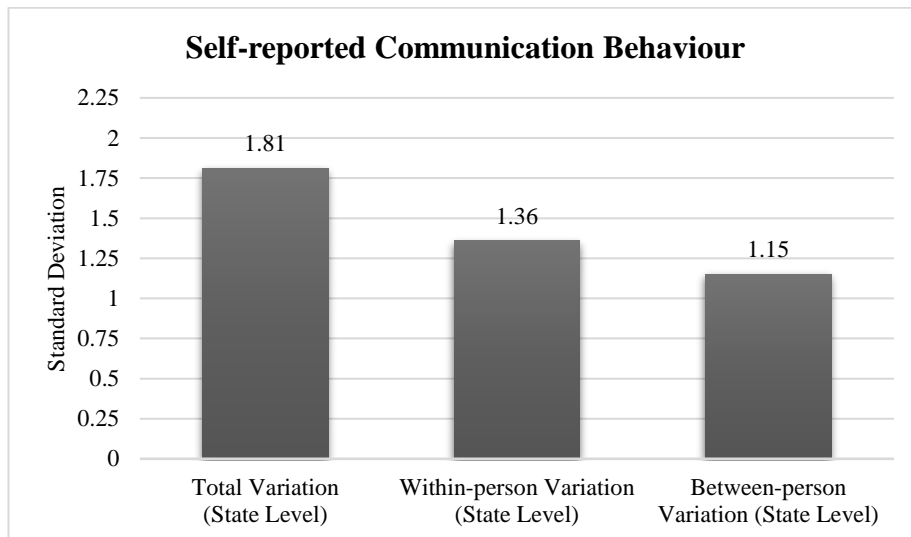


Figure 6.6. Variability in non-English major students' communication behaviour

6.3.3 Systematicity of variations in trait and state WTC

Whether trait and state WTC significantly correlated with some personal and situational variables was analysed. At the trait level, the focus was on the relationships between trait WTC and Big-Five personality traits, while at the state level, the focus was on the relationships between state WTC and the selected situation characteristics. To estimate the relationships between different pairs of variables, correlation coefficients were calculated. To note, no causal effect was inferred from the correlation coefficients (Coolican, 2014).

6.3.3.1 WTC and personality

The relationships between trait WTC and Big-Five personality traits have been presented in section 6.3.1.1 for making comparisons between trait WTC in English and in Mandarin. Trait WTC in English positively correlated with agreeableness ($r = .21$), conscientiousness ($r = .23$), and openness to experience ($r = .30$).

For comparison purposes, the relationships between WTC and personality were also analysed at the state level, i.e. whether mean state WTC was associated with mean personality states. Before calculating the correlation coefficients, I used HLM to estimate the amounts of variability in state feelings (including Big-Five personality states, and positive and negative affect). As shown in Table 6.2, both between- and within-person variability was found in each state, and within-person variability (e) was always

comparable in its amount to the respective between-person variability (r_0). That is to say, state feelings also fluctuated over the semester.

Table 6.2. Between- and within-person variability in personality states estimated by HLM

	e (within-person variability)	r_0 (between-person variability)
State extroversion	.76	.82
State agreeableness	.72	.75
State conscientiousness	.71	.78
State openness to experience	.68	.81
State emotional stability	.80	.74
Positive affect	.74	.78
Negative affect	.82	.77

After finding variability in personality states, their relationships with state WTC were tested. It was found that state WTC was significantly and positively related to the five major dimensions of state personality and positive affect (see Table 6.3). State WTC was significantly and negatively related to negative affect ($r = -.32$); however, the effect was much smaller than the effect between state WTC and positive affect ($r = .58$).

Table 6.3. Between-person correlations between state WTC and personality states ($N = 102$)

	Mean	SD	1.	2.	3.	4.	5.	6.	7.
1. State extroversion	4.82	.84	(.90)						
2. State agreeableness	5.05	.78	.90**	(.92)					
3. State conscientiousness	5.06	.80	.82**	.91**	(.93)				
4. State openness to experience	4.66	.83	.90**	.85**	.83**	(.84)			
5. State emotional stability	5.12	.80	.51**	.53**	.56**	.39**	(.78)		
6. Positive affect	4.64	.80	.84**	.85**	.86**	.86**	.47**	(.93)	
7. Negative affect	3.00	.84	-.23*	-.17	-.18	-.07	-.71**	-.15	(.90)
8. State WTC	4.95	1.19	.54**	.55**	.55**	.55**	.34**	.58**	-.32**

Coefficients in brackets represent Cronbach's α for the respective scales.

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

6.3.3.2 WTC and situation perceptions

At the trait level, the relationship between trait WTC and students' general perceptions of support received in their College English classrooms was investigated. It was found that trait WTC was positively related to general perceptions of classroom support ($r = .45$) and all its subscales (i.e. teacher support, peer support, and task support; see Table 6.1).

At the state level, the relationships between state WTC and the selected situation

characteristics (e.g. support, task-confidence, task-interest, task-importance, task-difficulty, and communication opportunity) as well as their major dimensions (e.g. duty, positivity, negativity, and sociality) were calculated. Everyone's scores for the above variables were averaged across all measurement occasions to calculate individual mean scores on each scale. The correlation coefficient between mean state WTC and each situational variable was tested across individuals, to ascertain whether between-person variability in state WTC was systematically related to perceived changes in the respective situation characteristic.

As shown in Table 6.4, state WTC significantly and positively correlated with both task-confidence ($r = .50$) and support ($r = .51$). When analysing support subscales, it was found that teacher support, classmate support, and partner support were all significantly associated with state WTC, although the impact of partner support ($r = .29$) was much lower compared to that of teacher support ($r = .53$) and classmate support ($r = .54$). Additionally, state perceptions of classroom support were related to trait perceptions of classroom support reported on the baseline measurement ($r = .51$). All other situation characteristics were related to the features of the tasks that participants were asked to complete in English lessons. As shown in Table 6.4, perceptions of task-importance ($r = .60$), task-interest ($r = .65$), and communication opportunity ($r = .50$) significantly and positively correlated with state WTC, whereas perceptions of task-difficulty did not. As for the underlying dimensions of situation characteristics, perceptions of duty, positivity and sociality were positively related to state WTC. Among them, positivity showed the strongest correlation with state WTC ($r = .64$), followed by a moderate correlation between sociality and state WTC ($r = .54$) and a weak correlation between duty and state WTC ($r = .24$). By contrast, perceptions of negativity were negatively related to state WTC ($r = -.23$).

The correlation coefficients presented above estimated the *between-person relationships* at both trait and state levels, indicating the systematicity of between-person variability in trait and state WTC. To investigate the systematicity of within-person variability in state WTC, the relationship between state WTC and each situational variable was calculated for each individual. These within-person situation-state relationships are referred to as 'situation contingencies' in the literature (Fleeson, 2007; see section 2.4.2).

Table 6.4. Between- and within-person correlations between state WTC and situation perceptions

	Mean	<i>SD</i>	1.	2.	3.	4.	5.	6.	7.	8..	9..	10.	11.	12.	13.	14.
1. Task-confidence	4.91	1.02														
2. Classmate support	4.96	.81	.56**													
3. Partner support	4.97	.75	.31**	.74**												
4. Teacher support	5.05	.94	.55**	.84**	.61**											
5. Overall support	4.99	.75	.53**	.95**	.85**	.92**										.25
6. Task-importance	4.84	1.01	.43**	.67**	.60**	.55**	.67**									.32
7. Task-difficulty	4.38	1.16	-.13	.26**	.37**	.08	.25*	.31**								
8. Task-interest	4.88	.97	.56**	.77**	.57**	.62**	.72**	.77**	.16							.30
9. Communication opportunity	4.86	.88	.48**	.75**	.64**	.58**	.72**	.62**	.20*	.67**						.23
10. Duty	4.56	1.14	.36**	.49**	.42**	.40**	.48**	.48**	.22*	.43**	.35**					.19
11. Positivity	5.07	.91	.47**	.72**	.54**	.62**	.69**	.73**	.12	.76**	.66**	.39**				.34
12. Negativity	3.15	1.07	-.21*	-.13	-.15	-.14	-.15	-.18	.30**	-.13	-.10	.06	-.22*			-.04
13. Sociality	4.99	.82	.47**	.72**	.52**	.56**	.66**	.60**	.19	.66**	.70**	.44**	.62**	.04		.32
14. State WTC	4.95	1.19	.50**	.54**	.29**	.53**	.51**	.60**	-.08	.65**	.50**	.24*	.64**	-.23*	.54**	

Note. Coefficients below the diagonal are between-person correlations ($N = 102$); coefficients above the diagonal are within-person correlations ($N = 1118$).

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

Firstly, contingencies were estimated by calculating correlation coefficients between state WTC and different situation perceptions separately for everyone. To note, participants who reported on less than 3 occasions, or whose scores for either variable were constant were excluded from these analyses. Task-difficulty was excluded for this analysis, as it did not show a significant relationship with state WTC in previous correlational analyses at the between-person level. On average, at the within-person level state WTC was moderately and positively related to all situational variables except negativity, although only 20% of the individuals' correlations were significant (see Table 6.4). Each coefficient above the diagonal of Table 6.4 is a mean of individuals' within-person correlations (i.e. contingencies), representing the typical within-person situation-WTC relationship.

However, individual correlation coefficients deviated from the mean, indicating differences in the within-person situation-WTC relationships. The standard deviations reported in Table 6.5 represent these differences but should not be confused with individual differences in trait WTC. Individual differences in trait WTC show that individuals differ in central communication tendencies; however, individual differences in within-person situation-WTC relationships show that individuals differ in how they respond to the properties of situations, which can be used to indicate individual difference in within-person variability (Fleeson, 2007).

To investigate the individual differences in within-person variability, the ranges of the within-person correlation coefficients that fell 68% of the sample (i.e. $\pm 1 SD$) were estimated. As shown in Table 6.5, the contingencies between state WTC and each situational antecedent (except for negativity) deviated from slightly negative to highly positive. This indicate that although a clear majority of individuals' state WTC was positively associated with the situational antecedents under study, some had higher correlation coefficients (i.e. stronger correlations) than others. That is to say, some individuals were more likely than others to adjust their state WTC when perceiving changes in the learning situation. Additionally, there were a small number of students whose state WTC was negatively correlated with one or more of these situational antecedents.

Table 6.5. Descriptive statistics for within-person situation-WTC correlations

	Mean	SD	Significance rate	Range of r
Support-WTC correlation	.25	.40	22%	$-.15 \leq r \leq .65$
Importance-WTC correlation	.31	.37	24%	$-.06 \leq r \leq .68$
Interest-WTC correlation	.30	.37	21%	$-.07 \leq r \leq .67$
Opportunity-WTC correlation	.23	.32	15%	$-.09 \leq r \leq .55$
Duty-WTC correlation	.19	.37	16%	$-.18 \leq r \leq .56$
Positivity-WTC correlation	.34	.38	27%	$-.04 \leq r \leq .72$
Negativity-WTC correlation	-.04	.43	16%	$-.47 \leq r \leq .39$
Sociality-WTC correlation	.32	.36	23%	$-.04 \leq r \leq .68$

However, it was interesting to find that state WTC was not associated with negativity at the within-person level ($r = -.04$). To note, the between-person correlation coefficient indicates that state WTC was significantly and negatively associated with negativity ($r = -.23$). This suggests that between- and within-person relationships between state WTC and negativity may differ. At the between-person level, learners with a more negative perception of a communication situation tended to be less willing to communicate in the target language (as compared to those with a less negative perception of the situation); however, at the within-person level, state WTC in different situations was not significantly influenced by negative perceptions of specific learning situations.

Secondly, within-person relationships between state WTC and different situational antecedents were further investigated through *conditional models of HLM*. The results of HLM were similar to those of the correlational tests presented earlier, as state WTC was positively associated with situational antecedents such as support ($\beta_{10} = 0.41, t = 6.03, p < .001$), task-importance ($\beta_{10} = 0.35, t = 6.69, p < .001$), task-interest ($\beta_{10} = 0.29, t = 6.17, p < .001$), and communication opportunity ($\beta_{10} = 0.19, t = 6.30, p < .001$). State WTC was also positively related to the underlying dimensions of situation characteristics, such as positivity ($\beta_{10} = 0.38, t = 8.23, p < .001$), sociality ($\beta_{10} = 0.31, t = 7.01, p < .001$), and duty ($\beta_{10} = 0.19, t = 4.40, p < .001$), but was not significantly related to negativity ($\beta_{10} = -0.06, t = -1.57, p = .12$).

The results of HLM analysis also showed significant individual differences in contingencies, although on average most of the situational antecedents positively correlated with state WTC at the within-person level. To further investigate these individual differences, the ranges of individual slopes (π_1) that fell 68% of the sample (i.e. $\pm 1 SD$) were estimated. A clear majority of participants positively adjusted their state WTC in response to positive changes in the selected situational perceptions (increases in

e.g. task-support); however, a minority negatively responded to these increases. For example, some individuals were slightly less willing to communicate when they felt more supported by the teacher or peers in the classroom.

Table 6.6. Descriptive statistics for situation-contingent WTC slopes estimated by HLM

	β_{10}	r_1	χ^2	df	Range of π_1
Support-contingent WTC	.41	.42	198.23	91	$-.01 \leq \pi_1 \leq .83$
Importance-contingent WTC	.35	.32	174.35	89	$.03 \leq \pi_1 \leq .67$
Interest-contingent WTC	.39	.30	174.61	88	$-.01 \leq \pi_1 \leq .59$
Opportunity-contingent WTC	.19	.08	107.55	89	$.11 \leq \pi_1 \leq .27$
Duty-contingent WTC	.19	.24	137.23	86	$-.05 \leq \pi_1 \leq .43$
Positivity-contingent WTC	.38	.27	171.91	87	$.11 \leq \pi_1 \leq .65$
Negativity-contingent WTC	-.06	.25	191.96	90	$-.31 \leq \pi_1 \leq .19$
Sociality-contingent WTC	.31	.27	179.48	89	$.04 \leq \pi_1 \leq .58$

Note. Level 1 model: State WTC = $\pi_0 + \pi_1$ (e.g. support) + e , where π_1 refers to each individual's slope. Level 2 model: $\pi_0 = \beta_{00} + r_0$ and $\pi_1 = \beta_{10} + r_1$, where β_{10} refers to the grand mean of slope, and r_1 refers to the deviation of individuals' slopes from the mean slope.

Both correlation coefficients and HLM outputs showed that some individuals responded more sensitively than others when perceiving changes in the situation, and some individuals responded inversely compared to the clear majority. To better explain individual differences in within-person variability, the contingency between state WTC and task-interest will be further illustrated as an example (see Figure 6.7). The typical individual's slope (β_{10}), everyone's slope (π_1), and three extreme cases of π_1 are depicted in the three panels of Figure 6.7, respectively. The typical slope (β_{10}) shows a clear positive relationship between state WTC and task-interest (panel 1); while individual slopes (π_1) show individual differences in the relationship between state WTC and task-interest (panel 2): although most participants have positive slopes, some have steeper slopes than others, and some have negative slopes. These individual differences are more evident when comparing the three extreme cases shown in panel 3. The participant represented by the yellow line has a positive slope, indicating that this person's state WTC tended to increase when perceiving a task as more important; while the participant represented by the blue line has a slope that approaches 0, indicating that this person's state WTC was not related to perceptions of task-importance. By contrast, the participant represented by the green line has a slightly negative slope, indicating that this person's state WTC tended to decrease when perceiving an increase in task-importance, which is rare, but did exist in this sample.

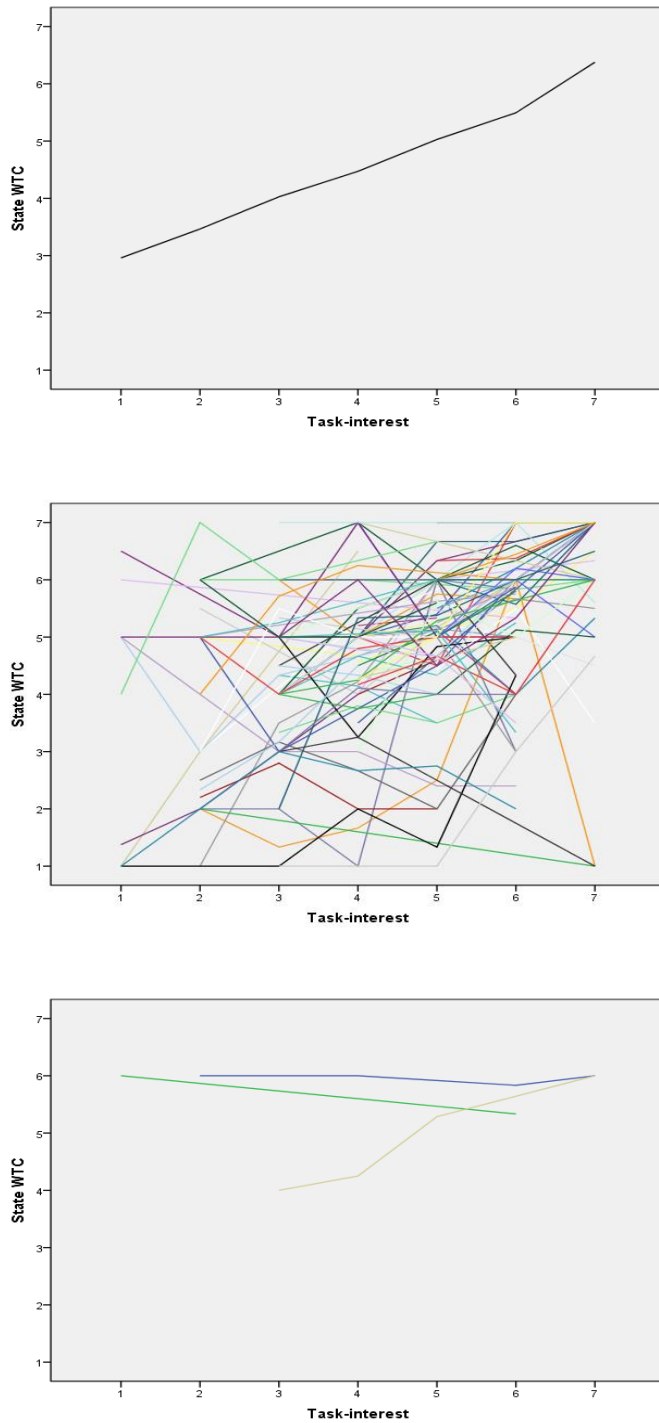


Figure 6.7. Aggregated and individual interest-contingent WTC slopes

6.3.4 WTC as a predictor of language learning performance

Another question addressed in this study was whether WTC could be a predictor of language learning performance. To collect data about language exam performance and, at the same time, protect the participants' anonymity, scores of the College English course were self-reported by the students. The baseline English scores were reported by 91

students (mean = 71.99; $SD = 8.93$), and the end-of-term English scores were reported by 70 students (mean = 71.97; $SD = 10.13$). Although participants differed more in end-of-term scores than in baseline scores, the mean of the end-of-term scores was very close to the mean of the baseline scores. The two sets of exam scores significantly correlated, $r(59) = .72, p < .01$ (two-tailed).

Other than reporting the relatively objective end-of-term exam scores, participants were also asked to self-assess their class performance on each measurement occasion. Hence, the relationships between WTC and language learning performance could be estimated both between and within individuals. As explained in the previous section, between-person relationships could be analysed at both trait and state levels: (a) the correlation coefficient between trait WTC and end-of-term scores, and (b) the correlation coefficient between mean state WTC and self-assessed performance over the semester.

6.3.4.1 Trait WTC and exam performance

A correlational test was conducted to investigate the relationship between end-of-term English scores and trait WTC in English, and a positive correlation found ($r = .49$). For comparison purposes, the relationship between end-of-term English scores and trait WTC in Mandarin was also tested; however, trait WTC in Mandarin did not correlate with end-of-term English scores. As suggested earlier, compared to L1 WTC (i.e. in Mandarin) that is mainly explained by extroversion or talkativeness, L2 WTC (i.e. in English) can be a more complicated decision-making process influenced by different personal and situational variables. The results of correlational tests indicate that it is intention to communicate in the target language, rather than talkativeness, that predicts language learning performance. Additionally, correlational tests were conducted between end-of-term English scores and Big-Five personality traits, as well as end-of-term English scores and general perceptions of classroom support. Among the five major dimensions of personality traits, only conscientiousness predicted end-of-term English scores ($r = .29$), while neither classroom support nor any subscale of it predicted end-of-term English scores (see Table 6.1).

As conscientiousness correlated with both trait WTC in English and end-of-term English scores, whether trait WTC in English predicted end-of-term English scores when controlling for conscientiousness was tested. Linear regression was employed using the 'enter' method, with end-of-term English score being the dependent variable, and

conscientiousness and trait WTC in English being the independent variables. It was found that, after controlling for conscientiousness, trait WTC in English still significantly predicted end-of-term English scores (see Table 6.7). That is to say, after the relationship between trait WTC and conscientiousness was removed, trait WTC remained as a significant predictor of language learning performance.

Table 6.7. Regression of performance on conscientiousness and trait WTC

	Unstandardised		Standardised	<i>t</i>	Sig.	Collinearity	
	Coefficients		Coefficients			Statistics	
	<i>B</i>	Std. Error	Beta			Tolerance	<i>VIF</i>
(Constant)	43.67	6.90		6.33	.00		
Trait conscientiousness	2.00	1.27	.18	1.57	.12	.95	1.06
Trait WTC in English	4.27	1.09	.45	3.91	.00	.95	1.06

a. Dependent Variable: End-of-term English score

6.3.4.2 *State WTC and self-assessed performance*

To test the between-person relationships between state WTC and language learning performance, everyone's scores for state WTC and self-assessed performance were averaged across all measurement occasions. Hence, everyone received a mean score representing their state WTC during the semester, and a mean score representing their self-assessed performance. It was found that state WTC correlated with both self-assessed performance, $r(100) = .57, p < .01$ (two-tailed), and end-of-term scores, $r(68) = .31, p < .01$ (two-tailed). However, after controlling for trait WTC, no significant relationship was found between state WTC and end-of-term English scores (see Table 6.8). This result suggests that state WTC in specific lessons and activities did not vary randomly, but largely relied on trait WTC that could significantly predict their language learning performance.

Table 6.8. Regression of performance on trait and state WTC

	Unstandardised		Standardised	<i>t</i>	Sig.	Collinearity	
	Coefficients		Coefficients			Statistics	
	<i>B</i>	Std. Error	Beta			Tolerance	<i>VIF</i>
(Constant)	50.06	5.49		9.12	.00		
Trait WTC	4.32	1.25	.46	3.46	.00	.72	1.38
State WTC	.59	1.12	.07	.53	.60	.72	1.38

a. Dependent Variable: End-of-term English score

For comparison purposes, correlation coefficients between self-reported communication behaviour and language learning performance were also tested at the between-person level. Although self-reported communication behaviour strongly correlated with self-assessed performance over the semester, $r(100) = .76, p < .01$ (two-tailed), it did not correlate with end-of-term scores. The findings show that state WTC can be a predictor of language learning performance, and indicate differences between state WTC (i.e. subjective intention to communicate) and actual communication behaviour.

In addition, the relationship between state WTC and language learning performance was analysed at the *within-person level*, i.e. whether within-person variability in state WTC and language learning performance over the semester were related. As language performance was not objectively assessed on different measurement occasions, self-assessed performance was the only way to indicate performance at the state level. Prior to correlational tests, it was necessary to show variability in self-assessed performance, especially within-person variability across different lessons across the semester. Following Fleeson's (2001) approach, amounts of between- and within-person variability in self-assessed performance were estimated. As shown in Figure 6.8, the amount of within-person variability was slightly higher than the amount of between-person variability, and these were comparable to the amounts of variability in state WTC and communication behaviour.

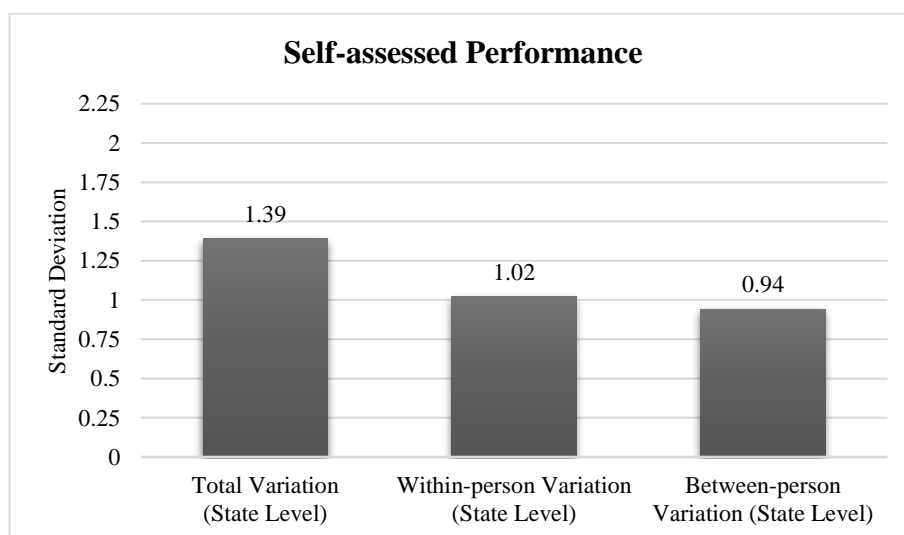


Figure 6.8. Variability in non-English major students' self-assessed performance

As a next step, the correlation coefficient between state WTC and self-assessed performance was calculated for each individual. It was found that, on average, state WTC

was positively related to self-assessed performance at the within-person level, although only around 15% of the correlations were significant. Nevertheless, the correlation coefficients ranged from -0.78 to 0.85 (mean = 0.22, $SD = 0.34$), indicating individual differences in these within-person relationships. Some individuals were more likely to adjust their self-assessed performance when their state WTC changed. Generally speaking, students tended to assess themselves performing better when their state WTC increased; however, there were a few individuals (less than 25%) who assessed themselves as performing worse when their state WTC increased.

Additionally, the within-person relationship between state WTC and self-assessed performance was further investigated through a conditional HLM model⁷. The results were similar to the results of correlation coefficients presented earlier, in that self-assessed performance was positively related to state WTC in general ($\beta_{10} = 0.29$, $t = 6.97$, $df = 92$, $p < .001$). The results of HLM analysis also showed significant individual differences in within-person relationships between self-reported performance and state WTC ($r_1 = 0.25$, $\chi^2 = 150.90$, $df = 86$, $p < .001$). As shown in Figure 6.9, individual WTC-performance slopes ranged from -0.88 to 0.88; however, for the vast majority of the sample (82%), within-person relationship between WTC and performance was not negative (i.e. $\beta_{10} \geq 0$).

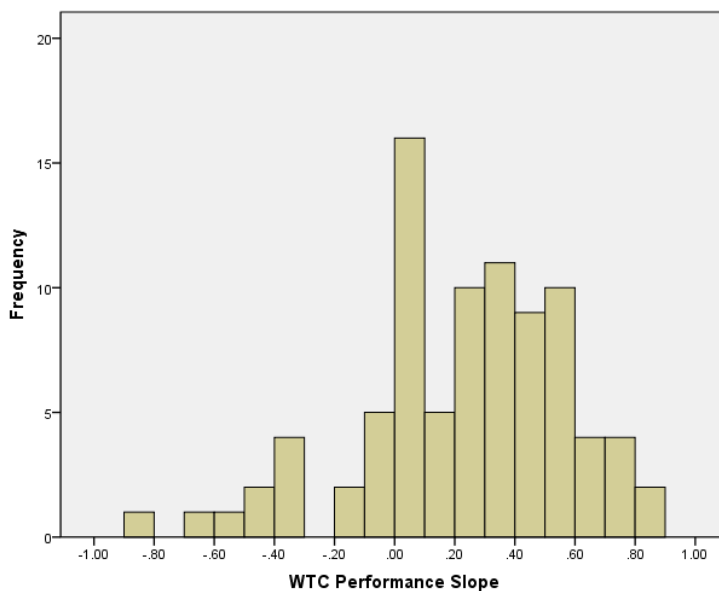


Figure 6.9. Distribution of individual WTC-performance slopes

⁷ Level 1 model: Performance = $\pi_0 + \pi_1$ (state WTC) + e . Level 2 model: $\pi_0 = \beta_{00} + r_0$ and $\pi_1 = \beta_{10} + r_1$, where β_{10} refers to the grand mean of slope for predicting performance by state WTC, and r_1 refers to the deviation of individuals' slopes from the mean slope.

6.3.4.3 Comparison between self-assessed and exam performance

It was interesting to find that self-assessed performance during the semester only slightly correlated with exam performance, as neither the correlation between self-assessed performance and baseline English scores, $r(88) = .22, p < .05$ (two-tailed), nor the correlation between self-assessed performance and end-of-term English scores, $r(68) = .22, p < .10$ (two-tailed), was strong. As baseline English scores significantly correlated with both trait WTC and end-of-term English scores, the effect of baseline English scores was removed using multiple regression analysis to test the relationship between self-assessed performance and end-of-term English scores. Results showed that, after controlling for baseline performance, self-assessed performance no longer correlated with end-of-term exam performance (see Table 6.9).

Table 6.9. Regression of end-of-term performance on baseline and self-assessed performance

	Unstandardised		Standardised	<i>t</i>	Sig.	Collinearity	
	Coefficients		Coefficients			Statistics	
	<i>B</i>	Std. Error	Beta			Tolerance	<i>VIF</i>
(Constant)	11.67	7.92		1.47	.15		
Baseline English score	.80	.11	.71	7.60	.00	.95	1.05
Self-assessed performance	.64	1.00	.06	.64	.53	.95	1.05

a. Dependent Variable: End-of-term English score

6.3.4.4 Situation-contingent WTC and exam performance

Another method to test whether systematic within-person variability in state WTC could predict language learning performance is assessing relationships between situation-contingent WTC (i.e. within-person relationships between state WTC and situation perceptions) and end-of-term English scores. As previously discussed, there are two ways of estimating contingencies (i.e. individual slopes) between state WTC and situation perceptions: calculating the correlation coefficients for everyone, and using the slopes provided in HLM outputs. Hence, the relationships between situation-contingent WTC and end-of-term English scores could also be analysed in two ways.

Correlation coefficients. Firstly, the contingencies representing individuals' correlation coefficients between state WTC and different situation perceptions, and the correlation coefficients between the contingencies and end-of-term English scores were calculated across individuals. As shown in Table 6.10, only interest-WTC contingency stood out, as

it positively and moderately predicted end-of-term English scores ($r = .32$). That is, students who showed higher state WTC when perceiving a task as more interesting tended to perform better in language exams and teacher judgments, and vice versa. Additionally, the support-WTC and positivity-WTC contingencies also seem to slightly predict end-of-term English scores; however, the correlations were relatively small.

Table 6.10. Correlations between performance and situation-WTC correlations ($N = 62-93$)

	1.	2.	3.	4.	5.	6.	7.	8.
1. Support-WTC correlation								
2. Importance-WTC correlation	.44**							
3. Interest-WTC correlation	.62**	.48**						
4. Opportunity-WTC correlation	.29**	.21*	.30**					
5. Duty-WTC correlation	.32**	.24**	.29**	.16				
6. Positivity-WTC correlation	.50**	.32**	.62**	.26*	.31**			
7. Negativity-WTC correlation	-.19	-.19	-.18	.00	-.01	-.18		
8. Sociality-WTC correlation	.42**	.37**	.53**	.18	.13	.52**	-.07	
9. End-of-term English score	.24	-.01	.32**	-.12	-.10	.22	-.05	.15

*. Correlation is significant at the 0.05 level (2-tailed).

** . Correlation is significant at the 0.01 level (2-tailed).

To clarify the relationship between the interest-WTC contingency and language learning performance, regression analysis was carried out with end-of-term English scores being the dependent variable and controlling for trait WTC in English and baseline English scores. As shown in Table 6.11, interest-WTC contingency still predicted end-of-term English scores when its relationships with trait WTC and baseline English scores were removed. In this model, both trait WTC ($\beta = .22$) and the interest-WTC contingency ($\beta = .21$) significantly predicted end-of-term English scores, although baseline English scores were much better predictors ($\beta = .61$). Altogether, 58% of the variability in end-of-term English scores was explained by baseline English scores, trait WTC, and the interest-WTC contingency. Regardless of individuals' trait WTC, those who had higher levels of interest-contingent WTC (i.e. an increase in state WTC in respond to an increase in task-interest) seemed to perform better in the end-of-term language exam. For example, among the three individuals depicted in Figure 6.7 (panel 3), the one represented by the yellow line (a positive slope) tended to perform best, whereas the one represented by the green line (a negative slope) tended to perform worst.

Table 6.11. Regression of performance on trait WTC and interest-WTC correlations

	Unstandardised		Standardised	<i>t</i>	Sig.	Collinearity	
	Coefficients		Coefficients			Statistics	
	<i>B</i>	Std. Error	Beta			Tolerance	<i>VIF</i>
(Constant)	11.91	6.85		1.74	.09		
Baseline English score	.69	.10	.61	6.63	.00	.82	1.22
Trait WTC in English	2.04	.87	.22	2.34	.02	.82	1.22
Interest-contingent WTC	5.69	2.33	.21	2.44	.02	.98	1.02

a. Dependent Variable: End-of-term English score

HLM coefficients. As individual slopes (π_1) provided in the outputs of conditional models of HLM were also indicators of contingencies between state WTC and its situational antecedents, their correlations with end-of-term English scores were calculated to confirm the above results. Similarly, only the interest-WTC contingency significantly and positively predicted end-of-term English scores ($r = .27$). In other words, individuals whose state WTC increased when perceiving a communication task as more interesting would probably perform better in the end-of-term language exam.

Table 6.12. Correlations between performance and situation-contingent WTC ($N = 63-69$)

	1.	2.	3.	4.	5.	6.	7.	8.
1. Support-contingent WTC								
2. Importance-contingent WTC	.39**							
3. Interest-contingent WTC	.58**	.46**						
4. Opportunity-contingent WTC	.33**	.20	.33**					
5. Duty-contingent WTC	.52**	.27*	.42**	.19				
6. Positivity-contingent WTC	.32**	.35**	.49**	.45**	.34**			
7. Negativity-contingent WTC	-.12	-.12	-.06	-.03	.00	-.08		
8. Sociality-contingent WTC	.47**	.41**	.49**	.32**	.23	.51**	-.03	
9. End-of-term English score	.11	-.01	.27*	-.10	.01	.15	.01	.07

*. Correlation is significant at the 0.05 level (2-tailed).

** . Correlation is significant at the 0.01 level (2-tailed).

Even when controlling for baseline English scores and trait WTC in English, the interest-WTC contingency still slightly predicted end-of-term English scores, although the effect was smaller and it seemed that the baseline English scores were much better predictors of end-of-term English scores (see Table 6.13). Altogether, 51% of the variability in end-of-term English scores was predicted by baseline English scores, trait WTC, and interest-WTC contingency. The results based on the systemic analysis conducted by HLM were in accordance with the correlation coefficient results presented earlier.

Table 6.13. Regression of performance on trait WTC and interest-contingent WTC

	Unstandardised		Standardised	<i>t</i>	Sig.	Collinearity	
	Coefficients		Coefficients			Statistics	
	<i>B</i>	Std. Error	Beta			Tolerance	<i>VIF</i>
(Constant)	11.52	8.09		1.42	.16		
Baseline English score	.62	.11	.54	5.68	.00	.94	1.06
Trait WTC in English	3.18	.93	.33	3.42	.00	.93	1.08
Interest-contingent WTC	3.34	1.76	.18	1.90	.06	.98	1.02

a. Dependent Variable: End-of-term English score

6.3.4.5 *Variability in state WTC and exam performance*

Lastly, the participants were divided into three groups based on individual mean of state WTC and amount of within-person variability in state WTC (see section 6.2.8.3). One-way ANOVA was conducted to investigate whether these three groups differed in their end-of-term English scores. Group 1's average level of state WTC was moderate (mean = 4.81), but its within-person variability in state WTC was relatively high (mean = 1.43). Group 2's average level of state WTC was the highest in the three groups (mean = 5.91), but its within-person variability in state WTC was relatively low (mean = 0.72). Group 3's average level of state WTC was the lowest (mean = 4.38), and its mean of within-person variability in state WTC was also relatively low (mean = 0.73).

The means of end-of-term English scores for three groups were 73.45, 75.57 and 66.29, respectively. These differed significantly, $F(2,59) = 5.19, p < .01$. Significant differences occurred between Group 1 and Group 3 and between Group 2 and Group 3, whereas Group 1 and Group 2 did not significantly differ. Figure 6.10 depicts three extreme examples. The yellow line represents a participant in Group 1, whose state WTC varied dramatically during the semester. The green line represents a participant in Group 2, whose state WTC was less variable and stayed at a relatively high level; while the blue line represents a participant in Group 3, whose state WTC was also relatively stable but stayed on a lower level. It seems that students in Group 1 and 2 (represented by the yellow and green line respectively) performed better in the end-of-term language exam than students in Group 3 (represented by the blue line). The results suggest that students whose state WTC was always high, as well as students whose state WTC was easily affected by the changes in the learning situation, tended to perform better than students whose state WTC was low and hard to be modified.

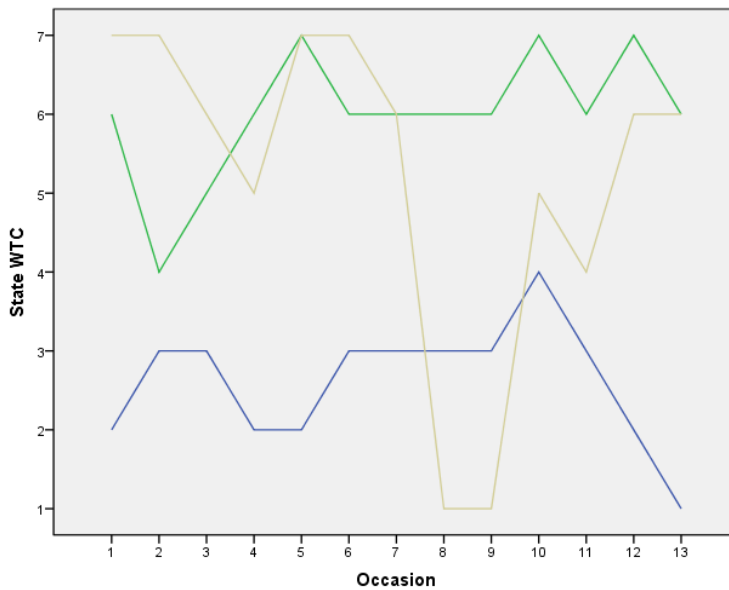


Figure 6.10. State WTC trajectories for three individuals

6.4 Discussion

The purpose of this study was to investigate whether there is between- and within-person variability in Chinese university students' trait and state WTC in English classrooms, whether this is systematically related to selected personal and situational antecedents (e.g. Big-Five personality traits and situation characteristics), and whether it could predict language learning performance. To achieve this, a high-density repeated measurement design was used. A set of questionnaires were used to measure a group of non-English major students' WTC and related variables at both trait and state levels. These trait and state data were analysed both between and within individuals.

The results indicate that there is not only between-person variability in both trait and state WTC, but also within-person variability in state WTC across different lessons during a semester. The amount of within-person variability in state WTC is comparable in size to the observed amounts of between-person variability in trait and state WTC. The results also show individual differences in within-person variability in state WTC, i.e. some individuals are more likely to change their state WTC than others, and a small number of individuals may change contrary to the clear majority. As a result, when studying variability in state WTC, more attention should be paid to the individual level: (a) identifying within-person variability in state WTC, and (b) comparing such variability across individuals.

This study is in line with previous research (e.g. MacIntyre & Legatto, 2011; Mystkowska-Wiertelak & Pawlak, 2014; Pawlak & Mystkowska-Wiertelak, 2015) which investigated within-person variability in state WTC on a moment-to-moment basis. While the current study was conducted, Mystkowska-Wiertelak (2016) published another study investigating the relatively long-term variability in state WTC over a semester. A group of 12 students' state WTC was measured every 5 minutes in a lesson for seven lessons in a semester to explore possible variability in state WTC both within each lesson and across different lessons. Although Mystkowska-Wiertelak (2016) found considerable variability in state WTC within specific lessons and activities, she did not find much variability across different lessons over the semester. However, Mystkowska-Wiertelak's (2016) finding does not contradict the finding of the current study, as she only analysed at the aggregate level without looking at individual state WTC trajectories.

The results of the current study suggest that between-person variability in trait WTC is systematically associated with some trait-like variables (e.g. Big-Five personality traits), while between- and within-person variability in state WTC is systematically associated with the interaction between personality and the situation, i.e. situation characteristics (e.g. support, task-interest, task-importance, etc.) and their underlying dimensions (e.g. duty, positivity, negativity, etc.). These findings are generally in line with previous research on trait (e.g. MacIntyre & Charos, 1996; Ortega, 2009) and state WTC (e.g. Dörnyei, 2009; Kang, 2005; Zhong, 2013). It indicates that within-person variability in state WTC across situations deserves further investigation rather than being simply ignored as error variance. However, this study does not find correlation between state WTC and task-difficulty, unlike the results reported by de Saint Léger and Storch (2009) and Eddy-U (2015).

As for predictability, both trait and state WTC correlated with language exam performance. State WTC also correlated with self-assessed performance at both between- and within-person levels. Generally speaking, students who are always willing to communicate will probably be the best language learners, and students who tend to modify their state WTC in response to changes in the learning situation seem to perform better than students who are always unwilling to communicate. It indicates that language learning performance is not only influenced by the average level of state WTC, but also the amount of within-person variability in state WTC. Compared to students whose state WTC is sometimes triggered and sometimes hindered by some situation cues and

characteristics, those whose state WTC is always high and can hardly be affected by the situation tend to perform better in language exams. However, it is worth noting that students who tended to be more willing to communicate when perceiving a task as more interesting seemed to perform particularly well.

A disadvantage of the questionnaire method might be that the participants' responses are confined by closed items and fixed options that had been predetermined (Cohen et al., 2011; Coolican, 2014). For example, it was impossible to know whether there were other important situational antecedents of state WTC that had not been included in the questionnaire, and why some individuals responded positively while others responded negatively towards the same items. Moreover, as state WTC and its situational antecedents were measured at the same time using the same questionnaire, no causal inference could be made. The correlation coefficients could only reflect relationships between variables, and these relationships could be subject to common method bias. Common method variance refers to 'the amount of spurious correlation between variables that is created by using the same method – often a survey – to measure each variable' (Craighead, Ketchen, Dunn & Hult, 2011, p. 578). One could argue that the participants might try to maintain consistency in their responses throughout the questionnaire, thus resulting in error co-variance besides the latent relationships (see Podsakoff, MacKenzie, Lee & Podsakoff, 2003). Hence, to acquire more informative details about how state WTC is influenced by the learning situation, further research should employ other data collection methods such as interviews.

Another limitation of this study was that all the data related to language learning performance were self-reported by the students. As explained in section 6.2.7, student matriculation numbers were not collected. Hence, the only way to link responses to exam performance was through self-reporting; however, this was not ideal, as it was impossible to know whether the exam scores were accurately reported. For example, one might not be able to remember the exact score, and thus reported a rough score instead. Another risk could be that some students, particularly those who did not perform well in the exam, might report dishonestly. However, as the report was anonymous, and no benefit was provided for any individual, there was no need to pretend to have higher scores. Hence, it is believed that most of the students reported their exam scores honestly and accurately. For future research, it would better to access official records or use teacher reports to collect exam scores.

At the state level, language learning performance was only self-assessed by the students, and this may not be a valid description of their actual performance. The findings suggest that, to some extent, the students assessed their current performance according to performance in previous exams; however, the relationship between their self-assessed class performance and performance in end-of-term exam is rather weak. Like the current study, de Saint Leger and Storch (2009) also reported a weak correlation between self-assessed class performance and teacher-reported grades. Hence, the next step may be exploring how and why student self-assessments of language performance differ from more objective assessments such as exam scores and teacher reports.

Another issue identified in the current study was the inconsistency between state WTC (i.e. subjective intention to communicate) and actual communication behaviour as reported by the participating students. Although the two groups of data significantly correlated, levels of actual communication behaviour seemed to be lower than levels of state WTC. This indicates that, in some situations, intention to communicate could be transformed into actual behaviour (e.g. when there are opportunities to communicate); however, in other situations, communication intention could not be put into practice. Additionally, this study also found that both trait and state WTC predicted language exam performance, but self-reported communication behaviour did not. As discussed in section 1.2, the differences between subjective communication intention and actual communication behaviour have not received enough attention. Hence, it might be worth further investigating the differences between WTC (particularly state WTC) and communication behaviour in future studies.

6.5 Summary

In this chapter, the methodology and results of Study 1 have been discussed. This study investigated a group of Chinese university students' trait WTC in English classrooms and explored their state WTC during a semester to understand relatively long-term variability within individuals. Both trait and state data were collected through questionnaires and statistically analysed. Systematic within-person variability (not entirely error variance) in state WTC has been found, which may be explained by the changes in learning situations (e.g. task-interest). This study has also found that students who have higher levels of WTC (both trait and state) tend to perform better in language exams, even when controlling for baseline language scores. Although students who have stable and high state WTC (i.e.

WTC that was not easily affected by learning situations) appear to perform best, students whose state WTC is easily modified by changes in learning situations (e.g. when tasks became more interesting) tend to perform better than students whose state WTC is less modifiable and low. Additionally, this study has observed differences between state WTC (i.e. subjective intention to communicate) and actual communication behaviour, as well as between self-assessed performance and actual language exam performance.

However, this study has not thoroughly explained the psychological processes of communication generation that occur within individuals, e.g. how personality interacts with the learning situation jointly influencing state WTC, and why communication behaviour could be generated or hindered in specific situations. Such communication generation processes will be explored in the next study, using interviews to acquire in-depth explanations of individual communication generation processes and how these psychological processes may change across different situations. The relationships and differences between state WTC and communication behaviour will be further explored. In addition, criteria for students' self-assessed performance and more objectively assessed performance (e.g. exam scores and teacher assessments) will be compared.

7 Study Two: Variability and Stability of L2 WTC during a Lesson

7.1 Introduction

Study 1 focused on relatively long-term within-person variability in Chinese university students' state WTC in English classrooms, i.e. variability across different lessons during a semester, which seems to be systematically associated with selected situation characteristics (e.g. support, task-interest, and task-usefulness) and their underlying dimensions (e.g. duty, positivity, and negativity). To have a more comprehensive understanding of Chinese university students' communication in English classrooms, it may be insightful to further explore the psychological processes of communication generation in different learning situations. Hence, a shorter-term study (Study 2) was carried out to (a) investigate relatively short-term within-person variability in Chinese university students' state WTC in English classrooms, i.e. variability across different communication activities during a lesson; and (b) provide in-depth explanations of why and how state WTC and communication behaviour are generated in some situations rather than the others.

Although a handful of studies have been carried out to investigate short-term variability in state WTC, they tended to collect data for a few minutes in labs (see section 3.3). While Pawlak and Mystkowska-Wiertelak (2015) focused on discussion in dyads, MacIntyre and Legatto (2011) focused on individual monologues and did not expect their participants to converse with the researcher. As both focused on one form of communication (i.e. dialogues or monologues), variability in state WTC across different types of communication activities is neglected. Mystkowska-Wiertelak and Pawlak (2014) used a similar approach to compare monologues and dialogues, and found a significant difference in state WTC, suggesting that state WTC tends to decrease during monologues but increase during dialogues; however, both tasks conducted by Mystkowska-Wiertelak and Pawlak (2014) were in dyads with partners assigned by the researcher.

It is important to note that authentic learning situations in real classrooms might be more flexible than in labs. In real classrooms, students are usually provided with various communication activities during a lesson, and may interact with different people in each activity. Hence, different situational variables (e.g. interlocutor, type of activity, topic, etc.) may interact and each variable's impact on state WTC can be either strengthened or weakened by the others, which may result in within-person fluctuations in state WTC

different from those found in labs. To investigate short-term within-person variability in state WTC in authentic classroom situations, Study 2 was carried out during normal lessons in real classrooms, serving as a complement to previous studies. Details of how Study 2 was conducted, and its findings will be presented in this chapter.

7.2 Methodology

7.2.1 Overview

Unlike Study 1 which was based solely on questionnaires, this study also used classroom observation and group interviews to obtain more in-depth explanations of the psychological processes of communication generation in different learning situations. The observation aimed at identifying important situations or issues for discussion in the follow-up interviews rather than collecting data for analysis. During the three weeks of data collection, participants' communication behaviour and performance in English classes were observed in groups. Each group was observed once. Altogether, eight lessons were observed. After each lesson, observed students responded to a brief questionnaire measuring state WTC in each communication activity and were interviewed to reflect their thoughts and feelings in different learning situations. The interviews were semi-structured to provide students with freedom to express what they really thought and how they interpreted their experience (Richards & Morse, 2007). An interview schedule was outlined in advance based on the proposed framework and results of Study 1, as they provided sufficient prior knowledge about within-person variability in state WTC and its situational antecedents (see section 7.2.7.2 for details). To stimulate recall during the group interviews, all the lessons were recorded to capture what was happening at different points in time. To triangulate student responses, this study also collected course teachers' observations and interpretations of their respective students' communication behaviour and language performance in each communication activity. This aimed at studying the same behaviour from different perspectives to increase the possibility of fully understanding the complexity of the behaviour (Cohen et al., 2011).

7.2.2 Context

It would have been preferred to carry out this study in the exact same context as study 1 (i.e., with the same cohort of students, or with students who major in the same subjects). However, unfortunately, this was not possible. Therefore, this study was carried out in the same university where Study 1 was done, but in the context of an English course for first

year undergraduates who learned English as their major subject. The English Department at this university not only takes charge of College English teaching for non-English major students, but also runs an English major programme with two pathways, BA English (English language and literature), and BA English (Energy English). Each year, the university enrolls about forty English-major undergraduate students, a clear majority of which tend to be female. They are randomly allocated into two classes, and required to live together in dorms with six people in each room. In the first two years, the classes take the same courses for 15 hours per week. In some modules (e.g. Listening Practice, Extensive Reading, Academic Writing, Translation, etc.), the two classes are taught together, while in other modules that require more oral interaction (e.g. Conversation Practice, Intensive Reading, Interpretation, etc.), they are taught separately by different teachers. At the beginning of the third academic year, students choose pathways for their speciality.

The course of interest is Intensive Reading (also known as Basic English in some other universities), which is a core course for BA English programmes throughout China. Although the course is named Intensive Reading, it is better understood as a comprehensive language course as the course that focuses on reading skills is named Extensive Reading. According to the national curriculum for English major undergraduates, the main objective of Intensive Reading is to improve early year English-major students' overall proficiency, including listening, speaking, reading, writing and translating skills (ELT Advisory Board under MOE, 2000). The curriculum requires course teachers to encourage students to participate in all kinds of communication activities (e.g. presentations, debates, book reviews, role-plays and interviews) to enable them to communicate in English, both orally and in writing. Although the curriculum does not provide specific requirements for course assessments, it suggests a balance between multiple choice and open-ended questions in end-of-term exams.

All BA English undergraduates in the university under study are required to attend Intensive Reading lessons twice a week during the first two years before choosing their pathways. Each lesson lasts for ninety minutes, with a five-minute break in the middle. In Intensive Reading lessons, the two classes of interest here are taught separately by two teachers. During these lessons, both teachers and students are supposed to use English only. In some lessons, students discuss the articles in the textbook, together with questions and exercises; in others, students use additional learning materials selected by their

respective teachers, such as English novels, poems, and newspapers. Communication activities often held in Intensive Reading lessons include group discussions, debates, presentations, news reports, speech imitations⁸, etc. The course assessment consists of two parts, a formative and a summative. The formative assessment is teacher judgment of class participation and performance during the semester, while the summative assessment is a paper-based language exam.

7.2.3 Participants

This study used a convenience sample of 39 first-year English-major undergraduates in two equivalent classes. Only four were male, two in each class. Seven of the students (four from Class 1 and three from Class 2) were recruited into the pilot study, and all others (16 from Class 1 and 15 from Class 2) participated in the main study. Data collected in the pilot study were excluded for analysis. One student withdrew before attending the group interview. Altogether, 31 students (28 females and 3 males) took part. All these were native mandarin speakers learning English as a L2. Their age ranged from 16 to 19, with an average of 18 ($SD = 0.67$). When data collection started, they had been attending Intensive Reading lessons for about three months. As it was the first semester, no previous English score could be reported as baseline information.

Both Intensive Reading teachers were Chinese female lecturers speaking English as a L2. They had MA degrees in relevant subjects (e.g. English Language and Literature) from top universities in China within the past five years and had been teaching Intensive Reading since obtaining their degrees. One teacher randomly allocated her students into five groups at the beginning of the semester and required them to work in the fixed groups throughout, while in the other class, students were free to choose where to sit in each lesson and collaborated with their neighbours when communicating in dyads or groups.

7.2.4 Ethical considerations

This study was reviewed and approved by the School of Education Ethics Sub-Committee at Durham University on 19th December 2016. The study was also approved by the head of the English Department at the university under study. As noted in Study 1 (see section

⁸ In this type of activity, students listen to recordings of speeches made by native speakers and try to imitate their pronunciation and intonation. After each student's imitation practice, the teacher and classmates will provide comments and suggestions on the student's pronunciation and intonation. This activity aims to make students aware of the problems in their pronunciation and speak English as native speakers do.

6.2.4), participation was completely voluntary, and anonymity and confidentiality were protected. A consent form like the one used in Study 1 (see Appendix E) was signed by the participants, to ensure their participation was voluntary and to protect their right to withdraw. All questionnaires used were anonymous, and confidential codes were used to link individual responses without revealing their names. The questionnaires were stored securely with password protection.

Although observation and recording conducted during lessons may affect normal classroom learning, the research did no harm to the students. During the study, students were aware of the presence of the observer and the recorder, which may affect their behaviour in the classroom. However, to minimise this effect, students were not informed about the exact date when they would be the focus for the observation. Not until the end of each lesson did they know whether they were in the group being observed during the previous lesson (see section 7.2.6 for details). This aims at making the study more ecologically valid.

The interviews were conducted in groups to collect a wider range of responses within limited time. Compared to individual, one-to-one interviews, the presence of others, particularly friends, might make participants withhold their real thoughts and feelings. For example, students may be afraid of ‘losing face’ in front of friends when talking about negative learning experiences, or feel under pressure to agree with the majority when holding different views. However, as students worked as groups in previous communication activities, it is believed that their interactions and discussions during group interviews may generate useful responses that could not be identified using individual interviews. Although participant anonymity could be compromised during group interviews, there was no way to connect the responses to relevant individuals after data collection.

All video recordings of lessons were taken from the back of the classroom, and all recordings of group interviews were audiotapes, on which individual identities could not be easily recognised. The recordings and other electronic data were saved on a password protected laptop and will be destroyed a year after the submission of this thesis.

7.2.5 Research design

Like Study 1, this study adopted a high-density repeated measurement design, measuring

state WTC in different communication activities during a lesson. However, for fear of disturbing students' normal learning in real classrooms, it measured state WTC at the end of a lesson rather than during (see section 7.2.6 for details). In each lesson, there were usually three or four communication activities. That is, for each participant in this study, three or four measurement occasions were collected.

To explore the diversity of thoughts, feelings and behaviours reflected by different individuals, observation in each lesson focused on a different group. Students in both classes were divided into five groups (about four students in each group). For the class randomly divided into five fixed groups by the teacher, existing groups were used. For the other class, students who sat close and did the communication activities collaboratively in class were grouped together. The students in the same group were interviewed together.

7.2.6 Procedure

The study lasted three weeks, from the last week of 2016 to the beginning of 2017. An introductory session was conducted for students and teachers. Students who agreed to participate in this study signed the consent form and completed a set of baseline questionnaires, on which they reported trait WTC and other trait-like predispositions and self-assessed their English proficiency (i.e. self-assessed baseline English score).

On the next day, both classes were observed. After these lessons, the teachers offered information relevant to the course, such as objectives suggested in the national curriculum, communication activities designed for the course, as well as their general impressions of student participation, performance, and English proficiency. Afterwards, the teachers responded to a brief questionnaire, rating each student's overall English proficiency (i.e. teacher-assessed baseline English score) and communication behaviour in previous lessons.

From the next lesson onwards, a camera on a tripod was placed in a rear corner of the classroom to record each lesson. As the camera was placed behind the students, it only captured a general picture of what was happening in the classroom. I, as a non-participant observer, observed students' communication behaviour and performance in different activities. Before each lesson started, I randomly selected a group as the focus of that specific lesson and informed the teacher, so she could pay more attention to those students'

communication behaviour and performance (which could be a supportive situation cue for those students). During the lesson, I sat close to the group of interest to observe each individual student's behaviour and performance and build rapport with the group after class.

While observing, I marked a structured observation schedule, and noted down any interesting points that I would like to discuss with the group. The students had regular classes as usual without any intervention. However, it should be noted that the presence of the researcher and/or the camera can make the learning situation less 'natural', as stress and embarrassment may be triggered, which may change what normally takes place (Richards & Morse, 2007). Not until the class finished, were students informed that they were chosen as the group under study, and a time slot was scheduled for a group interview. The participants were interviewed in groups to gain understanding of the thoughts and feelings of a relatively large number of participants in a short period of time. To maximise individual contributions rather than generating a group opinion, all individuals were given the chance to present their personal opinions.

After each lesson, I checked with the teacher to identify the major communication activities. Three quarters of observed lessons contained three major communication activities, and the remainder contained four. The teacher then rated on a scale from 1 to 7 concerning each individual's (in the group of that lesson) communication behaviour and language performance in each activity and stated her criteria for assessing language performance either orally or in writing. Conversations with the teachers were recorded with their permission.

As all lessons took place in mornings, all but one group interview was run within three hours of the lesson. The remaining one was carried out the next morning. All group interviews took place in the same classroom as the introduction session in Mandarin (the shared L1 of participants and researcher) and were recorded with participant permission. Each group interview normally lasted for about ninety minutes.

When piloting, students were asked to watch the lesson recording, reflecting on their thoughts, feelings and behaviours in each communication activity, and to stop the video whenever they wanted to talk. However, students in both pilot groups indicated that they could remember events in the activities, and were able to talk about their thoughts, feelings and behaviours without spending time watching the video. Hence, in the main

study, lesson recordings only served as cues for participants to recall what was happening at some points in time. The recordings were fast-forwarded to specific moments when students could not remember or disagreed about what had happened. When I asked the students to talk about occasions in class that I was interested in, I used the video clips to clarify which ones I was referring to.

All group interviews were conducted in a non-directive manner, with the interviewer serving as a group facilitator to prompt contributions rather than directing them. Participants had the chance to volunteer any response without any rule for talking or turn-taking. To encourage participants to respond honestly during the group interviews, they were kindly informed there were no right or wrong answers and no need to agree. Considering that a group interview may be dominated by one respondent, equal attention was paid to everyone to prevent any individual from dominating the interview. Some turns were deliberately offered to those who talked less. However, whoever verbally or non-verbally expressed a reluctance to respond was never forced to talk. Towards the end of each interview, students were asked to assess their performance in the lesson and explain their criteria for self-assessments. To ensure that all responses were understood correctly and completely, their contributions were briefly paraphrased soon after they talked, and affirmation or correction was asked. At the end of each interview, students were asked to complete a brief questionnaire, rating their state WTC, communication behaviour, and language performance in relation to each activity in that specific lesson.

A week after the last group interview, students took the paper-based final exam of the Intensive Reading course (i.e. the summative assessment). Soon after the exam, the course teachers reported the formative and summative assessment scores.

7.2.7 Materials

7.2.7.1 *Baseline measures*

Self-report. The baseline questionnaires were the same as those in Study 1 (see details in section 6.2.7.1). There were only two minor changes. One change took place in the instructions for creating confidential codes. The last two characters of student matriculation number were included in the confidential codes, as this allowed linking self-reports and teacher reports. The other was that students were asked to self-assess their English proficiency to indicate their baseline English scores, as they had not taken any exam when they responded to the questionnaire.

Teacher report. As this study aims to identify differences between WTC (both trait and state) and communication behaviour, as well as self-assessments and teacher assessments, teacher reports on communication behaviour and language performance were also collected. The baseline measure for either teacher was a list of the students in her class, on which the teacher rated everyone's communication behaviour on a seven-point scale and indicated baseline English scores by assessing everyone's English proficiency on a hundred-point scale (see Appendix F).

7.2.7.2 Momentary measures

Self-report. As the momentary questionnaire served as a supplement to the group interviews, it was much briefer than the one used in Study 1. Only the three items on state WTC, self-reported communication behaviour, and self-assessed performance were used (see details in section 6.2.7.2). Students responded to these questions about each communication activity, so to show their within-person variability in state WTC, communication behaviour and self-assessed language performance across different activities during a lesson (see Appendix G).

Teacher report. The momentary questionnaire given to teachers was similar to the one for students. The main difference was that the items on state WTC were not included, as it was impossible for teachers to observe student WTC as it is conceptualised as subjective communication intention. The teachers measured student communication behaviour and language performance in each communication activity in the lesson (i.e. 'He/She did communicate in English in this activity.' and 'He/She performed well during this activity. '; see Appendix H). Additionally, they were asked to comment on student language performance during the lesson and state their assessment criteria.

Group interview. The group interviews were semi-structured based on a brief schedule (see Appendix I). The schedule was adapted from Cao's (2014) stimulated-recall interview questions, which outlined general questions to lead into conversations (e.g. 'How did you feel/think about today's class?') and probing questions to explore thoughts and feelings at specific moments (e.g. 'Can remember what you were feeling/thinking when...?'). In addition, some questions were designed to focus on situational antecedents that showed correlations with state WTC in the results of Study 1 (e.g. support, task-interest, task-usefulness, communication opportunity, duty, positivity, negativity and sociality). For instance, support had been found as one of the most important situation

characteristics that significantly correlated with state WTC in the context of Study 1. Hence, questions about perceptions on peer and teacher support were added to the interview schedule (e.g. ‘Did you work well with your classmates/group members/partner during this activity?’ and ‘Did you feel supported by the teacher?’). Based on the findings of Study 1, a question was designed to explore whether and why the students showed lower levels of communication behaviour than state WTC (i.e. ‘Was there any point in time when you felt like to communicate but you did not talk? Why?’). Lastly, there was a question asking the students to self-assess their language performance and state their assessment criteria (i.e. ‘How will you assess your language performance in today’s class? Why?’). In practice, the interview schedule was flexible and could be adapted to the flow of the talk. The flexibility provided participants with opportunities to express any unpredicted responses.

Classroom observation. A structured observation schedule was designed based on Peng’s (2014) scheme, with some adaption to this context. The structured observation schedule (see Appendix J) was a blank table to be filled in during the observation. Each column represented an individual in the group, and each line represented a type of communication behaviour. Communication behaviour was categorised into three dimensions, nominated speaking, voluntary speaking, and non-verbal reaction. Under each, different kinds of communication behaviour were specified. For example, nominated speaking was subdivided into presenting in front of the class and answering a question. During the observation, I noted down the frequency of each type of communication behaviour observed from each student of interest. Additional notes about student behaviour were written under the table as a memo for the interview after the lesson.

7.2.7.3 *Language learning performance*

Baseline English scores. The self- and teacher-assessed English proficiency reported on the baseline measure were regarded as two indicators of baseline English scores. These assessments were on a hundred-point scale (see section 7.2.7.1).

End-of-term English scores. This study made use of the scores of the formative and summative assessments of Intensive Reading course to indicate exam performance. The formative assessment aimed at evaluating student participation and language performance in communication and other learning activities during the semester; however, it mainly relied on the course teachers’ subjective impressions. The summative assessment was a

paper-based exam collaboratively designed by the two course teachers to assess student English proficiency in terms of reading, writing and translation. The exam was also assessed by the two course teachers; however, as it mainly contained multiple-choice questions, the results seemed to be more objective than teacher subjective judgments. Both the formative and summative assessments were on a hundred-point scale, with scores below 60 as fail, between 60 and 70 as pass, between 70 and 80 as merit, and above 80 as distinction. Unlike the self-reported exam scores in Study 1, the scores in this study were reported by the course teachers.

Language performance in a specific lesson. Compared to the above assessments at the trait level, student language performance was also assessed at the state level. The self- and teacher-assessed performance reported on the momentary questionnaires were regarded as two assessments of student language performance in a specific lesson or activity. These were on a seven-point scale (see section 7.2.7.2).

7.2.8 Data analysis

7.2.8.1 *Statistical analysis*

Data collected through questionnaires were statistically analysed. As described in Study 1, descriptive analysis was conducted to quantify the amounts of variability observed in trait and state WTC, and correlation coefficients were calculated to analyse the relationships between WTC and language learning performance at trait and state levels, to show the predictability of WTC (see details in 5.2.8).

Unlike Study 1, this study collected statistical data not only through self-reporting, but also through teacher reporting. As a result, when applicable, student self-reports were compared to teacher reports. For example, the descriptive statistics of student self-assessed performance were compared to those reported by the teachers, to see whether there were differences. Additionally, correlation coefficients between self-reports and teacher reports were calculated to indicate relationships between these responses.

7.2.8.2 *Thematic analysis*

The data collected from group interviews and teacher conversations were analysed through thematic analysis, which involved researcher interpretation based on the literature (Cousin, 2009). The recordings were transcribed into Mandarin and, to get a general impression of them, were read through several times. After familiarising with the

data, thematic analysis was applied to systematically reduce the large amount of discourse into manageable data. The major feature of thematic analysis is coding, which is the translation of participant responses into themes for analysis (Kerlinger, 1970). Themes are defined as “recurrent and distinctive features of participant accounts, characterising particular perceptions and/or experiences, which the researcher see as relevant to the research question” (King & Horrocks, 2010, p. 150). A code is a label given to a piece of datum that helps the researcher to grasp the information at a quick glance and identify similar information from a large quantity of text (Cohen et al., 2011). Coding was systematically conducted in Mandarin throughout the transcripts line by line, and attention was paid to all individual contributions. A word or short phrase was written by the left margin of each piece of datum. Some pieces of text were given more than one code, as they conveyed two or more pieces of information that belonged to different labels. To keep the original meaning conveyed by the participants, their language was used at an early stage. At this stage, the aim was to identify the data that could be used to address the research questions, and the codes merely described participant responses rather than interpreting them (King & Horrocks, 2010). Hence, the initial coding was rough and descriptive, and some pieces of text too difficult to be coded were left pending for revision.

Through a process of re-reading and comparison, some codes from the earlier stage were subsequently modified, and the same codes were given to the pieces of text that conveyed the same meaning for consistency. Coding at this stage went beyond merely describing participant responses and emphasised researcher interpretation of the meaning (King & Horrocks, 2010).

After a few rounds of coding and re-coding, a manageable number of categories that seemed to capture participant responses were identified. The ‘core categories’ were critical factors of the experience being studied rather than incidental ones, and each core category was characterised by ‘sub-categories’ (Cousin, 2009). These thematic categories, more abstract themes than the interpretative ones, were drawn from existing theories in the literature (Cohen et al., 2011; King & Horrocks, 2010). In the literature review, the situational antecedents of state WTC had been summarised into a multi-layered framework based on findings in the literature. Hence, the situation characteristics directly influencing state WTC were grouped as core categories, and these were subdivided into sub-categories, the situation cues. Examples of core categories were task-usefulness, task-interest, lack of confidence and lack of knowledge; while examples of sub-categories

were type of activity, teaching style and learning material (see Figure 3.1). The links between categories were identified based on Rauthmann and colleagues' taxonomy of the underlying dimensions of situation characteristics (Rauthmann et al, 2014). The various situation characteristics were summarised into three dimensions, positivity, negativity, and duty. The frequency of each category and sub-category was counted. However, as noted by Cousin (2009), frequency is not the only measure of importance, as what are reported by most of the participants might not necessarily be the most important contributions. Some responses of the minority might offer a fresh angle, which could also be important as a complement to the majority. Lastly, the different categories and sub-categories were listed, to find out any relationships between them in accordance to previous findings in the literature.

Thematic analysis is systematic and replicable, as textual data are analysed step by step according to explicit rules of procedure (Mayring, 2004). It is also inductive and flexible, as pre-existing categories may be modified according to the data collected (Flick, 2014). Although thematic categories are theory-dependent, they must be supported by the interpretive themes coded based on data (King & Horrocks, 2010). During the process of data analysis, one may find that the theory does not perfectly fit the data and new themes emerge through researcher interpretation (Flick, 2014). Nevertheless, some researchers (e.g. Ezzy, 2002) note that the pre-existing categories might restrict researcher analysis. As researchers are aware of what they expect to find and have an agenda in mind of what categories will be applied, they might be confined by their agenda or imposition of interpretation rather than responding to interviewee responses (Ezzy, 2002). Another limitation of thematic analysis is that rigour might be sacrificed whilst reducing and categorising the large quantity of text into a few codes and categories (Cohen et al., 2011). As stated by Cohen et al. (2011), "the nuanced richness of specific words and their connotation" might be lost, as words with slightly different connotations or significance could be grouped into a relatively generic and ambiguous category (p. 573).

For the plausibility of interpretation of participant contributions, the coding of the textual data was checked by a second coder who is also in the field of SLA but not familiar with the construct of WTC. The second coder randomly selected one from the eight transcripts of the group interviews in the main study and coded throughout the transcript independently. Most codes given by the second coder were the same as those given by the first (about 85% overlap). For some data, although the codes given by the two coders

were different, the meaning was very similar. For example, the first coder labelled the cooperation between students in the same group as ‘peer support’, while the second coder coded it as ‘teamwork’. The reason might be that the second coder was not familiar with the literature on WTC, and thus could not always come up with the terminology used by the first coder. Whenever different codes were given to the same piece of datum, the two coders discussed until an agreement was made. This consistency between the two coders can be interpreted as evidence of the reliability of the thematic analysis (see Coolican, 2014; Cornish, Gillespie & Zittoun, 2014; Ryan, 1999).

After the thematic analysis, themes, codes, categories, sub-categories and selected quotations were translated from Mandarin into English. To ensure the accuracy of the translation, a professional translator was invited to back-translate the English versions into Mandarin. Any difference between the original transcripts and the back-translated versions was discussed with the professional translator to reach a consensus. The English translation was finalised based on the professional translator’s comments.

7.3 Results and Discussion

7.3.1 Results and discussion of the statistical analysis

The following sections will report key findings that emerged from the data collected through the questionnaires. I will present the descriptive statistics of trait and state WTC, quantify the amounts of observed variability in trait and state WTC, and report the correlation coefficients between trait and state WTC and different indicators of English learning performance. Inspired by the findings of Study 1, I will then compare the data concerning trait and state WTC to those concerning actual communication behaviour to show the differences between subjective communication intention and actual communication behaviour. I will also compare student self-assessments to teacher assessments, to investigate whether there is consistency between these two perspectives.

7.3.1.1 *Descriptive statistics*

Two sorts of data were collected in this study: trait data collected through the baseline questionnaires and state data collected through the momentary questionnaires. On the baseline questionnaires, students reported their trait WTC in Mandarin and in English. As shown in Figure 7.1 (panel 1), the mean score for the sample’s self-reported trait WTC in English was above 5 (mean = 5.15, *SD* = 0.79). This is higher than the average level of

trait WTC in English reported in Study 1 ($d = 0.81$). Additionally, only two students (6%) scored lower than 4 on a scale from 1 to 7 (min = 3.66), indicating that nearly all the English major students under study were relatively willing to communicate in English in general.

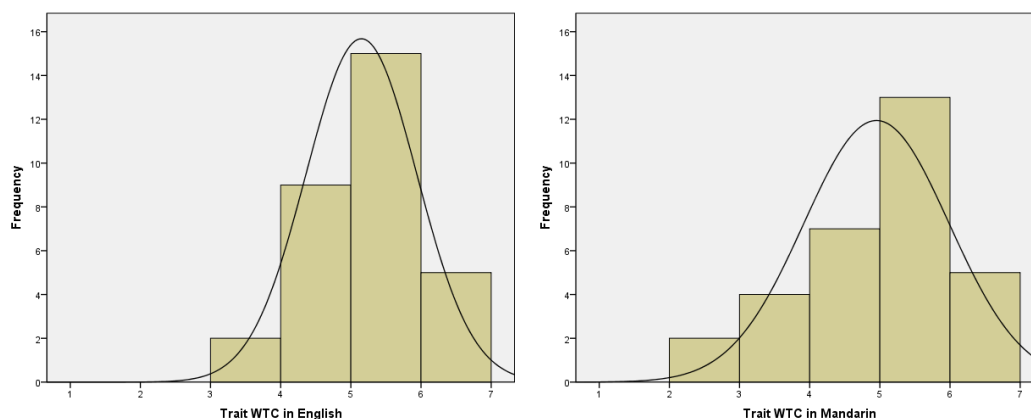


Figure 7.1. Distributions of English major students' trait WTC

In this study, trait WTC in English significantly and positively correlated with trait WTC in Mandarin ($r = .63$). However, it seems that (see Figure 7.1) not all students were as willing to communicate in Mandarin as they were in English, as the mean of trait WTC in Mandarin (mean = 4.95) was slightly lower than that in English ($d = 0.22$). Moreover, individuals differed more in trait WTC in Mandarin ($SD = 1.04$), with one fifth of the students scoring lower than 4 (min = 2.62). Nevertheless, this sample's average level of trait WTC in Mandarin was higher than that reported by the sample in Study 1 ($d = 0.39$), indicating that these English major students were generally more talkative than the non-English major students in Study 1.

At the state level, each student's scores for state WTC were averaged within-person across the three or four communication activities in a lesson, and thus each mean score represented an individual's average level of state WTC in that lesson (see Figure 7.2 for the distribution of the sample). The sample's grand mean of state WTC (mean = 5.43) was somewhat higher than that of trait WTC in English reported earlier ($d = 0.28$). Moreover, individual students differed more in their state WTC ($SD = 1.19$) than in trait WTC in English. The findings indicate that while most students were quite willing to communicate in English in general, some felt very willing and others felt unwilling to communicate in the specific lessons under study.

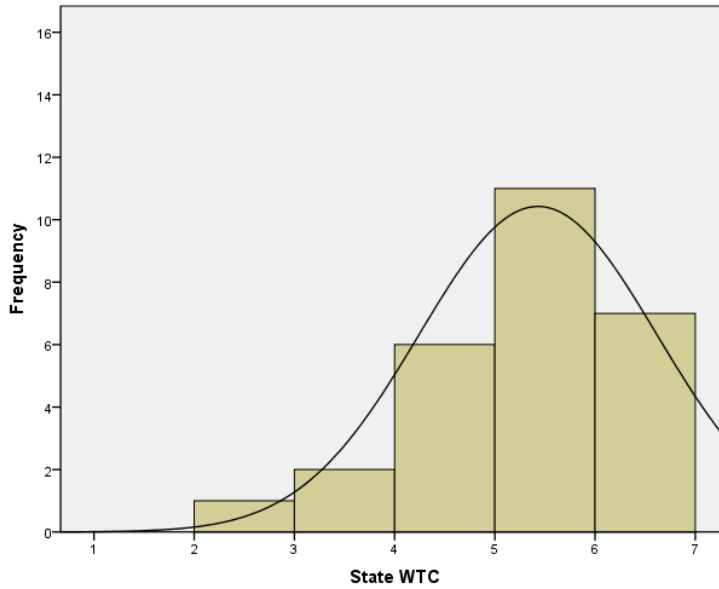


Figure 7.2. Distribution of English major students' state WTC (aggregate level)

7.3.1.2 *WTC and communication behaviour*

Relationships and differences between subjective communication intention (WTC) and actual communication behaviour were analysed at both trait and state levels. At the trait level, WTC was self-reported by students, while communication behaviour was reported by teachers, based on their general impressions. It was found that student self-reported trait WTC in English positively correlated with teacher observations of communication behaviour at the trait level ($r = .49$). However, the sample's average level of communication behaviour reported by teachers (mean = 4.55) was lower than its average level of self-reported trait WTC ($d = 0.56$). Moreover, the amount of between-person variability (i.e. individual differences) in teacher-reported communication behaviour ($SD = 1.29$) was higher than that in self-reported trait WTC ($SD = 0.79$). With a minimum score of 3.66, only two students (6.5%) scored slightly below 4 on the trait WTC in English scale, indicating that almost all students were relatively willing to communicate in English in general; however, teacher reports on communication behaviour showed that only about half of the students scored higher than 4, with a minimum score of 2, indicating that only a minority actively communicated and a few students hardly ever communicated (see Figure 7.3).

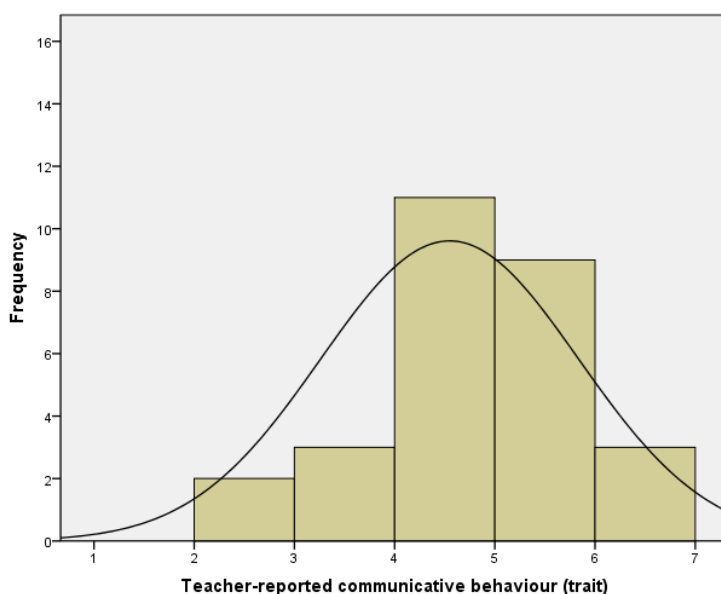


Figure 7.3. Distribution of teacher-reported communication behaviour (trait)

To further explore the differences between WTC and communication behaviour at the trait level, their relationships with Big-five personality traits, trait WTC in Mandarin, and self-assessed baseline English scores were tested. As shown in Table 7.1, trait WTC in English significantly correlated with openness to experience and extroversion, and the correlation between trait WTC and extroversion ($r = .41$) was weaker than that between trait WTC and openness to experience ($r = .55$).

However, teacher-reported communication behaviour at the trait level was only significantly related to extroversion ($r = .47$). This finding indicates that intention to communicate in a L2 might be jointly influenced by different personal variables, such as openness to experience and extroversion; however, actually communicating with others in the L2 largely depends on extroversion or talkativeness.

At the state level, student communication behaviour was reported both by themselves and their teachers through the momentary questionnaires. Scores for self-reported and teacher-reported communication behaviour were averaged across activities. Hence, everyone obtained two mean scores related to communication behaviour: one was self-reported, and the other was reported by a teacher. The distributions of individual mean scores are presented in Figure 7.4.

Table 7.1. Descriptive statistics for trait variables in Study 2 ($N = 30-31$)

	Mean	<i>SD</i>	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.
1. Trait extroversion	4.19	1.25	(.90)										
2. Trait agreeableness	5.58	0.73	.13	(.77)									
3. Trait conscientiousness	4.92	0.77	.06	.44*	(.68)								
4. Trait emotional stability	3.95	0.76	.32	.35	.51**	(.88)							
5. Trait openness to experience	4.83	0.65	.24	.05	-.06	-.28	(.68)						
6. Trait WTC in Mandarin	4.95	1.04	.77**	.38*	.26	.28	.25	(.95)					
7. Trait WTC in English	5.15	0.79	.41*	.27	.18	-.03	.55**	.63**	(.94)				
8. Teacher-reported communication behaviour	4.55	1.29	.47**	.04	.13	.30	.12	.45*	.49**				
9. Self-assessed baseline English score	62.73	17.30	.53**	.14	.23	.14	.19	.36*	.38*	.41*			
10. Teacher-assessed baseline English score	83.58	7.54	.08	-.06	-.08	.12	.07	.06	.24	.51**	.31		
11. End-of-term teacher judgment	89.39	5.99	.28	.06	.32	.28	.04	.33	.47**	.59**	.43*	.30	
12. End-of-term paper-based exam score	68.92	12.68	.29	.09	-.06	.12	.07	.27	.15	.43*	.47*	.61**	.50**

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

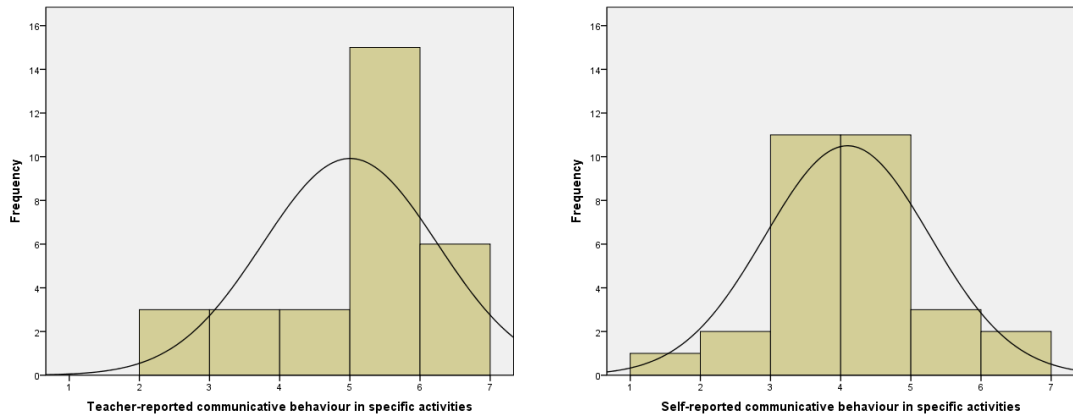


Figure 7.4. Distributions of teacher- and self-reported communication behaviour (aggregate level)

On average, self-reported communication behaviour (mean = 4.10) was lower than teacher-reported (mean = 5.00; $d = 0.74$), and state WTC (mean = 5.43) was higher than both self-reported ($d = 1.12$) and teacher-reported ($d = 0.35$) communication behaviour. Communication behaviour self-reported by the students and observed by the teachers positively correlated ($r = .43$); however, it was interesting that no significant correlation was found between student communication intention (i.e. state WTC) and actual communication behaviour at the state level, either reported by the students ($r = -.01$) or by the teachers ($r = .16$; see Table 7.2).

Table 7.2. Descriptive statistics for state variables in Study 2 ($N = 31$)

	Mean	<i>SD</i>	1.	2.	3.	4.	5.	6.
1. State WTC	5.43	1.19						
2. Self-reported communication behaviour	4.10	1.18	-.01					
3. Teacher-reported communication behaviour	5.00	1.25	.16	.43*				
4. Self-assessed performance in a lesson	4.33	1.08	.44*	.25	.16			
5. Teacher-assessed performance in a lesson	5.40	0.98	.09	.41**	.79**	.23		
6. End-of-term teacher judgment	89.39	5.99	-.00	.37*	.30	.18	.61**	
7. End-of-term paper-based exam score	68.92	12.68	.28	.43*	.49**	.15	.40*	.50**

*. Correlation is significant at the 0.05 level (2-tailed).

**.. Correlation is significant at the 0.01 level (2-tailed).

7.3.1.3 Variability in trait and state WTC

To investigate variability in state WTC at the aggregate level, state WTC in each communication activity was averaged across individuals. The aggregated trajectory shown in Figure 7.5 (panel 1) depicts the fluctuation in the sample's state WTC across different communication activities. It seems that state WTC did not vary much across different communication activities, as the grand mean slightly increased from 5.23 to 5.84

and then went back to 5.63.

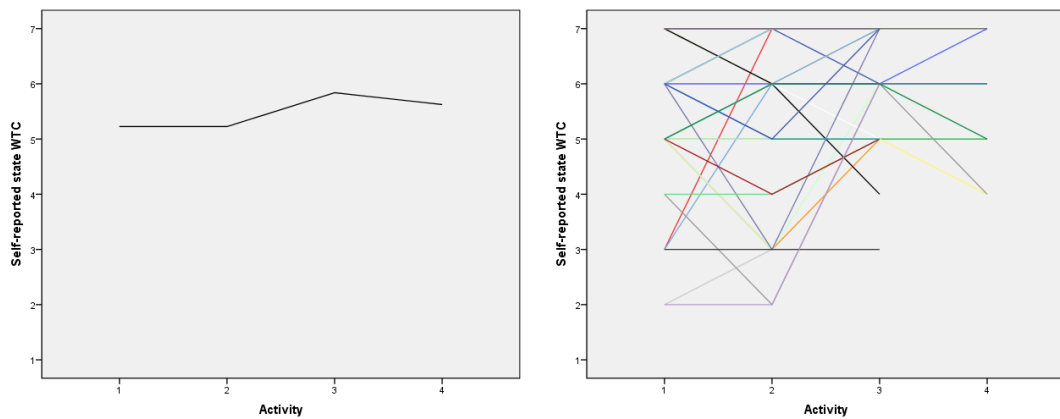


Figure 7.5. Aggregated and individual trajectories of state WTC within one lesson

However, when looking at individual trajectories, some variability was identified. To note, different individuals' state data may be collected in different lessons. Some lessons had four major communication activities, whilst others had three. Hence, several individual trajectories end at Activity 3 and do not continue to Activity 4 in Figure 7.5 (panel 2). This figure provides evidence for both between- and within-person variability in individual state WTC across different activities. Individual differences in within-person variability is also shown in Figure 7.5, although there are overlaps between individual trajectories.

Like the data analysis in Study 1, amounts of variability in trait and state WTC were quantified using standard deviations. The first bar in Figure 7.6 shows the total amount of variability in state WTC, i.e. the standard deviation of all scores across different communication activities regardless of individuals. The second bar in the figure shows a typical individual's within-person variability in state WTC. Everyone's standard deviation of scores for state WTC across the three or four communication activities within a class was calculated, representing individual within-person variability in state WTC. The standard deviations were then averaged across individuals, to show the typical amount of within-person variability. The right two bars show the amounts of between-person variability in WTC at trait and state levels, respectively. The between-person variability in trait WTC was quantified using the standard deviation of scores for trait WTC in English reported on the baseline questionnaire, while the between-person variability in state WTC was quantified using the standard deviation of individual mean state WTC across different activities.

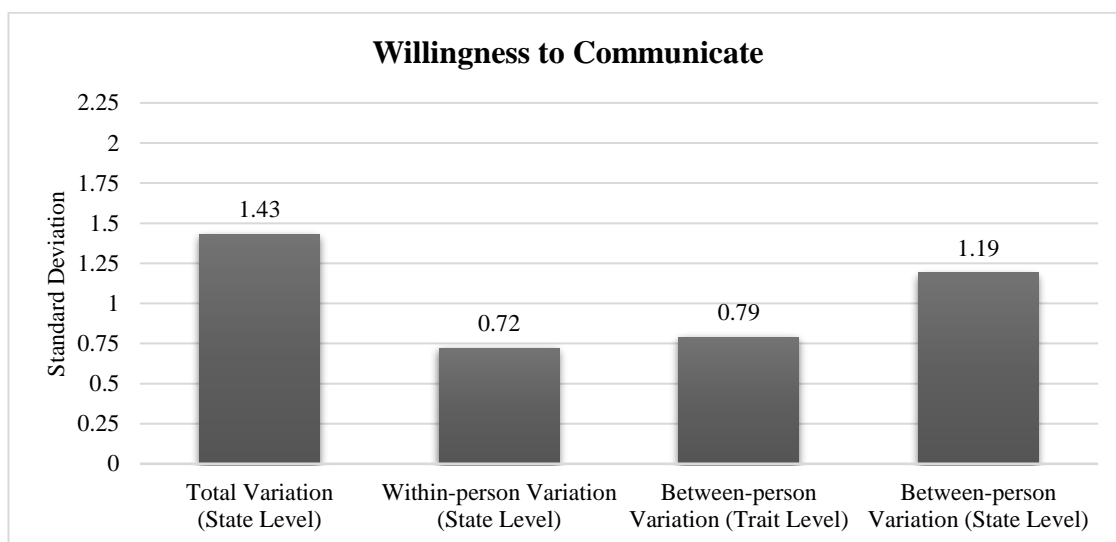


Figure 7.6. Variability in English major students' trait and state WTC

In sum, the sample showed not only individual differences in trait WTC, but also between- and within-person variability in state WTC. The amount of within-person variability in state WTC accounted for 60% of between-person variability. The findings suggest that state WTC tends to fluctuate across different communication activities during a lesson, and the amount of variability within individuals is somewhat lower but still comparable to the amounts of variability observed between them. Moreover, between-person variability in state WTC was higher than that in trait WTC. The amount of between-person variability in trait WTC only amounted to 66% of between-person variability in state WTC. That is, individuals tend to differ from each other in state WTC much more than they differ in trait WTC, which also indicates that state WTC in specific situations might deviate from trait WTC.

7.3.1.4 *WTC as a predictor of language learning performance*

The predictability of WTC was tested at both trait and state levels. At the trait level, language learning performance was indicated by scores of the formative (i.e. end-of-term teacher judgment) and summative assessments (i.e. end-of-term paper-based exam score) of the Intensive Reading course. Scores on the paper-based exam and teacher judgment significantly correlated ($r = .50$). However, the mean score on the paper-based exam (mean = 68.92, $SD = 12.68$) was 20 points lower than the mean score of the teacher judgment (mean = 89.39, $SD = 5.99$), and individual scores on the paper-based exam deviated much more widely than scores of the teacher judgment (see Figure 7.7). All but one student (97%) were assessed as distinction on the teacher judgment (min = 75);

whereas only five students (17%) received distinction on the paper-based exam and seven students (23%) failed (min = 41). That is, although nearly all students in this study performed very well in the communication and other learning activities held in the classroom during the semester, several did not perform well in the end-of-term exam.

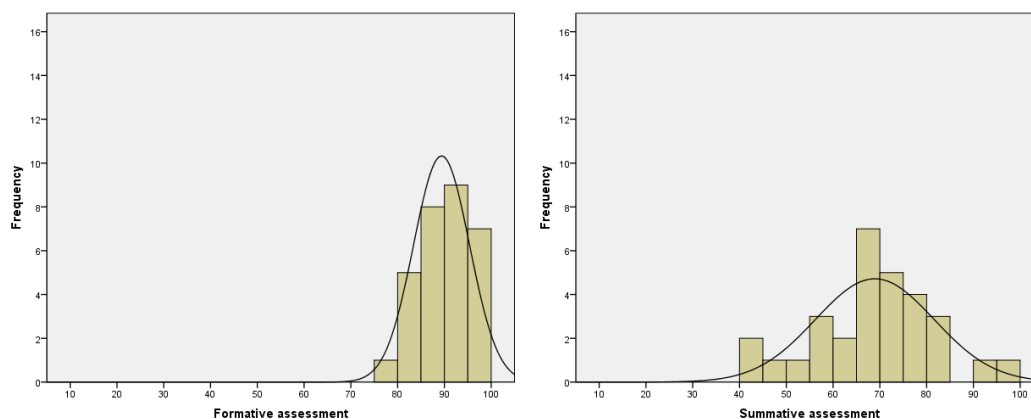


Figure 7.7. Distributions of English major students' end-of-term English scores

Correlational tests were conducted to find out the relationships between trait WTC in English and end-of-term English scores. As shown in Table 7.1, the relationship between trait WTC and performance subjectively assessed by teachers (i.e. end-of-term teacher judgment) was significant ($r = .47$), but the relationship between trait WTC and performance on the paper-based exam was weaker and insignificant ($r = .15$). To note, Study 1 found that trait WTC significantly correlated with end-of-term English scores (a combination of the paper-based exam and teacher subjective judgment) reported by the students.

At the state level, language performance in specific lessons and activities was assessed by both students and course teachers. After averaging everyone's state scores across different activities, correlation coefficients between state WTC and self- and teacher-assessed performance in a lesson were estimated (see Table 7.2). Self-assessed performance positively correlated with state WTC ($r = .44$); however, teacher-assessed performance was not associated with state WTC ($r = .09$), but significantly associated with actual communication behaviour, either reported by students ($r = .41$) or observed by teachers ($r = .79$). Moreover, the relationships between teacher- and self-assessed performance in a lesson were not statistically significant ($r = .23$).

7.3.1.5 Teacher- and self-assessed language performance

This study further compared student and teacher subjective assessments of performance, and their relationships with language exam scores. The comparison was made at both trait and state levels.

At the trait level, student baseline English scores were subjectively assessed both by themselves and their teachers. As shown in Figure 7.8, there were dramatic differences between self- and teacher-assessed baseline English scores. In general, the mean of the self-assessments was 20 points lower than the mean of teacher assessments ($d = 2.06$). Self-assessed baseline English scores varied remarkably from 10 to 90 ($SD = 17.30$), with a third lower than 60, another 15% exactly on 60, and only 27% higher than 80. However, from the teacher perspectives, 80% of students were above 80 and nobody was below 60. In other words, most of the students were not confident in their baseline English proficiency; however, the teachers were satisfied and encouraging, as their baseline assessments were highly favourable. As shown in Table 7.1, the self- and teacher-assessed baseline English scores positively correlated ($r = .31$); however, this was not statistically significant. Whether self- and teacher-assessed baseline English scores correlated with scores on the paper-based language exam was also tested. Although the average paper-based exam score (mean = 68.92) seemed to be closer to the average self-assessed baseline score, paper-based exam scores were more strongly associated with teacher-assessed baseline scores ($r = .61$) than self-assessed baseline scores ($r = .47$).

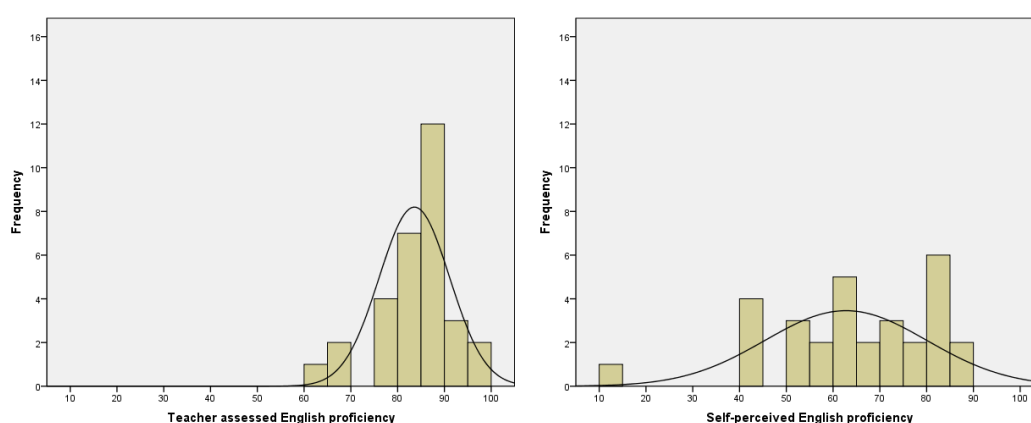


Figure 7.8. Distributions of teacher- and self-assessed baseline English scores

At the state level, language performance in different activities during a lesson was assessed both by students and their teachers. Individual scores of self-assessments and

teacher assessments were averaged across activities. Hence, each student received a mean score representing self-assessed performance and another representing teacher-assessed performance. As presented in Figure 7.9, the findings were like those at the trait level. On average, the mean score of self-assessed performance in a lesson (mean = 4.33) was considerably lower than that of teacher-assessed performance (mean = 5.40; $d = 1.04$). Only half of the students were satisfied with their language performance (i.e. scored higher than 4); while the teachers positively assessed around 90% of the students. As shown in Table 7.2, teacher-assessed performance in a lesson significantly correlated with the paper-based exam scores ($r = .40$); however, the correlation between self-assessed performance and exam scores was weaker and insignificant ($r = .15$). Moreover, teacher-assessed performance was strongly associated with their observation of student communication behaviour ($r = .79$); however, self-assessed performance was not significantly associated with self-reported communication behaviour ($r = .25$) but was significantly associated with state WTC ($r = .44$).

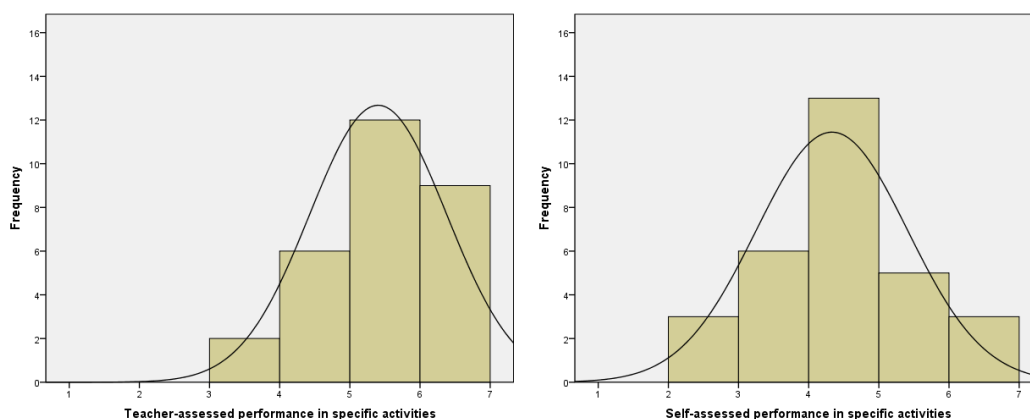


Figure 7.9. Distributions of teacher- and self-assessed performance in one lesson

7.3.1.6 Discussion of the statistical analysis

This study attempts to show between- and within-person variability in trait and state WTC during a lesson and examine between-person relationships between trait and state WTC and language learning performance. A group of English major undergraduate students and their language teachers were recruited to respond to a set of questionnaires measuring student WTC, communication behaviour and performance at both trait and state levels.

It seems that this sample of students are willing to communicate in English, although their level of actual communication behaviour tends to be somewhat lower than the level of

intention to communicate. Results of the baseline measurement indicate that trait WTC in English classrooms tends to differ between individuals, while results of repeated momentary measurements suggest that state WTC may not only differ between individuals, but also fluctuate within individuals across different communication activities during a lesson. In general, trait WTC could predict language performance, particularly communication performance subjectively assessed by teachers; however, in specific lessons state WTC could only predict self-assessed performance but not teacher-assessed performance. When comparing between teacher- and self-assessed performance, it seems that teacher-assessed performance was more strongly related to actual communication behaviour, whereas self-assessed performance was more related to subjective communication intention (i.e. WTC). Compared to student self-assessments, teacher assessments could be better indicators of actual language performance because they seemed to be more strongly related to the scores on the end-of-term language exam.

The above results are generally consistent with the results of Study 1 and previous research on short-term within-person variability in state WTC (e.g. MacIntyre & Legatto, 2011; Mystkowska-Wiertelak & Pawlak, 2014; Pawlak & Mystkowska-Wiertelak, 2015). While this study was conducted, other studies were carried out by Mystkowska-Wiertelak and Pawlak (e.g. Mystkowska-Wiertelak, 2016; Pawlak et al., 2016). These studies were published just before data collection of the current study. Like the current study, these collected data in real L2 classrooms; however, they measured state WTC from time to time during a lesson (e.g. every 5 minutes in a 55-minute period), which was different from the current study. Measuring on a moment-to-moment basis, they paid attention to within-person variability in state WTC during a lesson, i.e. both within and across different activities. For example, Pawlak et al. (2016) found that state WTC could be triggered by the onset of new activities, but state WTC during each activity tended to decrease with time. However, for fear of disturbing language learning in class, the current study only measured state WTC once in each communication activity, focusing on within-person variability in state WTC across different activities.

To obtain more in-depth explanations of the psychological processes of communication generation within individuals, as well as differences between state WTC and communication behaviour, and between teacher- and self-assessments, students were interviewed in groups after class. Findings from the data collected through the group interviews will be presented in the following section.

7.3.2 Results and discussion of the thematic analysis

This section will present the thematic findings of this study. The themes coded from the transcripts of group interviews will be elaborated, and quotations will be provided to support interpretations. To note, all quotations were originally in Mandarin, and were translated into English. Firstly, I will present themes relevant to subjective perceptions of the learning situation in the classroom (i.e. situation characteristics), and how – according to the students – these situation characteristics affected their state WTC. Moreover, explanations to the disparity between state WTC and communication behaviour will be presented. Finally, I will compare student criteria for assessing language performance in communication activities to teacher criteria, to illustrate why student self-assessed performance was not consistent with teacher-assessed performance. To reduce traceability, student gender will not be indicated in the quotations. End-of-term paper-based exam scores will be provided after the quotations using bands such as fail, pass, merit, and distinction (see section 7.2.7.3).

7.3.2.1 *Positivity*

Positivity refers to any positive feeling elicited by a classroom learning situation, which tends to trigger or enhance state WTC. In this context, positivity captures situation perceptions such as task-interest, task-usefulness, support, inspiration, and motivation.

Task-interest. A clear majority of students in this study reported that their state WTC increased when they felt the tasks were interesting, and their perceived task-interest was primarily influenced by types of activities. They tended to be more interested in dialogues in which their views and opinions were solicited (e.g. group discussions and debates) compared to monologues (e.g. giving talks and reports). In addition, some students said that they enjoyed being introduced to new activities (e.g. story-retellings) that they had seldom experienced before.

“When talking about my own opinions, though I hadn’t prepared beforehand, I felt extremely happy at that moment. It’s because I really enjoy... being asked by others about my viewpoints on a topic, just as I like the discussion.” (student 12, fail)

“It (story-retelling) was a new idea. I felt like I was having something new in each class. Every time there was something new in class, I felt a bit happier.”

(student 2, pass)

Apart from types of activities, learning materials used in activities also influenced student perceptions of task-interest. Almost all the students said that the textbook was boring, as the texts in it were old fashioned and there were too many exercises after each text. By comparison, they noted that they felt more interested in the authentic materials selected by their teachers (e.g. English novels, English newspapers, and storybooks for American pupils).

“Maybe I participated a bit more actively because I enjoyed when the teacher was talking about the novel. But once she talked about the exercises in the textbook, I felt very bored.” (student 2, pass)

“We all thought the newspaper was good. Content in the newspaper were interesting. Then, it seemed that we didn’t enjoy the textbook. The textbook... I didn’t enjoy it. The texts weren’t well-selected.” (student 8, distinction)

“I felt bored talking about the textbook. It wasn’t interesting, so I was absent-minded. But, for example, in previous sessions, we had another book... We learned from something like a reading material for American pupils. That one was relatively new, which was more likely to make me um... feel interested and engaged. But the textbook was always like that, which made me feel bored.”
(student 13, fail)

A few students pointed out that their interest in group activities was also influenced by familiarity with other group members. It was reported that working with unfamiliar classmates was more interesting than working with close friends. As students may have scarce opportunities to communicate with unfamiliar peers after class, group activities were seen as good opportunities to share ideas and seek others’ viewpoints. By comparison, familiarity between friends may reduce their interest in listening or talking in communication activities in class.

“Actually, I enjoyed talking with different people. I think, when talking with unfamiliar people, I was more willing to express my opinions and more interested in their opinions. If we had known each other very well, I wouldn’t have been... that willing to listen or to talk.” (student 4, distinction)

These results generally support Dörnyei's (2009) claim that interest is a prime motivator of task participation and are in line with Eddy-U's (2015) findings that students are interested in activities that break class routine and stimulate new perspectives. In contrast to previous studies (e.g. Cao & Philp, 2006; Eddy-U, 2015; Kang, 2005; Pawlak & Mystkowska-Wiertelak, 2015; Riasati, 2012), which reported that students are more willing to communicate with familiar interlocutors, this study found that students tend to be more interested in exchanging information and opinions with unfamiliar classmates.

Task-usefulness. A large body of students reported that their state WTC was high in some communication activities, as they felt the tasks were useful and they would benefit from participation. From this perspective, they preferred communication activities such as group discussions and questions and answers after presentations, in which their thinking and oral English could be trained simultaneously.

"I felt it was challenging standing up alone and reporting because the teacher would ask me questions. After my report, she might randomly ask me some questions. Then, I needed to think and, at the same time, organise my language. It was good. I quite like the Q&A, communication with the teacher. It was good." (student 24, merit)

"I thought the discussion could train my expression, and... my skills of thinking and drafting, which were important as well. Not only were the textbook, grammar, vocabulary, and exercises useful. I thought the former (the discussion) was also... If you did participate in it and learn from it, it could be useful." (student 4, distinction)

By contrast, some students reported that some communication activities were unnecessary because the tasks were too easy to trigger effective communication or learning. They suggested they did not intend to communicate during these tasks, as they knew that they could hardly benefit from the outcomes or the communication processes. They believed that everyone was able to accomplish these easy tasks without communication, so those who took the tasks seriously might be seen as showing off and be laughed at by others.

"The question was so easy that I felt unwilling to raise my hand. I thought others could answer it as well. I could answer it, but... if I had raised my hand, others would have... laughed at me... (as if) I was showing off." (student 2, pass)

These findings on task-usefulness are consistent with Zhong's (2013) and Peng's (2012) finding that participation in communication activities may be influenced by students' beliefs about the effectiveness of these activities. Students tend to be more willing to participate in activities that could contribute to their learning than activities that are interesting but meaningless (Peng, 2012).

Support. Most students reported that teacher support enhanced their state WTC in some communication activities, and they perceived it based on their teachers' verbal expressions such as encouragement and feedback. Some students expressed their eagerness for detailed feedback after talking in English. They preferred critiquing comments on content and language rather than short confirmatory phrases (e.g. very good, well done, etc.) because they regarded teachers' constructive advice and error corrections as ways of showing concern and giving assistance.

"In fact, before I raised my hand, I recited to myself. I did well, and I could remember a lot. When I stood up, I forgot everything. Yes, but the teacher encouraged me and said that I did quite well. Indeed, I thought I didn't perform well, because I stumbled, forgot something and got stuck. So, I thought the teacher was very kind." (student 1, merit)

"I felt it (the teacher's error correction) was okay because the teacher wasn't assessing me. I thought she wanted to help me, so it was acceptable. I didn't have any... negative feelings. And I felt that she cared about me." (student 25, pass).

Some students suggested that their perceptions of classmate support also affected their state WTC. By contrast to teacher support, which was manifest in verbal expressions; classmate support was mainly manifest in non-verbal encouragement as well as harmonious relationships among the students in the classroom.

"My classmates gave the mark to the person who made the most progress this time. I thought they did well... giving a bit more encouragement to others. I think um... Well, I felt my classmates were quite considerate. That's right. I think it was more important than the mark as such." (student 3, distinction)

"No, I wasn't afraid of losing face... My classmates might laugh while I was talking, but I knew they didn't mean to laugh at me. They just laughed. It didn't

matter.” (student 1, merit)

According to the students, support from group members made great differences to state WTC in group activities. Group support was perceived not only based on encouragement and cooperation within a group, but also on individual English proficiency. A few students reported that they felt extremely secure when communicating with more proficient learners because of the support they provided.

“All of my group members raised their hands, and I was the only one who didn’t. So, I raised my hand too. Well, they were extremely active. I was the only remaining one. Then, they encouraged me ‘come on, raise your hand!’” (student 5, merit)

“Originally, I merely wanted to answer one question, but the teacher asked many add-on questions. Because I already felt nervous when answering the first question and the teacher continued asking me more, my mind was a complete blank. Indeed, I couldn’t come up with anything, so I hoped he (another student in the group) could provide me with some hints. He did provide me with many hints. Um, I was very delighted and relieved.” (student 26, pass)

“XYZ (another student in the group) was my major source of security. Yes, because I thought she was really excellent in English. When I had nothing to say, XYZ took the turn. Then I listened to her and... I enjoyed talking with her. I could check with her whether I was correct. If she felt it (my language) was okay, then I felt secure. As the saying goes: I feel secure when I receive confirmation from a straight-A student.” (student 1, merit)

As for the more proficient learners, they sometimes deliberately gave up some communication opportunities to their classmates, to show support and avoid dominating the class. Rather than interacting with their teachers instantly, these students noted that they preferred waiting to see if someone intended to respond. It indicates that when more proficient learners kept silent, they were not necessarily unwilling to communicate. Instead, it was another form of peer support. Some students reported that they were delighted to hand over turns to their classmates and regarded these as opportunities to gain new perspectives from others.

“Um... I waited to see... to see whether someone else would like to talk. Well, it

was because I thought... if someone knew... the answer or someone wanted to talk, then he (or she) would stand up to talk. Um... also, because I didn't want to... I mean, answer the questions all the time." (student 3, distinction)

"As for that topic, so many people raised their hands, so I didn't. I left the opportunity to others, because they seldom... Our classroom was very silent indeed. It was a good opportunity for us to brainstorm." (student 11, merit)

Similar to the results of this study, previous studies have (e.g. Peng, 2012; Riasati, 2012) suggested that a relaxing classroom atmosphere, in which the teacher is supportive, and the students know and support each other, probably leads to higher levels of state WTC. In contrast to Kang's (2005) claim that L2 learners tend to be less willing to communicate with interlocutors who have higher L2 proficiency, this study suggests that there may be students who prefer to communicate with more proficient peers. This is in line with Cao's (2011), Eddy-U's (2015), and Yu's (2015) findings.

Inspiration. More than half the students indicated that, although they felt unwilling to participate at the beginning of some activities, they were inspired as the communication proceeded. For most situations, the students were inspired by the opinions expressed by their classmates and were eager to argue with or comment on classmates' opinions immediately and directly. Under these circumstances, teacher intervention could interrupt thinking and thus reduce state WTC.

"If you play back the video, you will find that my hand was also raised. Well, I hadn't prepared well, but after listening to them, I felt... I had much to say, especially about what was said by XYZ (one of the classmates)." (student 6, merit)

"When we were willing to talk, the teacher didn't need to call anyone. I mean, whoever wants to talk, okay, stand up and talk. When we wanted to argue, we preferred stating our arguments immediately after someone finished talking. She (the teacher) interrupted, which made me forget (my arguments). That's why I didn't talk." (student 5, merit)

Some students suggested that they felt inspired by teacher questions, instructions, and explanations. When reading by themselves, they could only understand the texts literally, and their attention was drawn to grammar and vocabulary rather than the information

conveyed in the texts. However, by thinking in-depth about the questions asked by their teachers or comprehending the texts based on teacher explanations and interpretations, some students started to pay attention to the underlying connotations of the texts and relate them to personal experiences. Hence, willingness to share or exchange thoughts and feelings were triggered.

“After we read it (the novel), the teacher asked some questions. I felt the questions asked by her were so deep that I had never thought about them. I merely... read through and checked the new words. I hadn’t dug into the characters or their thoughts such stuff.” (student 3, distinction)

“The course we were attending was called Intensive Reading, but in previous sessions the teacher didn’t deeply analyse the connotations of the texts and so forth. Um, I felt what we did in those sessions wasn’t intensive reading. Maybe those texts weren’t deep, so the teacher couldn’t find anything deeper for us to learn. The text we read today, I think, was very good, so she (the teacher) asked us to, how to say, find out the topic and supporting sentences, and state our thoughts, the viewpoints for ourselves. I felt it was quite good.” (student 19, pass)

“It was also relevant to how to behave yourself... I perceived something deep from the novel that I had never paid attention to. Then I felt, wow, we should behave in this way. I hope the teacher could add something like this when talking about the textbook.” (student 2, pass)

In the literature, there have been findings like the current ones in suggesting that L2 learners’ state WTC can be influenced by interlocutor contributions (e.g. Cao & Philp, 2006; Kang, 2005; Riasati, 2012). However, unlike previous studies that emphasise the active participation of interlocutors in prolonging the conversation, responses in the current study highlight the content of interlocutor discourse in deepening the conversation.

Motivation. Several students reported that they participated in communication activities because they were motivated by the rewards such as marks or gifts offered by their teachers. These students felt indifferent to communication activities held in class; however, they might be motivated when the teachers encouraged them by saying that additional marks or some gifts would be given as rewards for those who participated. These students explained that what they really needed were not the marks or gifts as such,

but motives for them to participate.

“The class was divided into two sides, and they competed against each other. For example, you say something and then I argue, like a debate. The teacher said that the side that was stuck would lose a mark. So, when it began, whatever the question was, there was always someone to answer it. Um... I quite liked it because I felt... I had the motivation (to communicate).” (student 3, distinction)

“I think we can afford the notebooks and pens as well, but if they are given by the teacher as gifts, they are special. Yeah, a reason for me to raise my hand.” (student 9, fail)

However, some students noted that, although they were motivated by rewards and participated in the communication activities, they did not enjoy the feeling of being compelled to communicate to gain marks or gifts. Additionally, a few students argued that unwillingness to talk could not be changed by rewards.

“At that time, I answered the questions to gain a mark. The feeling was weird. Well, I felt I was forced... fed up with these childish treats” (student 10, distinction)

“I felt I was the same as usual. I thought... although the teacher would give me additional marks, I wouldn’t participate when I was not willing to do so.” (student 1, merit)

These results generally support Eddy-U’s (2015) claim that marks could be regarded as short-term goals that motivate L2 learners to participate in communication activities; however, some other studies (e.g. Pawlak & Mystkowska-Wiertelak, 2015; Riasati, 2012) have suggested that marks and the feeling of being monitored could give rise to anxiety, and thus prevent students from freely expressing their thoughts and feelings.

7.3.2.2 Negativity

By contrast to positivity, negativity refers to any negative perception of learning situations that might hinder L2 learners’ willingness to communicate in classrooms. Negative perceptions reported by the students in this study include lack of confidence, lack of capability, lack of knowledge, inattentiveness, nervousness, embarrassment, and fear.

Lack of confidence. In this study, the most frequently reported negative perception

affecting state WTC was lacking confidence in conducting effective communication in English, particularly for students whose English proficiency was low. Although they were English major students, they were still at the beginning stage, and were not proficient English speakers. When learning English in secondary schools, they might have had few opportunities to practice their oral English, as English teachers in secondary schools in China tend to follow traditional grammar-translation pedagogy, focusing on grammar rules rather than communication skills (Song, 2000; Yu, 2001). As a result, most students reported that they lacked confidence in their oral English skills and were afraid of making errors or mistakes while communicating. For example, some students noted that after receiving feedback from teachers in the university, they realised that there were problems with their English pronunciation, which reduced their confidence in communication activities.

“Um... when we were in high school... we did reading comprehension exercises all the time. I mean, do (the exercises) and check (the answers), do (the exercises) and check (the answers), you know. Because of that, I mean, um... I was used to (the learning activities in) my high school. Then, I suddenly felt uneasy (when doing the communication activities) in the university. We used to learn loads of grammar points and remember vocabulary, to pass exams. For those exercises you just need to skim the texts, understand the main ideas, and then do the multiple-choice questions. However, it seems that the tasks are more difficult now because you may be asked to talk about your own understanding.” (student 10, distinction)

“For example, answering the question in this class. I was willing to answer, but I didn’t know how to use the language (English). I didn’t know. My oral English was poor, so I... was afraid of speaking and... making... mistakes... Well, I couldn’t master the language (English) as... fluently as other students did, to state personal opinions. I couldn’t. Although I had the ideas in my mind: I would like to say this. I couldn’t say it. I felt I couldn’t, so I didn’t raise my hand.” (student 14, fail)

“In my secondary school, the teachers didn’t use to correct my pronunciation. Since I came to the university, my pronunciation had been corrected, so I was afraid I couldn’t pronounce well. Then, I felt, um, not very good, so I was not willing (to talk). It was because I felt my pronunciation was inaccurate. I felt

somewhat unconfident, so I didn't want to talk." (student 9, fail)

A few students, particularly the students with lower proficiency, noted that the order of the activities within a lesson also influenced their task-confidence. They explained that unsatisfactory performance in an activity at the earlier stage of a lesson might lead to a lack of state WTC in later activities. According to these students, they felt defeated, and thus were afraid of making more mistakes and being negatively evaluated in the future. However, some students noted that the negative impacts of the unsatisfactory performance on their task-confidence did not last forever, as they gradually forgot the negative experiences while engaging in other activities.

"Maybe because I didn't perform well in the quiz (before the communication activity), I felt I was influenced. I had been thinking about the quiz. The teacher... well... so I said I... felt upset in her class... She must have a bad impression of me. I neither actively participated in the activity nor performed well in the quiz. It reminded me of the... final exam. After a while, I got better. I forgot it while listening to the class." (student 14, fail)

These responses are in line with the descriptive statistics, which indicate that this sample of students tend to be extremely unconfident in their English proficiency. Like the current study, a few other studies (e.g. Riasati, 2012; Eddy-U, 2015) have suggested that a lack of confidence is derived from a fear of making mistakes. It is believed that students would be more confident and thus become more willing to communicate if they did not need to worry about the accuracy of their L2 production (Liu, 2002; Riasati, 2012).

Lack of capability. Most of the students reported that they were unwilling to participate in some communication activities, as the tasks were so difficult that made them feel frustrated. Unlike the lack of confidence explained above, which seems to be a result of lacking of English proficiency, particularly in terms of oral communication competence, the perception of incapability was mainly because the tasks were beyond student capability. In these situations, the students were not confined by their English proficiency, but defeated by the difficulty of the tasks.

"At the end of the class, when the teacher asked us to write (personal interpretations of the novel), I was really... I did think that I was not very... good at... not good at reading comprehension. It was so difficult for me to analyse the

thoughts and emotions (conveyed by the texts). I couldn't even comprehend it in Chinese." (student 28, distinction)

In some communication activities, students were supposed to interact with their peers. However, it was reported that some of the peers' oral English was poor, so nobody but English teachers could understand what they were talking about. Thus, some students reported that they were not capable to interact with those who had relatively low English proficiency, so they waited for the teachers to interpret for them.

"Except for some students whose oral English was relatively good, others were hard to understand. Only the teacher could understand. So, when it wasn't my turn, I was absent-minded. Because they didn't talk clearly and sensibly, I felt I couldn't understand them." (student 26, pass)

Similarly, Cao (2011) has also reported that difficulty in comprehension tends to frustrate L2 learners and reduce their state WTC. However, Cao (2011) attributed comprehension problems to a lack of L2 proficiency in understanding discourse (e.g. limited vocabulary). However, results of this study suggest that comprehension problems may also be a consequence of low comprehension capability and interlocutors' poor oral English production (e.g. incorrect pronunciation).

Lack of knowledge. Nearly half of the students reported that when they were asked to talk about unfamiliar topics, they lacked ideas or knowledge and felt reluctant to communicate. At the same time, some students were aware of the difficulty for teachers to choose topics that matched everyone's background knowledge, as different individuals might be interested in different topics.

"In today's news report, we were asked to talk about the affair between Tang Yan and Luo Jin (two pop icons in China). I... really didn't know how to answer it, so none of us stood up. We kept thinking for a long time. Um... I hadn't watched the teleplay. I had no idea, so I could do nothing but sit there." (student 8, distinction)

"The people or things the teacher was talking about didn't match my background knowledge, so I kept silent. Um, well, I know it's difficult (to match everyone's background knowledge), because everyone's interest is different." (student 19, pass)

These responses support previous findings on topic familiarity, in that background knowledge could give rise to positive feelings in that specific situation and thus lead to higher state WTC (Kang, 2005).

Inattentiveness. Another commonly reported negative perception that hindered willingness to participate in communication activities was inattentiveness. Several students explained that, when they thought a lesson was neither efficient nor effective, they were easily distracted or absent-minded before or during communication activities, which resulted in low state WTC. They tended to attribute inattentiveness to their teachers' unreasonable time management and/or unconvincing teaching approach.

“The teacher spent much time on it (a communication activity), but some of the time was wasted. I mean, she didn't make full use of the time. There was a long period of silence. Then, um... my mind was wandering.” (student 12, fail)

“There was a text, followed by some questions. I thought I would never understand the point of this kind of teaching. I didn't feel I benefited and it was quite boring, so... I looked forward to finishing the class as soon as possible. Sometimes, the teacher looked at me, so I took some notes. If she hadn't looked at me, I wouldn't have taken any, as I thought the class was useless.” (student 13, fail)

Some students reported that the time when activities took place influenced their state WTC. For example, many students reported that they felt less attentive at the end of each lesson, especially when it was about lunch time. At that time, some students were thinking about what to eat at lunch, some were worrying about the crowd in cafeterias, and some were sending messages to their friends. These students admitted that, although they were still listening, their attention was distracted, which made them less active in participating in the communication activities.

“It was near the end of the class, so I wanted to leave as soon as possible. I didn't pay much attention to what the teacher was talking about. We should finish the class at 11:30, but at that time... the discussion on that topic just began. Well... I wanted to leave on time, so I didn't think much (about the topic).” (student 13, fail)

Student attentiveness was also influenced by personal states and affect at specific points

in time. For instance, whether they slept well the previous night, whether they were experiencing romantic affairs, or whether they were busy with extracurricular activities. The personal affairs listed above seem irrelevant to English learning; however, they did distract some students from attentively and actively participating in the communication activities in the classroom.

“I was not too bad during the speech imitation, but then... I became absent-minded gradually. I stayed up late yesterday. Because of some unexpected events, I didn't sleep well.” (student 28, distinction)

The above findings suggest that L2 learners' state WTC could be influenced by their personality states at a particular point in time. These results are in line with Mystkowska-Wiertelak's (2016) findings. It was also indicated that overlong waiting time could lead to inattentiveness and reduced state WTC, which contradicts Zarrinabadi's (2014) claim that extended waiting time could bring about higher state WTC.

Nervousness. Several students reported they felt nervous talking in public in the presence of many people, especially when they were not familiar with them. Hence, these students' state WTC in whole-class communication activities was lower compared to their state WTC in group activities. However, some students believed that their state WTC in whole-class communication activities would increase gradually over time, as they would become increasingly familiar with their teachers and classmates.

“Actually, I was shy, very shy. Well... I wasn't at ease when talking in front of a lot of people. I think I will become better next semester because we have been getting familiar. Whatever I say, I won't feel very uncomfortable. I was willing to be active, but... don't know why suddenly I became nervous. In fact, I am a rather active person, and I would like to be active. Because I... felt nervous, I did nothing. I also felt... not good.” (student 1, merit)

Some students were generally interested in participating in communication activities in class; however, they felt extremely nervous after listening to the talks delivered by one or two individuals whose oral English was excellent. The students who gave this response appeared to be those whose English proficiency was good but not the best in the class. These intermediate students were extremely unwilling to talk after listening to more advanced students, as they worried about their performance and felt afraid of being

compared with their more proficient peers.

“XYZ (the student whose oral English was the best in the class) went first. When I saw she stood up, I felt really nervous. I was no longer willing to stand up. If I talked immediately after she did, my inferiority would be obvious.” (student 30, merit)

“When I found someone, such as XYZ and the like who were excellent (in oral English), were asked to speak and did speak very fluently, I felt ashamed to speak, though I had something to say. Because I wouldn’t even say a single sentence smoothly while standing up, it would be embarrassing.” (student 24, merit)

Student nervousness was also influenced by types of activities. Almost all said that they felt less nervous when giving talks or presentations, as they had been given sufficient time to prepare and practice in advance; however, they felt more nervous in face-to-face conversations, particularly with their teachers or in their presence. Some students noted that when their performance in activities were assessed or compared, their feeling of nervousness was likely to increase. When doing competitive activities, they not only thought about the information they would like to impart, but also worried about how their performance would be evaluated by their teachers and peers.

“Um... I preferred those activities in which the teacher provided me with plenty of time for preparation, such as news reports and speech imitations. However, when she asked the questions in today’s class, especially those add-on questions, I felt nervous, so... I didn’t stand up until the very end of the activity. I had an answer in my mind, but it was difficult for me to say it. So, I never talked when I didn’t have to do so.” (student 25, pass)

“Because the task was to some extent competitive and we competed for the best, I hoped I could perform well. I didn’t like that kind of activity with too much competition or assessment. Talking in itself wouldn’t make me feel under such pressure.” (student 26, pass)

Additionally, how activities in a lesson were ordered seemed to make an impact on student state WTC. For example, when the students knew that there would be some important or stressful activities (e.g. a formal presentation or a quiz) later in a lesson, their state WTC in previous activities seemed to be reduced. Students reported that, during earlier

activities, they were worrying about their performance later and could not help but prepare for the presentation or quiz that was going to take place.

“I... um... was wondering when we would take the quiz. I wanted to finish it (the communication activity) as soon as possible because I was worrying (about the quiz). To get a higher mark, sometimes I glanced at the vocabulary list.” (student 2, pass)

“During the activity, sometimes I suddenly thought about the sentences I would like to say later (in the presentation at the end of the lesson) and suddenly could not come up with the words. Then I looked up the words and... felt nervous.” (student 26, pass)

The responses presented above are like those reported in Kang's (2005) security dimension. Like this study, Kang (2005) has suggested that the feeling of security could be influenced by situation cues such as the number of interlocutors, familiarity with interlocutors, interlocutors' L2 proficiency, and previous performance in the communication.

Embarrassment. Nearly half of the students stated that in communication activities when none of their classmates spoke or showed any intention to speak, they felt embarrassed. These students reported that, when the classroom was silent, they tended to hold their breath and looked down to avoid eye contact with their teachers. They felt that whoever suddenly broke the silence would attract attention, which was embarrassing. Hence, they were not willing to stand out and break the silence.

“Um... I felt others didn't want to talk, so I didn't want to talk either. Well, when others were rather silent, maybe I was also influenced by the atmosphere. I became silent too. The class was in silence... (How could) one suddenly raise a hand? It was embarrassing.” (student 2, pass)

“I think, for some of the questions asked by the teacher, all of us knew the answers, but we didn't speak up. Um, if a person had initiated (an answer), others would have followed him (or her). That's it. It must be a problem of the atmosphere. I felt that everyone could answer (the teacher's questions). I was willing to talk, but... Um... even if you looked up, the teacher would look at you. It was extremely embarrassing. Because others were all looking down, if I had looked up or moved,

the teacher would have gazed at me.” (student 9, fail)

A few students reported that some activities made them feel extremely embarrassed, because they were not sure what they were supposed to do and why they were asked to do it. For example, a teacher asked each group to choose a representative to give a report on a group project. Some students reported that they felt embarrassed because they were not chosen as the presenter and thus were not able to contribute. In some other communication activities, additional skills (e.g. singing and painting) were required, which might embarrass students who were not good at these and thus reduced their state WTC. Some students argued that these additional skills were not relevant to language learning, so they questioned why they should spend time on these tasks and, at the same time, lose face during the activities.

“In part 3, the teacher asked one person to present the results of the discussion among four of us, which made me feel uncomfortable. Only one person! So, let her (the presenter of the group) talk. I decided not to talk, although I had many ideas to share. I mean... the rule was only one person could present. If that’s the case, the remaining ones had nothing to do.” (student 19, pass)

“Um... how to say, I mean, the various activities that the teacher arranged for us, such as painting and singing a rap, were beyond my capability, to some extent. Um... talking about English learning, I don’t think I can learn much from these activities.” (student 18, merit)

“I felt quite embarrassed. I thought, ‘poor me, need to sing a rap in the class!’ Well, the teacher held a wide range of activities in the class, but, sometimes, the activities were outside my expectations. Actually, this activity was beyond our imagination, out of expectation. To some extent, her class made me wonder what I was supposed to learn.” (student 16, fail)

The results show that state WTC in L2 classrooms can be influenced by classroom atmosphere, which is in line with previous research (e.g. Eddy-U, 2015; Peng, 2012; Riasati, 2012). A participant in Peng’s (2012) study commented “Maybe it was a little funny, but too ridiculous. It just provided us with fun” (p. 208). Similarly, this study also suggests that students tend to be embarrassed and unwilling to communicate in ‘funny but ridiculous’ activities.

Fear: According to the students, state WTC in the classroom was not only influenced by peer participation, but also by teacher comments on and reactions to their participation. A few students reported that the teacher seemed to be frustrated when no one interacted with her, which led to an extremely stressful classroom atmosphere. They said that they were reluctant to speak when the teacher looked unhappy, because they were not sure whether they could meet the teacher's expectations and were afraid of being criticised.

“Well, when the teacher became serious, we felt... um, how to say that, a little bit scared and didn't dare to talk. Then, the classroom atmosphere became quieter, which made her even more serious and upset. Then we didn't dare to say anything.” (student 7, fail)

“For example, the teacher asked, ‘what does it mean?’ No one spoke. Not a single person responded. Then she tried to let us talk, but we still kept silent. Then she became angry and sometimes a little bit... harsh. For example, (the teacher said) ‘You can't behave like this.’ Then we became extremely silent, throughout the class. Um, it was very depressing.” (student 4, distinction)

“Sometimes when the teacher commented on you, she never considered your feelings at all. (Under such circumstances,) I no longer wanted to take part in the class. I was willing to say something and break the ice, but I was afraid that she would continue criticising.” (student 20, missing)

In line with the current study, Kang (2005) has suggested security (defined as an antonym of fear) as an important psychological antecedent of state WTC that seemed to be influenced by interlocutors' verbal and non-verbal responses, such as teacher smiles and feedback. Similarly, other studies (e.g. Lee, 2009; Riasati, 2012) have noted the impact of the teacher on classroom atmosphere and student state WTC; however, these studies focused on situation cues related to the teacher (e.g. teacher behaviour, attitudes and teaching styles) instead of subjective situation characteristics reported in the current study (e.g. students' perceptions of teacher attitudes and feedback).

7.3.2.3 Duty

Sometimes students did not have any positive feeling or even negatively perceived the learning situation; however, they were still eager to communicate because they felt they had a duty to communicate in the classroom. These perceptions of having to communicate

in classroom activities were categorised as responsibility/duty. The feeling of responsibility is neither positive nor negative but can stimulate state WTC. One might feel a responsibility to meet teacher expectations, to learn from class learning activities, or to support group members.

Responsibility towards teachers. For most of the students who felt a duty to participate in the communication activities, their rapport with the teachers was the main source of this. These students believed that their teachers expected them to volunteer responses. Hence, to please the teachers or give a good impression, they felt that they had to communicate to meet teacher requirements and expectations.

“When the teacher was happy, we felt very happy. Um, to make her happy, well, we did everything we could to make her happy, like... voluntarily answering the questions to make the atmosphere more active and doing whatever she asked us to do.” (student 7, fail)

“I was afraid that the teacher would give me a low score, so I sometimes had to stand up. I stood up and responded to her mainly because I wanted to give her a good impression.” (student 20, missing)

However, a small number of students doubted whether teachers always expected responses when posing questions. A student claimed that the teacher performed as if the class was student-centred because it was required by the curriculum; however, in reality the teacher preferred dominating the class. As a result, the student wondered whether actively communicating in the classroom was a good way to please the teacher, if the teacher did not expect any responses.

“Seriously? I thought the teacher quite enjoyed talking alone. In my view, in fact, she looked quite happy when talking alone. She merely wanted to go through those procedures (questions and answers with the students).” (student 15, pass)

These responses generally correspond to Bernales’s (2016) claim that some student participation in communication activities seems to be affected by teacher expectations, even though these may not match students’ own desires.

Responsibility to learn. Nearly half the students reported that the feeling of responsibility was partly caused by their determination to learn from class communication tasks. They

suggested that the feeling of having to communicate was influenced by learning materials. Although it was presented earlier that most students perceived the textbook as less interesting than other learning materials, they tended to believe that the textbook contained knowledge that was important. Hence, they believed they had to learn from the textbook, and decided to participate in the related communication activities even though the learning situation did not trigger any positive perception.

“The textbook-related tasks were boring, um, but I still... um, had to learn attentively and take lots of notes. Um... anyway, I think it (the textbook) was less interesting than the novel. But I must study attentively. Yes, because the teacher was teaching us some important vocabulary and phrases. I thought they must be very important, so I had to participate.” (student 3, distinction)

“Although we said we preferred talking about the text, we just wanted to gain knowledge. We didn’t mean to say that we were happy learning from it. I think nobody was happy when talking about it, but we knew we had to learn.” (student 4, distinction)

“When we were learning knowledge (from the textbook), I felt... Even though I was neither interested nor pleased, I could still do it peacefully and attentively.” (student 7, fail)

Similarly, both Bernales (2016) and Kang (2005) have suggested that L2 learners’ feeling of being responsible to communicate in classroom activities could be related to their personal L2 learning goals and motives.

Responsivity towards group members. For students who worked collaboratively in fixed groups throughout the semester, a feeling of responsibility could also be triggered by their group members. One might initiate a talk in a communication activity when seeing that other members in the group did not intend to speak. It was reported previously that when the classroom atmosphere was extremely quiet, most students tended to feel embarrassed and thus were not willing to break the silence; however, when working collaboratively in groups, they tended to feel that everyone had a duty to break the silence so to avoid causing embarrassment to other members. In other words, students might have some negative perceptions (e.g. embarrassment) that could be potential obstacles to state WTC; however, a feeling of responsibility could help them to overcome these.

“Actually, I think, well... The teacher threw a question and nominated our group (to answer that question). If everyone in our group had kept silent, it would have been embarrassing. Then... actually, I had no idea what I ought to say, but I stood up and said something. I was not sure whether it was correct or not.” (student 3, distinction)

“Because the teacher asked for volunteers, we might... people like myself would not stand up. But if we had been divided into (fixed) groups, like those in the other class, it would have seemed like talking in turns... like sort of compulsory... but not exactly... I mean... it would have made everyone take a turn.” (student 5, merit)

These results generally support Kang’s (2005) and Eddy-U’s (2015) claim that L2 learners’ willingness to communicate, as well as a related feeling of responsibility, might be results of their desires to achieve personal goals and maintain good interpersonal relationships with interlocutors.

7.3.2.4 WTC and communication behaviour

The statistical analysis showed inconsistency between state WTC and communication behaviour, as the mean of communication behaviour (reported either by teachers or students) was lower than the mean of state WTC. Similarly, some responses during group interviews implied this inconsistency. Students suggested that, in some rare situations, they might communicate even without readiness or intention to enter into discourse; however, in most situations, they had intention to communicate but did not show actual communication behaviour. In this section, I will explain how and why this inconsistency between state WTC and communication behaviour arises from student perspectives.

Nominated talking. A few students indicated that some of their responses to the teachers were not voluntary but nominated. By contrast to voluntary speeches, which were highly dependent on student communication intention, nominated speeches were not the results of student desires but teacher-required. Thus, it might be the case that a student was reluctant to answer a question asked by a teacher because of some negative perceptions illustrated above; however, the student was nominated by the teacher, and thus responded reluctantly. In other words, the student did communicate, even without any willingness to communicate in that specific situation. Some students noted that, when they were not ready to talk but were asked to do so, their originally negative perceptions of the learning

situation tended to be aggravated.

“Yes! I was highly unwilling to talk! I didn’t even know what I was saying. What I felt was ashamed. I really felt ashamed. If I had been asked to do something that I was good at, I would have been willing to do so. But I had to do a task that I was not able to complete and had to do it in public. Gosh, I felt so ashamed!” (student 15, pass)

However, some students argued that if one was not willing to talk, then one could keep silent even when nominated. In other words, the teachers had never forced any student to talk in the classroom. Whenever one did communicate, be it voluntary or nominated, one must have some extent of state WTC.

“No, I didn’t say anything (when nominated by the teacher). I just waited for the teacher to nominate someone else. I felt unable to say anything, that’s why I didn’t raise my hand.” (student 14, fail)

“For the translation discussed in this class, I did it rather attentively. Maybe because I was confident, when the teacher asked me to share, I felt I would be fine sharing with others. Otherwise, I would have refused determinedly.” (student 18, merit)

Opportunity to talk. It was often the case that students desired to communicate but were not provided with adequate opportunities to speak at that moment. To conduct any communication behaviour, one needs not only a desire to communicate (i.e. state WTC), but also an immediate opportunity to communicate. The communication opportunities in classrooms tended to be managed by teachers. Although a large number of students believed that the teachers generally expected them to communicate in the classroom and offered sufficient opportunities, some did report that it was a shame they had no chance to express their thoughts and feelings in some activities.

“After I raised my hand, the teacher called another student, called XYZ. Then I... planned to talk after listening to XYZ. Because XYZ was talking, I couldn’t interrupt her. When XYZ finished her talk, I thought it was my turn to talk. But the teacher gave some concluding remarks. Then... she talked about the textbook. I didn’t get the chance to talk. What a pity!” (student 6, merit)

“Not sure why the teacher suddenly, I mean, began to interrupt me today. Because I... spoke English as a second language, sometimes I paused for a while before responding, searching for the words in my mind. Then, when I was almost ready, she said the answer for me! I felt quite upset. Just as if you are about to find out who is the murderer, but suddenly someone tells you that this person is the murderer.” (student 20, missing)

Necessity to speak. Despite the perception that they were expected to communicate in class, many students expressed confusion as to why they were supposed to talk in some activities, such as responding to teacher questions or reporting to the class after a group discussion. They claimed that the process of thinking and discussing with peers was more important than presenting the results of this. Some students explained that they had rehearsed their answers or speeches in their minds several times, in case they were nominated by the teachers. As previously presented (section 1.1), WTC is defined as readiness to enter into discourse. From this perspective, these students had state WTC in those activities because they thought actively, rehearsed in their minds, and were ready to speak; however, they did not volunteer any communication behaviour unless nominated. Some students argued that they did communicate during the communication activities, even though they did not talk with anyone. They believed that thinking and drafting speeches in mind was another effective way of communication, even though it could not be observed by others.

“It didn’t matter, didn’t matter whether I stood up to answer it or not. I had thought about the teacher’s questions in my mind. So... in fact, I had communicated in my mind. At that time, I was extremely... interested, so the report of our group was drafted by myself. Then... I felt my willingness to communicate reached the peak at that time.” (student 12, fail)

“In fact, I thought the report at the end was nothing. What was important was not the final report, but the process of discussing. I thought we had discussed very well (outside the classroom). That’s enough. The process of discussing, rather than the final presentation, was the important part. It didn’t matter who presented in the classroom.” (student 20, missing)

“It’s okay to let the teacher nominate. I didn’t care, because I had prepared. I didn’t raise my hand, but if she had chosen me, I could have done it. Um... maybe

because I didn't realise why I ought to raise my hand." (student 8, distinction)

Other than the perception that speaking is not necessary for language learning, a few students expressed concerns that their speeches would not be beneficial to other students. One claim was that one's questions or opinions were one's own affairs that would not be useful for others and thus should be resolved personally after class. Another claim was that some thoughts were like those presented by others and were not good enough to be repeated. Hence, these students believed that, in either situation, speaking in class would be a waste of others' time. Therefore, they did not actually communicate whether they had state WTC or not.

"My interaction with the teacher was only the interaction between the teacher and myself. Sometimes my questions were not very useful for others, so I waited until the teacher finished her teaching. Um, I think... I preferred talking with the teacher after class. But, I think, I wouldn't mind if others initiated any questions." (student 13, fail)

"When they were talking about their supporting opinions, actually I... was managing my language in my mind as well. But... I found they talked a lot, and some of their opinions were similar to mine. Um, finally, I didn't talk. Um, my opinion was... the same (as what had been presented), and my examples or, let's say, my language was also common, nothing new. The ideas had been said by others, so I thought it was unnecessary to repeat." (student 7, fail)

In accordance with the results of the statistical analysis, these results suggest that not all intention to communicate could be successfully transformed into actual communication behaviour. However, as some responses have indicated that students might communicate reluctantly when nominated by teachers, we cannot overlook the possibility that communication behaviour might be generated without the presence of state WTC in some situations.

7.3.2.5 Teacher and student criteria for language performance

Teacher assessment. To compare with student responses, reports from the two course teachers were collected. The teachers subjectively assessed their students' language performance in specific lessons and activities and stated their criteria for these assessments. It seems that the assessments were mainly based on quantity and quality of

language production.

Both teachers reported that their priority was voluntary participation in communication activities. It seems that they paid more attention to teacher-student rather than student-student interaction when assessing language performance. One teacher clearly distinguished between voluntary responses, stimulated responses, and non-public responses in teacher-student interaction. According to this teacher, voluntary responses referred to interaction without intervention from the teacher and were the best form of teacher-student interaction. In some situations, although the students did not actively participate in teacher-student interaction at first, they initiated responses after being encouraged by the teacher. The teacher referred to these as stimulated responses and believed they were satisfactory. Additionally, the teacher noted that there were some students who responded non-publicly by whispering or slightly moving their lips. The teacher reported that these non-public responses were better than not participating at all, although they might not be actual communication behaviour because nobody could receive the messages.

“For example, I asked a question, and waited for volunteers to answer it. There were some students who never cared whatever you asked, while some others might respond silently because I could see their mouths were moving. These two students... raised their hands after I encouraged them to do so. As for XYZ (another student in the class), whenever I asked a question, he/she always responded, even without any encouragement.” (teacher 2)

Both teachers reported that they might also take the quality of language production into account when assessing student performance, including pronunciation, communicative competence, as well as general English proficiency.

“XYZ’s English productive skills were poor, no matter speaking, pronunciation or writing. Both his/her written and spoken outputs were terrible, and his/her pronunciation was difficult to understand.” (teacher 1)

“Generally speaking, he/she was indeed... the one whose English proficiency was the lowest in this class. Um, his/her English proficiency was the lowest; however, he/she was extremely eager and willing to improve. So, I think he/she... was quite active.” (teacher 2)

The teachers noted that their assessments were also affected by impressions and expectations of individual students. They tended to compare one's performance in a specific class or activity to his or her general English proficiency and past performance. For example, they valued progress made by students who had lower English proficiency or performed less actively in previous lessons; however, they were not easily satisfied by more competent learners, as high expectations towards these students were hard to meet.

"XYZ was relatively active. As you might observe, when I asked for volunteers, he/she was always the one representing her group. Although his/her English proficiency was not very high, he/she made great progress. He/she was much much better than before." (teacher 2)

"Generally speaking, he/she was not bad, but he/she did not perform well today. I felt he/she was a bit inattentive and not in the mood. Maybe because his/her English is relatively good, I hope he/she can be better. I might have higher expectations of him/her." (teacher 1)

One of the teachers mentioned student attentiveness; however, the teacher reported that she only paid attention to this when the students did not actively participate. According to the teacher, when students did not produce any communication behaviour during an activity, she observed whether they were listening attentively through non-verbal cues, such as facial expressions and eye contact.

"In the classroom, he/she seldom made eye contact with me and always looked down. Sometimes I wondered whether he/she was listening, and what he/she was thinking. Maybe he/she lacked confidence to actively share, talk or make eye contact with teachers." (teacher 1)

Self-assessment. For most students, attentiveness was one of the most important criteria for good performance. From their perspective, whether or not they talked, as long as they listened and thought attentively, their performance could be satisfactory. To note, some students did realise that their attentive (not necessarily active) performance might be negatively assessed by teachers or other observers.

"On the one hand, I was inactive. On the other hand, I felt not too bad. Because I... wasn't distracted... and was quite attentive. I was always listening attentively whatever the teacher taught. Maybe I looked bad from the outside, but not too

bad from the inside.” (student 8, distinction)

Participation in communication activities was another major criterion when the students assessed their performance. Compared to the teachers, who paid more attention to teacher-student interaction, some students noted that student-student interaction could be more important than interaction with teachers. Moreover, rather than mentioning non-public responses as the teachers did, the students took nominated speeches into account.

“I participated quite a lot during the group discussion. Not only did the communication with the teacher matter, but also the communication during group activities.” (student 25, pass)

“My (only) criterion for assessment is responding to the teacher voluntarily... or even reluctantly.” (student 21, merit)

Interestingly, most students stated that whether they had been well-prepared before attending a class was a crucial criterion of their self-assessment. They explained that their preparation (e.g. completing teacher-assigned homework, previewing texts, and looking up new words) helped them understand and learn better in class. Hence, self-assessed performance in communication activities seemed to be influenced by performance in preparation work beyond the classroom.

“A very important criterion is... whether I had finished the preview tasks assigned by the teacher. I think, if I had finished the tasks before attending the class, I mean, if I had already thought and summarised in advance, I would have felt totally different. In that situation, my learning outcome after the class would be totally different as well. So, the most important criterion is this one.” (student 23, pass)

Some students reported whether they made mistakes made a difference to their self-assessed performance. To some extent, they defined good class performance as communicating and learning as required by teachers without making mistakes or errors.

“I mean... what I did in the class. Did I do as the teacher required? Things like that. Um, I mean, I didn’t make any big mistakes.” (student 17, pass)

“I made errors. I made errors. Then, I didn’t talk fluently, and there were errors here and there. My criterion is not making any errors.” (student 22, pass)

A few students assessed their performance based on perceived progress or learning outcomes after doing the communication activities. Some thought they performed well, as they challenged themselves during the activities and realised their progress and problems after the activities, while others assessed themselves as not performing well because they gained nothing.

“I think I was not too bad, because, um... the news report was a kind of breakthrough to myself. I did it on the stage without notes! Then... when talking about the text, I found, in fact, I was short of knowledge. Also, my oral English wasn't perfect, far from perfect, and my vocabulary... My criterion is... my progress compared to myself in the past. As for the vocabulary, I didn't... have much progress, but at least I realised my shortcomings. Then, I could move forward step by step in the future.” (student 1, merit)

“Um... after XYZ's (a student from the same group) talk, the activity finished. I didn't learn anything. Then, when the teacher was teaching, I felt a little bit absent-minded. So, I don't think I performed well. Um... neither my interaction with the teacher nor my perceived learning outcome was good.” (student 23, pass)

Another criterion of assessing performance reported by a small number of students was their emotional feelings. For example, whether they felt happy or engaged when doing a communication activity. A student explained that positive feeling in an activity might indicate that one worked attentively and effectively.

“How often I looked at my watch indicated (the extent to which) I hoped the teacher finished the class as soon as possible. When I didn't want the teacher to finish the class quickly, I felt I performed well.” (student 20, missing)

In general, both the self-assessments and teacher assessments were based on student participation in communication activities, quality of communication, and progress compared to past performance. However, it seems that students tend to stress their subjective thoughts and feelings (e.g. attentiveness, perceived progress, and positive affect), whereas teachers mainly focus on observable cues (e.g. actual participation and quality of language production). These are consistent with the results of the statistical analysis that self-assessed language performance significantly correlated with trait and state WTC, whereas teacher-assessed performance significantly correlated with student

actual communication behaviour and language exam scores rather than trait and state WTC.

7.4 Summary and Conclusion

This chapter has presented the research design and results of Study 2. This study investigated a group of Chinese university students' trait and state WTC in English classrooms, and whether, how and why state WTC fluctuates within a relatively short period of time, i.e. across different communication activities during a lesson. Trait data were collected using baseline questionnaires, whilst state data were collected using momentary questionnaires and group interviews. By statistically analysing data collected through questionnaires, this study has found that state WTC fluctuates across different activities during a lesson, and this within-person fluctuation differs from one individual to another.

This study is in line with Pawlak et al.'s (2016) recent study in finding both within-person variability in state WTC during a lesson and individual differences in within-person variability. However, compared to Pawlak et al.'s (2016) study that measured state WTC 13 times during a lesson, the current study only had three to four measurement occasions in each lesson for fear of disturbing normal learning in the classroom. Hence, it was impossible to track the moment-to-moment fluctuations in state WTC, and this is a limitation of this study. One might argue that, if there were more measurement occasions, more fluctuations in state WTC might have been observed. Moreover, as very little data were collected at the within-person level, within-person correlations between state WTC and its related variables (e.g. language learning performance) could not be estimated. However, it needs to be noted that the statistical data collected through the questionnaires were only a small part of this study, and the focus was on student responses in group interviews.

Student responses in the group interviews have been thematically analysed, providing in-depth explanations of student communication generation processes in different learning situations. The current study has found inconsistency between state WTC and actual communication behaviour, which is in line with the finding of Study 1. Some communication intention (i.e. state WTC) could not be translated into actual behaviour because of lacking communication opportunities or motivation. However, in situations when students were not willing to communicate, they might do so to meet teacher

expectations.

Different situation characteristics suggested by students as either triggering or hindering state WTC and communication behaviour have been categorised into three major dimensions: positivity, negativity, and duty. In addition, differences have been observed in student responses, indicating potential individual differences in communication generation processes. These findings are generally consistent with the systematic review, which categorised the situational antecedents of state WTC reported in literature into the above three major dimensions, and the findings of Study 1, which found significant correlations between state WTC and a set of situation characteristics (e.g. task-interest, task-usefulness, support) as well as the three underlying dimensions. However, unlike the result in Study 1, which showed no correlation between state WTC and task-difficulty, a number of students in the current study reported that their state WTC was influenced by task-difficulty, and this is consistent with the findings reported by de Saint Léger and Storch (2009) and Eddy-U (2015).

One may argue that what participants report could be different from what they think, particularly when they are interviewed in groups. Students may have concerns about the confidentiality of their responses, withhold their true thoughts and feelings and falsify responses (Arksey & Knight, 1999; King & Horrocks, 2010). However, the current study was only interested in thoughts and feelings related to English learning experiences, and the questions asked during the group interviews were not sensitive. As students had worked together during the lesson before each group interview, the group is believed to be a relatively secure situation for them to share their thoughts and feelings. Under these circumstances, group members could support each other in exploring their experiences, and discussions could stimulate a wider range and better expression of ideas than conducting one-to-one interviews (Cohen et al., 2011; King & Horrocks, 2010).

This study has found that language learning performance could be predicted by trait and state WTC; however, compared to results of the summative assessment (i.e. the end-of-term paper-based exam), results of the formative assessment (i.e. the end-of-term teacher judgment of participation and language performance) were more strongly related to trait WTC. This result is in line with Naderifar and Esfandiari's (2016) and Yashima's (2002) findings and supports Savignon's (2005) claim that WTC seems to better predict oral language performance than written language performance. Like Study 1, this study found inconsistency between student self-assessments and teacher assessments. It has been

found that students tend to consider their subjective thoughts and feelings (e.g. state WTC and affect) when assessing language performance, while teacher assessments seem to be based on observations of student behaviour and learning outcomes (e.g. the quality of language production, general English proficiency, etc.).

Another limitation of this study is that the very small sample size might have affected results. A sample size of 31 may not be big enough to detect statistically significant results. The larger the sample, the more likely to obtain significant results if they exist (Coolican, 2014). Hence, one may argue that if a bigger sample had been used, the results of correlational analyses might be different. Additionally, the interpretation of teacher perspective was based on responses from only two teachers, which is too small to make any generalisation. Future research may recruit larger numbers of students and teachers to test the results reported in this study.

8 Discussion

8.1 Introduction

This chapter will summarise the main findings across the two studies and discuss them in relation to the literature. I will firstly summarise my key findings and then refer to previous research that either support or contradict them. After discussing the results related to the three research questions, some unexpected outcomes will be discussed, including the inconsistency between teacher and student assessments, as well as the distinction between subjective communication intention (i.e. WTC) and actual communication behaviour. Afterwards, limitations of the research will be discussed, which will lead on to directions for future research. Lastly, conclusions will be drawn together with a set of theoretical, methodological and pedagogical implications for other researchers and practitioners.

8.2 Research Question 1: Variability

The first research question on variability in Chinese EFL learners' WTC in language learning classrooms, has been addressed at both trait and state levels. The results showed that: (a) there was between-person variability (i.e. individual differences) in trait and state WTC; (b) there was within-person variability in state WTC, that was nearly the amounts of variability observed at the between-person level; and (c) there were individual differences in within-person variability in state WTC. In other words, learners differed in their amounts of variability observed in state WTC. In the following sections, I will discuss between- and within-person variability in trait and state WTC as well as the individual differences in within-person variability in relation to existing findings in the literature.

8.2.1 Between-person variability in trait WTC

The samples in Study 1 and 2 showed different levels of trait WTC in English. The students in Study 1 were generally not so willing to communicate in English as they were in Mandarin, although the effect was small. This result is in line with the results reported by Chu (2008) and Liu and Jackson (2008), who both suggest that Chinese university students are generally less willing to communicate in English than in Mandarin. However, contradicting the above findings, the sample in Study 2 was quite willing to communicate in English, and showed a slightly higher level of trait WTC in English than in Mandarin

though the effect was also small.

As the samples in Study 1 and 2 were made up of first-year undergraduate students in the same university, the inconsistency in the results might, at least partially, be explained by students' different major subjects. In Study 1, as well as the two studies discussed above (Chu, 2008; Liu & Jackson, 2008), the participants were all non-English major students in Chinese universities; however, Study 2 was conducted with a group of English major students. The findings indicate that students who learn English as a major subject seem to be more eager to communicate in English than their non-English major counterparts. The reason might be that, compared to other EFL learners, those who decide to learn English as a major subject are more likely to have positive attitudes towards language learning and high motivation towards English communication. To my knowledge, no previous research has compared trait WTC between English major and non-English major students. However, the results of the current research correspond to Yashima and Zenuk-Nishide's (2008) finding that students who choose to study on more intensive L2 programmes (i.e. programmes with more L2 exposure) tend to be those who stand out from their peers in terms of L2 learning motivation, L2 WTC, and L2 proficiency. Similarly, English major programmes in Chinese universities can provide students with higher levels of English exposure than the College English courses for non-English major students. However, as the current research was conducted in only one university and the sample sizes were small (particularly for Study 2), the result cannot be generalised to students in other contexts. Similar studies may be conducted to compare trait WTC of major and non-major students in other L2 learning contexts.

The different levels of trait WTC revealed by the two studies might also be attributed to gender differences, as the non-English major students in Study 1 were mostly male whereas the English major students in Study 2 were mostly female. The sample in Study 2 not only scored higher on the trait WTC in English scale, but also showed a slightly higher level of trait WTC in Mandarin than the sample in Study 1. Hence, one might claim that females are generally more talkative than males, and this distinction becomes salient in the L2 learning context. Several large-scale personality surveys conducted in different cultural backgrounds have found that females are generally more extrovert or talkative than males (e.g. Costa, Terracciano & McCrae, 2001; Feingold, 1994; Lippa, 2010; Weisberg, DeYoung & Hirsh, 2011). Although relevant findings in L2 WTC literature have been inconsistent, most of them are in line with the results of the current research,

showing small gender differences favouring females (e.g. Baker & MacIntyre, 2000; Li, 2012; MacIntyre et al., 2002). As issues relevant to gender differences are not the focus of the current research, they have not been further analysed. Future research may test gender differences in trait and state WTC across different L2 learning contexts.

Within either group (i.e. English major or non-English major) there were individual differences in both trait and state WTC. These results support previous research on trait WTC (e.g. Liu, 2002; Oz, 2014) in identifying individual differences within specific culture backgrounds or L2 learning contexts. Other than showing the existence of individual differences as previous research did, this research has moved further by using standard deviations to quantify individual differences (and within-person variability, see the next section) to estimate the extent to which individual students differed in trait and state WTC. The results indicate that there are considerable amounts of individual differences in both trait and state WTC, accounting for about half of the total amount of variability in WTC. As this research was carried out, Pawlak et al. (2016) published a paper that used standard deviations to estimate the amounts of individual differences in state WTC over a 60-minute-period. They compared four groups of learners and found that some groups varied more than the others; however, in general, Pawlak et al. (2016) found large amounts of individual differences in L2 learners' state WTC, like the current findings.

8.2.2 Within-person variability in state WTC

When looking at the aggregate level (i.e. regarding a sample as a whole), both the non-English major sample in Study 1 and the English major sample in Study 2 seemed relatively stable. It seemed that state WTC did not fluctuate much either across different activities during a lesson (Study 2) or across different lessons during a semester (Study 1). Study 1 showed a slightly decreasing trend in state WTC over a semester, which corresponds to the result reported by Mystkowska-Wiertelak (2016), the only other study as far as I know using a high-density repeated measurement approach to study within-person variability in state WTC over a relatively long period of time (i.e. a semester). As suggested by Mystkowska-Wiertelak (2016), the generally decreasing trend in state WTC over a semester might be attributed to student tiredness after a term's work, the tension caused by end-of-term exams, and/or a feeling of boredom with repeated measurements. For within-person variability in state WTC during a lesson, it is hard to suggest any trend because state WTC was only measured three or four times in Study 2. With more frequent

measurements, Mystkowska-Wiertelak (2016) reported an initially increasing and then decreasing trend with a very small number of participants (i.e. 12). However, Pawlak et al. (2016) and Mystkowska-Wiertelak and Pawlak (2017) did not find any uniform trend in the change of state WTC during a lesson.

Although state WTC seemed to be relatively stable at the aggregate level, individual trajectories showed more remarkable within-person variability and no individual's trajectory was exactly the same as the aggregated trajectory. It seems that one could fluctuate from extremely unwilling to extremely willing to communicate both across different communication activities during a lesson and across different lessons during a semester. The amounts of within-person variability during a lesson and during a semester were comparable and were as much as the amounts of between-person variability (i.e. individual differences) in trait and state WTC. The results of Study 2 are line within those of Pawlak et al.'s (2016) study, which also showed within-person variability in state WTC at both aggregate and individual levels. However, to my knowledge, individual state WTC trajectories over a semester have not been studied previously, as the only long-term study (i.e. Mystkowska-Wiertelak, 2016) that used a high-density repeated measurement design to investigate within-person variability in state WTC over a semester only looked at the aggregated rather than individual trajectories.

8.2.3 Individual differences in within-person variability

By comparing between individuals, it was found that the fluctuations in individual trajectories were not consistent and no clear pattern could be detected. This inconsistency could be evidence for individual differences in within-person variability in state WTC as some individuals might vary more frequently and sharply than others and might respond differently towards the same situation.

Among the handful of studies interested in within-person variability in state WTC, few paid attention to individual differences, except for the small-scale studies conducted by Cao (2013) and Pawlak et al. (2016), in which individual state WTC trajectories were compared. Like Study 1, a small-scale study by Cao (2013) found individual differences in within-person variability in state WTC over a five-month period, as three state WTC trajectories showed gradually increasing trends while the other four fluctuated. However, as Study 1 recruited a much larger sample and collected state data more frequently than Cao (2013) did, more fluctuating individual trajectories were found, and it was hard to

identify any underlying pattern across individuals.

Study 2 shows that individual state WTC also fluctuated widely during a lesson without any observable pattern, which contradicts Pawlak et al.'s (2016) finding that most individual state WTC trajectories within a lesson showed very similar patterns (e.g. students tended to be more willing to communicate at the onset of new activities and less willing to communicate towards the end of the lesson). The reason might be that Pawlak et al. (2016) measured state WTC thirteen times in a lesson, and thus could track the changes in state WTC and observe underlying patterns. By comparison, Study 2 only measured state WTC three to four times in a lesson, which made it impossible to observe the patterns reported by Pawlak et al. (2016). However, it should be noted that measuring state WTC thirteen times over a 60-minute-period might also have limitations. For example, it might disturb normal learning during the lesson and increase participant fatigue.

The results discussed in the above two sections support Schmitz's (2006) claim that only focusing on the aggregated trajectory could be misleading, as the aggregated trajectory may not be a good representative of individual trajectories and cannot reflect individual differences in within-person variability. It indicates that different amounts of variability may be observed when studying people as a group or as individuals. Hence, when dealing with data collected through repeated measurements, analyses would be better carried out at both aggregate and individual levels (de Vaus, 2001).

8.3 Research Question 2: Systematicity

As Study 1 and 2 showed large amounts of between- and within-person variability in trait and state WTC, the next step was to test whether this variability, particularly in state WTC, could be systematically explained by personal and situational antecedents. If the variability in state WTC was systematic, it deserved more attention rather than being simply seen as random error. To answer this, Study 1 used correlation and regression coefficients to estimate relationships between trait and state WTC and different personal and situational antecedents suggested in the literature, and Study 2 acquired student in-depth explanations of their communication generation processes in different learning situations. Study 1 found that state and trait WTC was systematically related to some personality traits, states and situation characteristics, which generally corresponds to the themes generated from the responses in Study 2. The following sections will discuss the

personal and situational antecedents of trait and state WTC, with a focus on the situational characteristics summarised in the framework proposed.

8.3.1 Personal antecedents of trait and state WTC

At the trait level, it was found that both L1 and L2 WTC was significantly influenced by personality, which generally supports the claim in the literature that personality has an enduring influence on WTC in both L1 (e.g. McCroskey & Richmond, 1990) and L2 literature (e.g. MacIntyre et al., 1998). However, unlike the findings reported by MacIntyre and Charos (1996) and Oz (2014), extroversion did not stand out from Big-Five personality traits in relationships with L2 WTC. The results of Study 1 and 2 were similar, showing that L1 WTC was strongly related to extroversion and moderately related to agreeableness, while L2 WTC was more related to openness to experience rather than extroversion or agreeableness. These results support Chu's (2008) claim that the relationship between extroversion and L2 WTC could be much weaker than that between extroversion and L1 WTC. Moreover, this research is in line with a recent large-scale study conducted by Piechurska-Kuciel (2018) with 534 secondary school students in Poland, suggesting that openness to experience can be regarded as a significant predictor of L2 WTC.

The findings suggest that compared to L1 communication, L2 willingness to communicate seems to depend less on talkativeness or kindness, and more on creativity and curiosity. The reason might be that L2 communication has more uncertainty and challenges than L1 communication, because L2 communicative competence varies widely and most people cannot speak a L2 as fluently and effectively as the L1 (MacIntyre et al., 1998). Hence, L2 WTC is less driven by eagerness to talk than attitude towards uncertainty. Talkative people are not necessarily willing to communicate in foreign languages because they may be conservative. In comparison, people who are curious about new ideas and experiences are more likely to be interested in the adventure of learning new languages and seek opportunities to communicate in them. It seems important to further investigate the distinction between L1 and L2 WTC; however, to my knowledge, not many studies have paid attention to this.

At the state level, Study 1 measured personality states and affect in different lessons and activities and analysed their relationships with state WTC. It was found that Big-Five personality states and affect changed over time and significantly correlated with state

WTC. For instance, students who felt more positive in the classroom during the semester tended to be more willing to communicate in English. These findings support MacIntyre and Charos's (1996) claim that L2 learner personality and affect jointly constitute the psychological context for L2 communication. Although personality has been widely studied as a basic antecedent of L2 WTC, previous studies (e.g. Cetinkaya, 2005; Chu, 2008; Ortega, 2009; Oz, 2014) tended to focus on the trait level and little attention has been paid to the state characteristics of personality and their relationships with state WTC.

The relationships between state WTC and personal antecedents reported in Study 1 was at the between-person level. At the within-person level, Study 2 suggested that some changes in state WTC were less affected by changes in the learning situation, and more by personality states on the day or the moment. For example, it was observed that one of the most active and proficient learners kept silent in a lesson. Afterwards, the student explained that she failed to concentrate because she did not sleep well and was struggling with 'unexpected affairs'. However, after studying the relationship between state WTC and L2 learners' emotional states (focusing on anxiety), Macintyre and Legatto (2011) reported that changes in anxiety do not necessarily bring about changes in state WTC. Future studies interested in the impacts of personal antecedents on within-person variability in state WTC may estimate the within-person relationships (i.e. contingencies) between different state variables (e.g. Big-Five personality states and affect) and state WTC.

8.3.2 Situational antecedents of trait and state WTC

At the trait level, Study 1 and 2 showed inconsistent results on the relationship between WTC and classroom support. Study 1 found that trait WTC significantly correlated with the support received in College English classes. More precisely, students who perceived their English teachers, peers and learning tasks as more supportive tended to be more willing to communicate in English. This is in line with the clear majority of previous studies in the literature (e.g. Khajavy et al, 2014; Peng & Woodrow, 2010) in suggesting that trait WTC is significantly related to classroom support. However, Study 2 did not find significant correlation between trait WTC and classroom support. Among the subscales of classroom support, only peer support significantly correlated with trait WTC. That is, students who perceived their partners and classmates as more engaging and cooperative tended to be more willing to communicate. However, student perceptions of teachers and tasks did not make a great difference to their trait WTC. As far as I know,

Peng's (2007a) study is the only one that did not find a significant relationship between classroom support and trait WTC.

By comparing the above studies, it was found that the inconsistency in findings with regard to the relationship between trait WTC and classroom support might be attributed to differences between different L2 learning programmes. Most studies (e.g. Khajavy et al, 2014; Peng & Woodrow, 2010) that reported significant relationships between trait WTC and classroom support were conducted with students on traditional L2 learning programmes, such as the College English course for non-English major undergraduates in Study 1. In contrast, Study 2 and Peng's (2007a) study, were conducted with students on intensive L2 programmes. Although the participants in Peng's (2007a) study did not learn English as a major, they took intensive English courses⁹ and had large amounts of English exposure like the English major students in Study 2. Hence, it might be concluded that trait WTC is significantly related to perceptions of classroom support in traditional EFL learning contexts rather than intensive L2 learning contexts. Future research might be carried out to test this assumption and explore why it might be the case.

At the state level, Study 1 found that a set of situation characteristics, based on the framework proposed in the systematic review, significantly correlated with state WTC. These situation characteristics were related to interlocutors (e.g. teacher support, classmate support, and partner support) and tasks (e.g. task-usefulness, task-interest, etc.), and could be categorised into three major dimensions as suggested in the proposed framework (i.e. positivity, negativity, and duty). When comparing between individuals, students who reported more positively on these situational characteristics tended to be more willing to communicate in the English classroom during the study. For example, students who received more support from the teacher and peers tended to be more willing to communicate. These results are consistent with studies such as Khajavy et al. (2014) and Peng and Woodrow (2010). However, as Khajavy et al. (2014) and Peng and Woodrow (2010) only measured WTC at the trait level using cross-sectional studies, the correlation between WTC and its situational antecedents reported by them could not indicate relationships at the state level.

At the within-person level, previous studies tended to use small samples together with

⁹ According to Peng (2007a), this intensive programme increased English teaching to 20 hours per week, while on traditional programmes it was 6 hours per week.

data collected through interviews or learning journals to suggest how state WTC could be generated or hindered by perceptions of specific situations (i.e. situation characteristics). As a complement to previous research, Study 1 was the first to statistically analyse the co-variance between state WTC and its situational antecedents within individuals using a larger sample and a high-density repeated measurement design. It was found that state WTC in different situations co-varied with changes in student situation perceptions. Study 2 further explored the psychological process of communication generation within individuals. The themes drawn from participant responses were generally consistent with the situational antecedents of state WTC suggested in Study 1, with more detailed and in-depth explanations. In the following sections, I will discuss the within-person relationships between state WTC and different situation characteristics based on the findings of Study 1 and 2. With regard to each situational characteristic, details of how state WTC could be generated or hindered in different situations will be illustrated.

8.3.2.1 *Teacher support*

Study 1 suggested that one of the most important situation characteristics correlated with state WTC was support. In the current context, support offered by teachers played an important role in influencing state WTC, which might be a result of students' relatively low L2 proficiency in general. Students might encounter difficulties in understanding discourse or expressing themselves, which could possibly prevent them from participating in communication activities. In these circumstances, teacher assistance and encouragement could be extremely helpful. Similarly, previous studies (e.g. Pawlak et al., 2016; Peng, 2012; Riasati, 2012) suggest that state WTC in classrooms is primarily influenced by teachers, including their attitudes towards students, rapport with students, and teaching styles.

However, unlike previous studies (e.g. Peng, 2012; Peng et al., 2017; Zarrinabadi, 2014) that focused on teacher-related situation cues such as a smile, a confirmatory phrase (e.g. very good, well done, etc.), or a joke, student responses in Study 2 did not highlight any particular verbal or non-verbal behaviour as a cue for teacher support. Students noted that a smile from a teacher did not necessarily bring about a positive communication situation. Rather, they believed that concern and encouragement could be perceived from teacher critiquing and constructive feedback. That is, what directly affected students' state WTC were not teachers' words and behaviour as such, but student subjective interpretations of teacher attitudes towards students. Teacher attitude has been reported by previous

research as an important situational antecedent of state WTC (e.g. Cao, 2013; Lee, 2009; Riasati, 2012). For example, it was suggested that students felt better when their teacher paid equal attention to everyone in the class rather than showing special concerns for a few individuals (Riasati, 2012).

The rapport between teachers and students or the desire to establish rapport could make students feel responsible for meeting teacher expectations through teacher-student interaction and activity participation. Similarly, Bernales (2016) found that student communication behaviour in some situations might be triggered by teacher expectations, although these might not correspond to students' own desires. Bernales (2016) explained this by referring to ought-to L2 self (see 1.3.2.2). That is, students may not want to participate in some communication activities; however, as they believe that their teachers expect them to communicate, they decide to enter into discourse in order to please the teachers. When proposing responsibility as one of the most important psychological antecedents of state WTC, Kang (2005) also claimed that one reason for the feeling of responsibility was unwillingness to "ruin an interpersonal relationship" (p. 285). In the current study, the interpersonal relationship could be interpreted as teacher-student rapport.

Teacher support could also be perceived based on teaching styles. For example, it seemed that students' state WTC could be boosted when teachers promised small gifts or bonus marks to active communicators. Some students explained that it did not matter whether it was a pencil or a bonus mark, because what they needed was a motive that could persuade them to take part in the activities. The impact of gifts or praise has not been studied in L2 WTC literature, although how marks influence state WTC has been explored. Some studies (e.g. Eddy-U, 2015) reported favourable results, suggesting that marks could be regarded as short-term goals that motivated students to take part in communication activities. However, others (e.g. Pawlak & Mystkowska-Wiertelak, 2015; Riasati, 2012) argued that marks were closely associated with assessments, which could make students feel more anxious and thus less willing to communicate. That is to say, how state WTC changes depends on how an individual student interprets marks. For those who perceive marks as motivators, their state WTC could be stimulated accordingly, whereas for those who perceive marks as results of competition and evaluation, negative feelings might be triggered, and state WTC would probably decrease.

8.3.2.2 *Classroom atmosphere*

In L2 classrooms, students may not only receive support from teachers, but also from classmates. Student responses suggested that the classroom atmosphere also influenced their state WTC. Classmates' active participation in a communication activity could increase one's eagerness to enter into discourse. However, when the class was silent and no one actively participated, students might feel embarrassed to talk. Some students expressed that they were expecting someone to break the silence and then they would continue the conversation; however, they all felt embarrassed to be the first one to initiate communication. This result corresponds to Eddy-U's (2015) finding that an active classroom atmosphere could push students to join in communication activities. The reason might be that students like to stay conform with the majority in the classroom, so they prefer to follow others rather than standing out (Peng, 2012).

The classroom atmosphere was not only influenced by student engagement, but also co-constructed by the teacher and students. Some students responded that the teacher seemed to be unhappy when they did not interact with her, which made the classroom atmosphere discouraging. In other words, teacher reaction towards student engagement could make a difference to student state WTC. Like the current study, Riasati (2012) suggested that a positive classroom atmosphere could be built collaboratively by the teacher and students. Some students in Lee's (2009) study believed that the classroom atmosphere was largely affected by teacher attitudes, and it was the teacher's responsibility to construct a positive learning environment. These findings suggest that teacher support and classroom atmosphere are closely associated. When students do not actively communicate in a lesson or activity, the teacher needs to be more encouraging and supportive to make the atmosphere more relaxing instead of showing anger or dissatisfaction.

Additionally, state WTC was associated with the number of interlocutors, relationships with them, and their L2 proficiency. As suggested by previous research, students may feel nervous speaking English in front of a large audience, particularly when they are not familiar with the audience (Kang, 2005). Similarly, in the current study, students' state WTC in whole-class activities tended to be low, especially at the beginning of the semester when they did not know each other well. Some students noted that their feelings of nervousness in whole-class activities gradually decreased, as they became increasingly familiar with classmates. Comparisons between whole-class, group and dyadic activities have been made by a few studies (e.g. Cao, 2011, 2013; Cao & Philp, 2006; Lee, 2009;

Mystkowska-Wiertelak, 2016; Riasati, 2012; de Saint Leger & Storch, 2009; Zhong, 2013). Like the current study, previous studies suggested that students felt less nervous and more willing to communicate in dyads or small groups than in whole-class activities. Moreover, current findings support Cao and Philp's (2006) and Riasati's (2012) claim that student WTC would change with familiarity.

8.3.2.3 *Peer cooperation*

In group or dyadic activities, peer cooperation can be more effective than teacher support. In these situations, the interlocutor(s) one works and communicates with matters more than the class. Some students may not be ready to talk at the very beginning of a communication activity but decide to participate after being inspired by their interlocutors' discourse, particularly when they have different or contradicting opinions. In the literature, there were similar findings suggesting that L2 learners tended to be more willing to communicate with interlocutors who were cooperative and actively contributed to the progress of communication (e.g. Cao, 2013; Cao & Philp, 2006; Kang, 2005; Riasati, 2012). It indicates that cooperation between peers could not only contribute to a harmonious atmosphere for communication, but also help to generate new ideas and perspectives.

Sometimes state WTC dropped because students were unable to understand their classmates, as most students were not fluent English speakers and had strong accents. As a result, some reported that they preferred communicating with interlocutors who had higher English proficiency. They believed that more proficient interlocutors could provide them with effective language support helpful for L2 development. However, other students reported they were extremely unwilling to talk after listening to advanced learners' talks, because they felt inferior and worried about being negatively evaluated by others. Like the current results, previous findings concerning the impact of interlocutor proficiency on state WTC are not straightforward. Some studies (e.g. Cao, 2011; Eddy-U, 2015; Yu, 2015) reported that students preferred communicating with more proficient interlocutors; however, Kang (2005) reported that students were reluctant to communicate with more proficient L2 learners due to a feeling of insecurity and fear of making mistakes. It is worth noticing that some students in the current study argued that it was the most proficient interlocutors who provided them with feelings of security, which is not incompatible with Kang's (2005) finding, and confirms the positive role played by security on state WTC. It suggests that what directly affects state WTC are not objective

situation cues (here interlocutors' L2 proficiency), but individual perceptions of situations, i.e. situation characteristics (here the feeling of security).

From advanced learners' perspective, sometimes they were willing to communicate but deliberately gave up communication opportunities to boost others' state WTC, as they would like to know other opinions and gain new perspectives. Students with lower L2 proficiency usually need more time to organise their language and build up confidence. However, it is believed that while less proficient learners are hesitating or preparing for communication, more proficient learners may take up the opportunities (Zhong, 2013). According to de Saint Leger and Storch (2009), more proficient learners tended to behave more confidently and readily, so communication in classrooms might be dominated by them, making others unable to enter into discourse. However, advanced learners in this study made efforts to give up their dominance and make the classroom environment less competitive as another way of showing peer support.

In group or dyadic activities, when interlocutors do not display intention to talk, a feeling of responsibility and state WTC may be boosted. As previously discussed, a silent classroom can make students feel embarrassed in whole-class activities (see section 8.3.2.2). However, when silence occurs in group or dyadic activities, students tended to feel more responsible than embarrassed. They believed that they had a responsibility to cooperate with partners or group members, so had to take the turn to avoid embarrassment when others did not intend to talk. This supports Kang's (2005) claim that feelings of responsibility are negatively associated with the number of interlocutors. Similar results have been reported by Peng (2012). In her study a student responded that she felt like talking because she wanted to "avoid the awkward silence" and "break the ice" (p. 208). The feeling of responsibility here may result from a desire to maintain good relationships with others in the same group (Eddy-U, 2015).

8.3.2.4 *Task-interest*

Among the situation characteristics related to tasks, task-interest was of vital importance. This research suggests that students are more likely to communicate in activities that they are interested in, supporting Dörnyei's (2009) claim that interest is a prime motivator in task participation. The perception of task-interest is based on types of communication activities (e.g. games, presentations, role-plays, etc.), learning materials used (e.g. novels or textbooks), and relationships with interlocutors.

Rather than specifying any type of activity, students in this study emphasised their subjective thoughts and feelings by reporting that they were more interested in innovative activities that they had not experienced before and activities that elicited deep thinking and new viewpoints. This is in line with Eddy-U's (2015) study, in which students reported that they were more interested in activities such as games, group discussions, and role-plays, because games were something that could break the class routine and group discussions and role-plays could stimulate new perspectives. Although Eddy-U (2015) put more emphasis on specific activities, the interpretation of why students were interested in some activities but not others was like the results of this study. It indicates that attention would be better shifted from specific activities (i.e. situation cues) to student subjective perceptions of situations (i.e. situation characteristics) when exploring situational antecedents of state WTC (Peng, 2012). For example, students' task-interest may be triggered and thus become more willing to communicate when experiencing a novel activity. However, it does not really matter whether the activity is a game.

'Learning material' is another situation cue that influences student perceptions of task-interest. In the current context, students tended to be bored with the textbook but interested in other learning materials such as novels and newspapers selected by teachers. Learning material is not a commonly reported situational antecedent of state WTC. The low interest in the textbook rather than other learning materials might, to some extent, be explained by topic-related differences. Compared to lively plots of novels and up-to-date news in newspapers, the topics chosen for the textbook might be too serious or old-fashioned for students to discuss. The impact of topic on task-interest and state WTC has been noted by previous studies. Like the current findings, Wolf (2013) found that students had significantly greater interests in self-selected topics than in textbook-assigned ones.

As for the influence of interlocutors on task-interest, the current result differs from previous findings. Previous studies (e.g. Cao & Philp, 2006; Eddy-U, 2015; Kang, 2005; Pawlak & Mystkowska-Wiertelak, 2015; Riasati, 2012) suggested that students preferred to communicate with familiar interlocutors, whereas students in Study 2 said that they were more interested in exchanging information and opinions with classmates that they were not very familiar with. One reason might be that the students in the current research were not proficient in English, so they were afraid of speaking and losing face in front of their friends. Both Cao (2011) and Kang (2005) reported similar findings, although their studies were conducted in English-speaking countries where students of different

nationalities learned together. They found that international students were less interested in speaking English with classmates of the same nationality, because they were more likely to know each other in person and thus felt embarrassed to show their low English proficiency. Another explanation to the inconsistency in findings might be the different degrees of familiarity discussed in different studies. Previous studies (e.g. Cao & Philp, 2006; Pawlak & Mystkowska-Wiertelak, 2015) mainly compared friends, acquaintances and strangers as different types of interlocutors. However, as the students in the current study lived and learned together every day (see section 7.2.2), any interlocutor in the classroom might be a friend, and nobody was a complete stranger. What the students meant to say might be that they were bored talking with the same interlocutors and were interested in seeking new communication experiences and different perspectives.

8.3.2.5 *Task-usefulness*

Although task-interest could be an important facilitator of state WTC, students noted that they were unwilling to communicate in activities that elicited nothing useful. For example, they wondered why they should be asked to sing or paint in a language classroom and how this kind of activity could contribute to language learning. Like these concerns, a participant in Peng's (2012) study commented "Maybe it was a little funny, but too ridiculous. It just provided us with fun." (p. 208). It seemed that interesting but meaningless activities could reduce student willingness to communicate.

Students preferred activities that could provoke meaningful communication and contribute to L2 development. Although the textbook was generally seen as boring (see section 8.3.2.4), some students noted that they felt a responsibility to participate in and learn from textbook-related tasks due to the perceived importance and usefulness of the textbook. This result is in line with Wolf's (2013) finding that students perceived textbook-assigned topics as less interesting but more important than self-selected topics. As suggested by Bernales (2016), the feeling of having to learn could be understood as an ideal L2 self. That is, sometimes state WTC can be boosted by L2 learning goals (see section 1.3.2.2). Similarly, Kang (2005) suggested that the feeling of responsibility in L2 communication tended to be related to personal motives, which could be interpreted as L2 learning goals in this context. Hence, it would be better for teachers to reconsider the learning objective of each communication activity they design and explicitly introduce it to students, so to raise awareness of the importance and usefulness of each activity and how it may help to achieve personal L2 learning goals.

It was found that student state WTC tended to increase when they believed in the effectiveness of a task in developing their communicative competence and English proficiency. This is consistent with Zhong's (2013) claim that state WTC in group activities was associated with perceived effectiveness of collaborative learning. Other than effectiveness, efficiency of a task could also affect perception of task-usefulness. Some students argued that too much time was wasted while waiting for responses in communication activities, which made them inattentive and unable to concentrate. This result contradicts previous findings on teacher wait time that suggest L2 teachers should be more 'patient' (e.g. Pawlak & Mystkowska-Wiertelak, 2015; Zarrinabadi, 2014; Zhong, 2013). Previous studies suggested that students required more time to arrange responses, so extended wait time could reduce feelings of unease and thus increase state WTC. However, this study suggests that redundant wait time may distract students from participating in communication activities. Hence, how time can be used in each activity might be something for language teachers to consider when preparing a lesson.

8.3.2.6 Task-confidence

As proposed in MacIntyre et al.'s (1998) heuristic model, Study 1 found that task-confidence significantly correlated with state WTC. Similarly, one of the most frequently occurred themes in Study 2 was lack of confidence, which seemed to be a result of low L2 proficiency together with intolerance of mistakes. Some students argued that they had to be 'a hundred percent' sure about the correctness of their discourse before speaking up in the classroom, although their proficiency was not high enough to guarantee this. Hence, they preferred not to talk rather than make mistakes and losing face. Task-confidence has been widely suggested by previous studies as one of the most important factors influencing state WTC; however, only a few studies (e.g. Eddy-U, 2015; Liu, 2002; Riasati, 2012) have revealed that lacking confidence is due to intolerance of mistakes. In line with the current study, Liu (2002) and Riasati (2012) found that when students worried excessively about grammatical accuracy and feared making mistakes, they were less likely to communicate confidently in L2 classrooms. Similarly, Eddy-U (2015) suggested that intolerance of mistakes might seldom occur in some Western countries but appeared to be a major obstacle to the L2 acquisition of learners in EFL learning contexts like China. It might be because English teaching in these contexts has long been dominated by the traditional grammar-translation approach, which puts emphasis on accuracy (Hu, 2002).

Student task-confidence changed across different activities. Due to their relatively low English proficiency, students required plenty of time to arrange discourse before talking. Hence, they preferred monologue activities such as presentations rather than dialogues that featured spontaneous responses without much time to think. This is consistent with Mystkowska-Wiertelak and Pawlak's (2014) finding that students were generally more willing to communicate in monologues than dialogues and Cao's (2013) finding that students preferred talking in presentations that offered plenty of time to prepare and practice in advance. However, a different result was reported by Peng (2012), who argued that students were very willing to communicate in dialogue activities that provoked authentic interactions and were effective in facilitating L2 development.

Sometimes, students' task-confidence and state WTC may be influenced by their previous L2 communication experiences in the classroom and their previous performance. For example, if students did not perform well in an activity at an earlier stage of a lesson, their task-confidence might be negatively affected and state WTC in later activities might decrease. Similar findings have been reported by Kang (2005), showing that after experiencing difficulties in expressing themselves or understanding others, students tend to become less secure and thus less willing to communicate. That is, a negative communication situation may not only affect student state WTC and language performance at that point in time, but also reduce task-confidence in later communication. However, some students did note that negative effects of previous unpleasant communication experiences were not enduring but could gradually reduce over time.

8.3.2.7 *Task-difficulty*

Interestingly, unlike the literature, in which task-difficulty was suggested as an important situation characteristic that could affect L2 learners' state WTC in classrooms, Study 1 found that task-difficulty was the only selected situation characteristic that was not significantly related to state WTC, either at the between- or within-person level. However, as task-difficulty was a recurring theme in Study 2, a more detailed explanation of how state WTC could be influenced by it will be discussed in this section.

On the one hand, state WTC could be reduced if a task is too easy. Some students reported that questions asked by their teachers during communication activities were so easy that everyone was able to answer them. In this situation, students preferred keeping silent because they were afraid of being seen as showing off by peers. On the other hand,

difficult tasks beyond student capability could also reduce state WTC. These results support Eddy-U's (2015) claim that tasks at appropriate levels of difficulty (i.e. neither too difficult nor too easy) are more likely to promote state WTC. It indicates that the relationship between task-difficulty and state WTC may be nonlinear, which may explain why Study 1 did not find a significant correlation between them. More evidence is needed to clarify the relationship between state WTC and task-difficulty.

The difficulty of a task could be manifest in two aspects: the topic being discussed and requirements of the task. Student state WTC tended to be low when discussing unfamiliar topics because they lacked background knowledge and relevant ideas. A familiar topic could enable students to think and talk without obstacles, whereas an unfamiliar topic could increase task-difficulty, making students nervous and unwilling to talk. However, some students realised that it is not easy to come up with a topic that can meet everyone's expectation, because individuals differ in interests and expertise. To find out which topics that most students are familiar with and have background knowledge of, L2 teachers could negotiate with students when selecting topics for discussion (Zarrinabadi, 2014). Evidence has been provided by Wolf (2013) showing that students perceived more background knowledge related to their self-selected topics than assigned topics.

In the L2 learning context, a task is a piece of work that learners are asked to complete by communicating in the target language (Ellis, 2003; Nunan, 2004). This task requires both linguistic and non-linguistic skills. Previous studies on the relationship between task-difficulty and WTC (e.g. Eddy-U, 2015; de Saint Léger & Storch, 2009) tend to focus on linguistic difficulties (e.g. in terms of fluency, pronunciation, and vocabulary). By contrast, this study found that non-linguistic requirements of tasks also affected student state WTC. Students reported that some communication tasks required advanced comprehension skills (e.g. understanding metaphors in articles), which were beyond their capability and could reduce their state WTC. Similarly, a student in Cao's (2014) study reported that he was not willing to communicate in a group discussion because he had difficulty in comprehending the reading material. When a comprehension difficulty is experienced by most of a class, it cannot not be simply attributed to students' low L2 proficiency or comprehension ability but may indicate a mismatch between task-difficulty and student ability.

8.3.3 Individual differences in within-person relationships

Although within-person variability in state WTC is generally associated with changes in the above situation characteristics, individual differences in these within-person relationships cannot be overlooked. By comparing between individual contingencies, Study 1 found that (a) some individuals were more likely than others to adapt their state WTC to the changes in the learning situation, and (b) some individuals might modify their state WTC inversely compared to the majority. These individual differences were also observed in responses in Study 2. Two examples will be illustrated in this section.

Firstly, individual differences occurred in responses towards gifts and bonus marks offered by teachers. Generally, rewards like these boosted state WTC; however, different students seemed to experience different psychological processes (see section 7.3.2.1). Some appreciated these rewards, eagerly engaging in corresponding communication activities; while some noted that they did not enjoy the feeling of being compelled by rewards and were ‘fed up with these childish treats’, although their state WTC did increase. Additionally, some students whose state WTC was relatively stable were not affected by these rewards. They reported that their state WTC was ‘the same as usual’ whether teachers provided rewards or not.

Secondly, the relationship between peer support and state WTC was not straightforward. Most students who were not highly proficient in English usually received support from peers. With language support offered by more proficient learners, they could feel more secure and thus more willing to communicate. From this perspective, a supportive atmosphere could bring about higher state WTC. However, this pattern may not apply to advanced learners. Rather than receiving support from peers, advanced learners were more likely to be those who offered support. These students suggested that they provided support by restraining their own state WTC to leave more preparation time and communication opportunities for less proficient students. Hence, it could be argued that peer support may also lead a decrease in some students’ state WTC in certain situations. The results indicate that different individuals may interpret the same situation from different perspectives, and thus may adjust their state WTC differently.

8.4 Research Question 3: Predictability

As variability in trait and state WTC was not random error but can be systematically

explained by personal and situational antecedents, the last step is to test whether systematic variability in trait and state WTC could predict language learning performance. In the current research, language learning performance was measured through two types of assessments: student self-assessment and teacher assessment (including teacher subjective judgments and more objective paper-based exam scores). It was found that intention to communicate (i.e. WTC) did not always significantly correlate with teacher-assessed language learning performance; however, it significantly correlated with student self-assessed performance. In the following sections, trait and state WTC's relationships with language learning performance will be discussed in detail.

8.4.1 Trait WTC and language exam performance

Study 1 found that both trait and state WTC significantly predicted end-of-term scores of the College English course. To be more specific, students with high and stable state WTC tended to score higher than students with more flexible WTC, and students with low and stable WTC tended to score lower. However, Study 2 found that trait WTC significantly correlated with scores subjectively assessed by teachers at the end of the semester (i.e. end-of-term teacher judgment), but the correlation between trait WTC and scores on the end-of-term paper-based exam was not statistically significant. The results of the two studies are not necessarily incompatible. The end-of-term scores in Study 1 were self-reported by students, while the scores in Study 2 were reported by teachers according to records. As discussed in section 6.4, it was impossible to know whether the students in Study 1 self-reported their English scores honestly and accurately. Additionally, as scores that these students could access were combinations of results of the teacher subjective judgment and the paper-based exam, it was impossible to know the separate correlations between trait WTC and paper-based exam, and trait WTC and teacher judgment.

For both Study 1 and 2, formative assessments were teacher judgments of student participation and language performance during a semester. Although these largely relied on teacher subjective impressions, the focus was usually on oral language performance in communication activities; whereas summative assessments were paper-based language exams held at the end of the semester. The result of Study 2 indicates that students who are more willing to communicate do not necessarily perform better in paper-based exams; however, they are very likely to perform better in oral communication in terms of quantity, fluency and comprehensibility of oral language production (Savignon, 2005). Similar results have been reported by previous studies: Naderifar and Esfandiari (2016) found

significant correlation between trait WTC and oral English proficiency; however, Yashima (2002) did not find significant relationships between trait WTC and L2 proficiency and attributed this to the absence of an oral exam to measure speaking skills.

Results may have been different if oral exams had been carried out together with paper-based exams to assess speaking and communication skills together with other language skills (e.g. reading and writing). When assessing oral language performance, the emphasis tends to be placed on the ability to conduct effective information exchanges (i.e. communicative competence), which is likely to be developed through frequent interpersonal communication. Paper-based exams tend to pay more attention to the accuracy of grammar and vocabulary, which is not necessarily accompanied by communication practice. In other words, the paper-based exam scores that are highly valued in this context may not be able to assess communicative competence. Although the teacher judgements to some extent assessed communicative competence, they were largely based on subjective impressions. As communicative competence is regarded as an important learning goal for both English major and non-English major students in China (see section 5.2), oral exams are suggested to complement paper-based exams to make language assessments more comprehensive and objective.

8.4.2 State WTC and self-assessed language performance

Between-person relationships between trait and state WTC and self-assessed language performance were tested in Study 1 and 2, and statistically significant correlations were found. These results support MacIntyre et al.'s (1998) claim that perceived competence could be more effective than actual competence in predicting L2 WTC. Similarly, several studies have reported significant correlations between L2 WTC and self-assessed communicative competence, although most only focused on the trait level (e.g. Fallah, 2014; MacIntyre et al., 2003; MacIntyre & Doucette, 2010; Mystkowska-Wiertelak & Pawlak, 2014; Peng & Woodrow, 2010; Yashima, 2002).

At the within-person level, Study 1 found that state WTC and self-assessed language performance was also positively correlated in general, although there were individual differences in these within-person relationships. That is, students tended to perceive themselves performing better when their state WTC increased in a specific situation. Alternatively, it might be that students became more willing to communicate when they were more confident in their performance. The within-person relationships between state

WTC and self-assessed L2 performance in specific situations have seldom been studied in the literature. Two relevant studies to my knowledge are Cao (2014) and Freiermuth and Jarrell (2006), showing that low perceived competence could bring about negative thoughts and feelings (e.g. losing confidence in speaking English) that hindered state WTC in some communication activities. Similar to the findings of the current study, they also indicated within-person relationships between state WTC and self-assessed L2 performance. However, causation cannot be inferred based on results of correlational analyses.

8.4.3 Situation-contingent WTC and language exam performance

The current research is heuristic in being the first effort to investigate whether language learning performance could be predicted by situation-contingent WTC. Study 1 found that language exam scores significantly correlated with the contingency of state WTC on task-interest. That is, students who were more willing to communicate when perceiving communication tasks as interesting tended to perform better in the end-of-term language exam. Although no previous research has investigated the contingencies of state WTC on situation characteristics in the field of SLA, the contingencies of personality states on situations have been studied by researchers in the field of personality science. Fleeson (2007) was the first to use contingencies to explain within-person variability and found that Big-Five personality states were contingent on a set of different situation characteristics, such as friendliness of interlocutors and task orientation. The effects of these contingencies on performance have also been studied. For example, Minbashian et al. (2010) found that task-contingent conscientiousness could predict adaptive performance, i.e. individuals with higher levels of task-contingent conscientiousness were more likely to maintain their performance as task complexity increased. The contingencies of state WTC on situation characteristics and their relationships with language learning performance may be further explored by future research interested in within-person variability in state WTC.

8.5 Unexpected Outcomes

8.5.1 Distinction between WTC and communication behaviour

8.5.1.1 *WTC vs communication behaviour comparison*

At the trait level, significant correlations were found between WTC and communication behaviour. That is, students who were more willing to communicate in English tended to

communicate more in English classrooms. This is in line with the result of Mystkowska-Wiertelak and Pawlak's (2014) study, in which significant correlations between trait WTC and frequency of communication in English was found. However, unlike Mystkowska-Wiertelak and Pawlak (2014) who asked students to self-report their frequency of communication at the same time when reporting trait WTC, Study 2 measured communication behaviour through teacher observation, which is arguably more objective than self-reports.

Results of relationships between state WTC and communication behaviour at the state level were not consistent. A statistically significant correlation between state WTC and communication behaviour was observed in Study 1; however, this was not observed in Study 2, which might be attributed to the small sample size as well as the fewer measurement occasions used in Study 2. Nevertheless, both Study 1 and 2 found that the average level of communication behaviour (reported either by students themselves or by teachers) was always lower than the average level of WTC, indicating that not all intention to communicate can successfully transform into actual communication behaviour. However, it does not exclude the possibility that students may communicate even without communication intention in some situations. In the following section, I will discuss (a) why some state WTC cannot transform into actual communication behaviour; and (b) why students may communicate without state WTC.

A major factor that prevents state WTC from transforming into communication behaviour is the lack of communication opportunities. To initiate any voluntary communication behaviour, one must have state WTC and at the same time be given an opportunity to communicate. However, state WTC and opportunities not always appear simultaneously. When a teacher asks a question, several students may raise their hands. Raising a hand in the classroom can be seen as a nonverbal cue to show state WTC (MacIntyre et al., 1998). However, only one student may be called upon by the teacher to give the answer. Hence, only the student who gets the chance to communicate finally conducts communication behaviour. Although students in the current study responded that their teachers generally expected them to communicate and provided them with sufficient opportunities, it did not indicate that every communication intention was able to meet an immediate opportunity. In some specific situations, students felt that they were not given enough opportunities to express their thoughts and feelings, whereas in other situations, too many opportunities were offered when students were not willing to communicate.

In addition, not all students can realise the necessity of verbalising thoughts and feelings and communicating with others. Some students claimed that the process of thinking could be more important than presenting or talking with others. They believed that they could participate in communication activities without verbalising their thoughts, because constructing and rehearsing discourse in their minds was also an effective form of communication. This claim questions whether speaking is the only form of active participation in communication activities and whether those who talk less must be passive learners who are reluctant to use and learn the target language (Marlina, 2009). Speaking may not be the only approach to participate in communication activities, and silence may not indicate mental disengagement. Other than external speech which is a social function, there is also inner speech, an egocentric function influenced by outside factors, which influences thoughts (Vygotsky, 1986). That is, students who are not verbally participating in an activity are not necessarily absent-minded or passively receiving knowledge but may be actively thinking and constructing discourse in their minds (Shi, 2006). This might also be seen as state WTC, as it shows eagerness and readiness to enter into discourse, although, only when discourse is verbalised is communication behaviour visible.

On the other hand, students seldom showed unintended communication behaviour, unless they were nominated by teachers. Communication required by teachers can be different from the voluntary communication discussed above. In the context of SLA, communication opportunities are provided to make students engage in communicating in the target language and learn from positive communication experiences (Finocchiaro & Brumfit, 1983). During communication activities, a teacher is not the class authority but serves as a co-communicator who is equal with students (Hu, 2002). If teachers trigger communication by forcing students to talk, negative thoughts and feelings that tend to hinder state WTC are likely to be aggravated, which runs counter to the original purposes. However, some students claimed that, even when nominated, they would keep silent if they were not willing to communicate. This indicates that students cannot be forced to communicate without state WTC. Whether they communicate with reluctance, they must be willing to talk at that moment, because state WTC represents a final psychological step before communication behaviour (MacIntyre, 2007). Hence, teacher nomination could be a facilitator of communication generation. From this perspective, teacher nominated communication may not be significantly different from voluntary communication, as they both are effective information exchanging experiences that have the potential to develop communicative competence.

The current project found a significant correlation between WTC and language exam scores, which indicates that the subjective intention to communicate is relevant to language learning, regardless of whether this intention is translated into actual, observable behaviour or not. Hence, language teachers may not need to be worried if their students do not show observable communication behaviour or seem ‘apparently unwilling to communicate’. Teaching can be more effective than it seems to be, because language learning relies less on what students say, but more on their subjective intention to communicate which can be difficult to detect. Students may experience a silent phase first, thinking and phrasing in their minds, and then gradually become more confident and proficient L2 speakers.

This also raises a question: whether we should focus on ‘speaking’ and see it as the most important form of communication. Listening, reading, writing, and thinking are also valid and effective forms of communication (Marlina, 2009). Given the current findings, it is conceivable that thinking and phrasing in mind may be more effective than actually talking in promoting language learning, though this requires further investigation. Future research may further investigate the relationships between WTC, communication behaviour, and language learning performance to fully understand whether it is necessary for L2 learners to transform communication intention into actual behaviour.

8.5.1.2 Chinese students’ ‘silence’ revisited

Aware of the distinction between WTC and communication behaviour, Chinese students’ silence or ‘apparent unwillingness to communicate’ discussed in the literature could be understood as a lack of communication behaviour observed by some researchers or language teachers. However, the absence of observable communication behaviour does not imply unwillingness to communicate. WTC refers to a subjective intention that is notoriously difficult to be observed and thus better self-reported. One may have sufficient intention to communicate; however, one’s intention may not be successfully transformed into actual communication behaviour (e.g. speaking) due to personal and situational factors, such as lacking communication opportunities. Hence, it is not wise to claim that a student is unwilling to communicate just because of the absence of observable communication behaviour. Additionally, it is worth noting that the impacts of Confucianism and other cultural traditions on Chinese students are declining. Contemporary Chinese students have many characteristics in common with their Western contemporaries. For example, Shi (2006) has provided evidence showing that

contemporary Chinese students are more critical and communication-oriented when learning English in classrooms, and show little difference from their Western counterparts. Hence, previous findings or stereotypes about Chinese culture and students may not be applicable to contemporary Chinese students without considering the rapid social changes (Shi, 2006).

Moreover, as discussed in section 5.3, the relatively unwillingness to communicate in a L2 seems to be a problem shared by learners from different cultural backgrounds. Hence, it is not wise to attribute an apparent lack of WTC to Chinese traditions or Confucian heritage. Individual differences in WTC within a cultural background and potential within-person fluctuations in WTC over time may be more pronounced than culture-related differences. At the between-person level, not all Chinese students are unwilling to communicate. For example, a majority of the students studied in the current project reported that they were relatively willing to communicate in English classrooms. At the within-person level, willingness to communicate is not fixed but changes across different situations. Investigating within-person variability in WTC is of vital importance because it might indicate that WTC is malleable. As results have shown that WTC levels can predict language learning performance, the malleability of WTC may provide researchers and language teachers with opportunities to promote L2 learning performance, that is, by enhancing WTC. Therefore, future research should pay more attention to the variability and malleability of WTC rather than restricting their focus to cultural differences between students.

8.5.2 Inconsistency between teacher and student assessments

Study 1 and 2 were consistent in finding that the relationships between self-assessments and teacher assessments (including end-of-term paper-based exam scores) were very weak. Study 2 showed that results of teacher subjective judgments were significantly related to student communication behaviour (both self-reported and observed by teachers); however, results of student self-assessments were not associated with actual communication behaviour but with state willingness to communicate. The findings suggest that teachers tended to assess student language performance based on actual behaviours in communication activities, whereas students tended to assess themselves based on thoughts and feelings rather than actual behaviours. This distinction between self- and teacher assessments was also shown in the assessment criteria stated by students and teachers in Study 2.

Firstly, students believed that the prime criterion for language performance in the classroom was attentiveness during a class or activity, while teachers mainly assessed performance based on student participation and language production in communication activities. Most students in this context were satisfied with their performance if they listened attentively during the class, which might be a result of the teacher-centred learning tradition rooted in Chinese culture. Many teachers in Chinese secondary schools still adopt teacher-centred pedagogy, not only because it is a quick approach to transmit knowledge to students, but also because it is an easy way to manage classes containing large numbers of adolescents (Littlewood, 2006). Therefore, students might have been used to passive roles of receiving knowledge from their teachers due to past learning experiences, which could be an obstacle to the implementation of student-centred communication activities in university language classrooms (Hu, 2002; Littlewood, 2006). By comparison, the English teachers interviewed in Study 2 believed that actively participating in communication activities was more important than listening attentively. Only when some students did not show intention to communicate, did the teachers raise concern as to whether these students were listening attentively. These results support Liu and Littlewood's (1997) claim that there may be a "mismatch between teachers' and students' perceptions of learner role" (p. 377).

Another criterion frequently reported for assessing language performance in the classroom was L2 improvement, which can be understood in relation to Dörnyei's (2005) L2 Motivational Self System (see section 1.3.2.2). Students evaluated their improvements based on perceived learning outcomes after participating in an activity and compared these to their ideal L2 selves (i.e. personal L2 learning goals). For a student who aimed to gain more confidence in speaking English in public, the experience of presenting in front of the class might be an encouraging step forward; while for one who aimed to gain more grammatical knowledge, acquiring new vocabulary and grammar rules might be more important. By comparison, teachers perceived student learning improvements based on language proficiency and past performance, which might be interpreted as students' ought-to L2 selves (i.e. extrinsic expectations). For a student with relatively low English proficiency, even a little progress may be valued and praised by the teacher; while for a student with relatively high English proficiency, the teacher may have higher expectations and requirements.

Lastly, teachers and a few students took quality of language production into account when

assessing language performance in the classroom, however, they focused on different aspects. Students evaluated their language production according to grammar rules, whereas teachers comprehensively evaluated from different facets besides accuracy, including pronunciation, fluency, comprehensibility, etc. Communication activities in L2 classes are usually designed to achieve a balance between learning grammatical forms and communicating meaning to develop L2 communicative competence (Littlewood, 1981). Hence, when assessing performance, teachers not only focused on grammatical accuracy but paid attention to effectiveness of information exchanges. However, student emphasis on the grammatical forms might be a result of their concerns about the paper-based exam at the end of each semester as well as national English proficiency tests (e.g. Tests for English Majors and College English Tests). Although class activities have been designed to be communication-oriented, students still have to take “more traditional, form-oriented examinations which will determine their educational future” (Littlewood, 2006, p. 245). Hence, student concerns about the exams are likely to influence their criteria for self-assessment.

The above results reveal that there may be a mismatch between the assessments and learning objectives in this context. According to the curriculum, the major goal of modern English learning is to develop competence for intercultural communication, and thus communicative language teaching (CLT) is adopted as the principal approach for English teaching (MOE, 2003). However, the exams that are highly valued in this context fail to keep pace with such developments in the curriculum. The exams are still traditional, paper-based assessments, emphasising grammar accuracy that is not the focus of CLT. This may have a washback effect on the implementation of CLT. As it is claimed by Littlewood (2006), assessments may be the major constraint on student participation in classroom communication activities, and student concerns about the assessments may be the main obstacle to WTC generation. Hence, to improve WTC and meet the requirements of the curriculum, the assessment system needs to be reformed by adding an oral exam component to measure communicative competence.

8.6 Limitation and Direction for Future Research

As both Study 1 and 2 used non-probability samples recruited from the same university and the sample size of Study 2 was very small, the results reported in this thesis may not be generalised to other EFL learning contexts. As the university is a key university

specialising in science and engineering, the entrance requirements for English major students and the quality of EFL teaching in this university are not amongst the best in China. Hence, most students recruited in this research (both English major and non-English major students) may be seen as intermediate EFL learners. As results have shown individual differences in both the levels of trait and state WTC and within-person variability in state WTC among students with different levels of L2 proficiency, one may argue that the results could be different if this study was conducted in other universities with more beginning or advanced EFL learners. Additionally, it has been found that student state WTC could be influenced by teachers (e.g. attitudes towards students, rapport with students, teaching styles, etc.); however, the English teachers of the four classes under study were all young females who seemed very similar. For a more comprehensive portrayal of possible variability and stability in WTC, future studies might recruit students taught by different types of teachers in different L2 learning contexts (e.g. male teachers, more experienced teachers, teachers in other universities or institutions).

The results of Study 1 and 2 have been compared and a few differences have been found. For example, the sample in Study 2 (English major students) had generally higher and less variable WTC than the sample in Study 1 (non-English major students). However, as they were not equivalent groups, the comparison was rather tentative and not the main purpose of this research. Hence, it is still not clear whether the differences in the results of the two studies were due to the differences between English major and non-English major students or other factors such as gender differences. In China and other countries (e.g. Iran, Poland, etc.), EFL learners are divided into two types depending on whether they learn the language as a major subject or not. Both types of learners have been studied in previous research on WTC; however, not much attention has been paid to differences between them. As presented earlier in this thesis (see section 6.2.2 and 7.2.2), language learning situations for English major and non-English major students in the same university can be quite different, which may result in the differences in variability and stability in trait and state WTC. Future research might further explore differences in WTC between English and non-English major students, how these differences are generated, and how they influence language learning performance.

To note, a portion of within-person variability in state WTC observed in the current studies may be meaningless and unpredictable error variance. The observed states could be the results of latent states plus random effects. In other words, the large amount of

within-person variability in state WTC found in the current studies is systematic variability combined with random error. However, it is hard to know the amount of meaningful within-person variability, as what can be observed is only *manifest states* rather than *latent states*. Latent states cannot be observed, as they are always accompanied by random effects (Schmitz, 2006). That is to say, the effects of random error in within-person variability have not been removed during data analysis. However, the hierarchical linear modelling approach has modelled the random error at all levels, to improve the accuracy of the analysis (see Nezlek, 2001). The within-person correlations between state WTC and its situational antecedents (e.g. interest-contingent WTC) reported in Study 1 suggest that observed within-person variability in state WTC was not entirely error variance but systematic and meaningful.

As it has been found that state WTC is impacted by general personal characteristics or traits (i.e. personality) and perceptions of a specific situation (i.e. situation characteristics), the next question may be the malleability of state WTC. Although a few strategies for language teachers to train learner state WTC within and beyond L2 classrooms could be proposed based on current findings (see section 8.7.2), the current studies only identified relationships between state WTC and its situational antecedents and provide explanations for these relationships. The causality, i.e. whether it is selected situational antecedents or unknown variables that cause changes in state WTC, is yet to be clarified. Therefore, experimental research could be very useful to test the malleability of state WTC, because the existence of control groups allows controlling for the effects of third variables to clearly identify the cause for an effect (Gorard, 2013; de Vaus, 2001).

However, to my knowledge, only Munezane (2015) has conducted quasi-experiments with 373 Japanese EFL learners to examine whether L2 WTC can be enhanced via classroom interventions such as visualisation and goal setting. Munezane (2015) used visualisation activities to help students imagine their ideal L2 selves as proficient English speakers. It was found that the visualisation intervention alone was not effective in improving state WTC; however, when it was combined with a goal setting intervention (i.e. activities that helped students develop L2 learning goals), a significant increase in state WTC was found. Nevertheless, Munezane (2015) emphasises the impact of personal characteristics, such as ideal L2 self and self-regulated learning, rather than the impact of systematically shaping classroom situations to enhance state WTC. The potential malleability of state WTC is the most practical implication of research on WTC for

language teachers, and thus deserves more attention.

8.7 Conclusion and Implication

The current research has distinguished state WTC from trait WTC and has emphasised potential variability and stability in state WTC. Firstly, evidence has been provided to show the dynamic nature of WTC. The findings show that WTC not only varies across different individuals but may also fluctuate within individuals over time and across different situations. Although there are individual differences in within-person variability, its average amount can be comparable to the amount of between-person variability. It indicates that there can be a large amount of within-person variability in state WTC that cannot be neglected. Secondly, it has been found that within-person variability in state WTC is at least to some extent meaningful, as it is systematically correlated with situational antecedents. To better understand situational antecedents of state WTC, the concepts of situation cues and characteristics have been introduced to distinguish subjective perceptions from objective features of situations. A multi-layered framework has been proposed to systematically categorise the situational variables that may influence state WTC into three interlinked layers: situation cues (i.e. teacher, class, peers, activity, and topic), situation characteristics (i.e. support, cooperation, and objectives), and underlying dimensions (i.e. negativity, positivity, and duty). The findings suggest that within-person variability in state WTC cannot be simply seen as error variance but deserves further investigation. Thirdly, it has been found that both trait and state WTC could predict language learning performance, particularly when emphasising communicative competence. Correlations have also been found between state WTC and student self-assessed performance in English lessons; however, as students tend to focus on their thoughts and feelings rather than actual behaviours when assessing their own language performance, self-assessed performance may not be seen as a good indicator of actual language learning performance.

8.7.1 Theoretical implications

The framework proposed provides a theoretical implication for future research on state WTC and its situational antecedents. To better facilitate future research, the main situational antecedents of state WTC commonly reported in the literature have been identified and integrated into a coherent framework. The framework distinguishes situation characteristics (subjective perceptions of situations) from situation cues

(objective features of situations), and systemically categorises different types of situational variables into three interlinked layers. It is the first attempt to arrange previously suggested situational antecedents of state WTC together. By suggesting the framework, I hope to contribute to a better understanding of the situational antecedents of state WTC and plead for more consistency in using terminology for situation cues and characteristics in future research. The validity of the framework has been tested by my empirical studies in a Chinese EFL learning context. However, further studies are required to further refine the framework and adapt it to other L2 learning contexts.

The current research also has methodological implications for future research on state WTC. Firstly, it seems that a high-density repeated measurement approach may be most appropriate to study within-person variability over time. This promising, relatively new approach in SLA makes it possible to capture state WTC in specific situations and monitor its fluctuation over time. However, only a handful of recent studies have employed this approach to study state WTC (MacIntyre & Legatto, 2011; Mystkowska-Wiertelak, 2016; Mystkowska-Wiertelak & Pawlak, 2014; Pawlak & Mystkowska-Wiertelak, 2015; Pawlak et al., 2016). While most of these studies repeatedly measured state WTC over very short periods to capture moment-to-moment fluctuation, the current research suggests that a high-density repeated measurement approach may be adapted to study fluctuations in state WTC over longer periods of time. More studies using a high-density repeated measurement approach to study within-person variability in state WTC are required as a complement to cross-sectional surveys and experimental studies. However, it should be noted that studies using this approach can be time- and energy-consuming, and intensive repeated measurements may make participants feel bored and thus increase the dropout rate. Hence, studies using this approach need to be carefully-designed.

Secondly, this research has introduced the hierarchical linear modelling approach to test the systematicity of within-person variability in state WTC. Hierarchical linear modelling is a systematic approach to process data collected through repeated measurements, and it analyses data at both between- and within-person levels. Before the current research, the coefficients between WTC and its related variables have only been estimated between individuals. The systematicity of within-person variability in state WTC tended to be thematically analysed by previous studies using qualitative data collected through interviews and learning journals. However, this research has statistically analysed relationships between state WTC and its situational antecedents within individuals (i.e.

state WTC's contingencies on situation characteristics) and how they predict language learning performance. For a more in-depth understanding of the systematicity and predictability of state WTC, future research might shift to within-person relationships between state WTC and its situational antecedents, and the hierarchical linear modelling approach may be employed to systematically estimate these contingencies at the within-person level.

8.7.2 Practical implications

This thesis sheds light on the different types of situational antecedents that may trigger or hinder L2 learners' state WTC, which could provide a basis for a set of pedagogical implications for language teachers who would like to facilitate student state WTC by managing appropriate learning situations. Before conducting communication activities in classrooms, teachers may negotiate with students when choosing learning materials or topics for discussion. By taking student suggestions into consideration, learning materials and discussion topics selected are more likely to match student interests and background knowledge. It is evident in the findings that the perception of task-interest is a potential facilitator of state WTC, and background knowledge on a topic tends to reduce the perception of task-difficulty and thus increase state WTC. At the same time, the processes of negotiation on learning materials and discussion topics can be good opportunities for teachers to build rapport with students. According to the findings, teacher-student rapport is also likely to increase student state WTC in classrooms.

At the beginning of each communication activity, teachers are advised to provide clear instructions of what students are expected to do and explicitly state the rationale for each task. The findings have shown that students may withdraw from a group project when they are unsure about their roles in the group or duties during the project. Hence, detailed task instructions may assist students to understand how they can contribute and thus make them more willing to participate. After introducing to students what they are supposed to do, it may be necessary to explain why they are asked to do so and how it may contribute to L2 acquisition. As it has been found that state WTC can be influenced by perceived task-usefulness, explaining the rationale aims to raise student awareness of the importance and usefulness of the task to boost their state WTC.

When students are expected to communicate in groups, the composition may be something for teachers to consider. It has been found that students may get bored and

become less willing to communicate with the same interlocutors throughout a semester. Hence, it would be better to group students differently in each class or activity instead of allocating them into fixed groups or always letting them talk with their neighbours. Moreover, students' L2 proficiency should be considered. As a group may consist of students with different levels of L2 proficiency, it is necessary to prevent more proficient learners from dominating communication. It may be a good strategy to let more proficient learners initiate discourse but make sure that the less proficient have enough time and opportunities to join the conversation. This aims at boosting the less proficient learners' state WTC by offering them more time to prepare and giving them more confidence, which in turn could increase the more proficient learners' chance to gain new perspectives and deepen their thoughts.

After each communication activity, it would be better for teachers to give some constructive feedback to each student. It seems that students perceive teacher support based on feedback they receive. They prefer critiquing feedback that is more informative and constructive than short affirmative phrases such as 'well done'. As suggested by Hattie and Timperley (2007), negative feedback or disconfirmation can be more effective than positive feedback or praise. After communication activities, students expect to receive feedback that could stimulate L2 development, such as pointing out their mistakes and providing suggestions for future learning. However, it does not mean that teachers should always criticise students and highlight their mistakes. Positive comments can also be useful for students to build confidence and improve learning motivation, if the comments are concrete and informative. Concrete and helpful feedback from teachers may make students feel supported and become more eager to engage in class communication activities, whereas praise or confirmation that contains little learning-related information may disappoint them and make them less willing to communicate.

I hope that this thesis will raise researchers' awareness of relevant situation characteristics that may affect state WTC, and lead to a more comprehensive understanding of variability and stability in state WTC, i.e. how L2 learners' state WTC co-varies with their subjective perceptions of learning situations and how this influences language learning performance. As practical strategies have been proposed to facilitate state WTC by managing appropriate learning situations, this thesis is also insightful for language teachers who would like to trigger student state WTC in language classrooms.

References

- Adamson, B., & Morris, P. (1997). The English curriculum in the People's Republic of China. *Comparative Education Review*, 41(1), 3-26. Retrieved from <http://eprints.ioe.ac.uk/4755/1/Morris1997TheEnglish1.pdf>
- Ahmadian, M., & Shirvani, K. (2012, May). The role of gender and academic experience in EFL students' willingness to communicate in English in academic context. Paper presented on *A National Conference on Issues in English Literature and Language Teaching: New Trends and Criticisms*, Lorestan, Iran. Abstract retrieved from https://s3.amazonaws.com/academia.edu.documents/32288862/WTC__Ahmadian__Shirvani.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1512061173&Signature=NFZmI2uHTSUuTyDe09X%2FCttoAw%3D&response-content-disposition=inline%3B%20filename%3DThe_Role_of_Gender_and_Academic_Experience.pdf
- Aliakbari, M., & Mahjoob, E. (2016). The relationship between age and willingness to communicate in an Iranian EFL context. *Journal of Applied Linguistics and Language Research*, 3(1), 54-65. Retrieved from <https://www.researchgate.net/publication/292128678>
- Arends, R. I., Winitzky, N. E., & Tannenbaum, M. D. (1998). *Exploring teaching* (2nd ed.). Boston: McGraw-Hill.
- Arksey, H., & Knight, P. T. (1999). *Interviewing for social scientists: An introductory resource with examples*. London: SAGE.
- Asmalı, M., Bilki, U., & Duban, C. A. (2015). A comparison of the Turkish and Romanian students' willingness to communicate and its affecting factors in English. *Journal of Language and Linguistic Studies*, 11(1), 59-74. Retrieved from <http://jlls.org/index.php/jlls/article/viewFile/301/211>
- Baker, S. C., & MacIntyre, P. D. (2000). The role of gender and immersion in communication and second language orientations. *Language Learning*, 50(2), 311-341. Retrieved from http://faculty.capebretonu.ca/pmacintyre/research_pages/journals/gender_L2com2000.pdf

- Barrett, L. F., & Barrett, D. J. (2001). An introduction to computerized experience sampling in psychology. *Social Science Computer Review*, 19(2), 175-185. doi:10.1177/089443930101900204
- Bolton, K., & Graddol, D. (2012). English in China today. *English Today*, 28(3), 3-9. doi:10.1017/S0266078412000223
- Beckmann, N., Beckmann, J. F., Minbashian, A. & Birney, D. P. (2013). In the heat of the moment: On the effect of state neuroticism on task performance. *Personality and Individual Differences*. 54(3), 447-452. doi:10.1016/j.paid.2012.10.022
- Beckmann, N., Minbashian, A., & Wood, R. E. (2011). Integrating the study of within- and between-person variability in personality at work. In S. Boag, & N. Tiliopoulos (Eds.), *Personality and individual differences: Theory, assessment and application* (Chapter 2). New York: Nova.
- Beckmann, N., & Wood, R. E. (2017). Dynamic personality science. Integrating between-person stability and within-person change. *Frontiers in Psychology*, 8, 1486. doi:10.3389/fpsyg.2017.01486
- Beckmann, N., Wood, R. E., & Minbashian, A. (2010). It depends how you look at it: On the relationship between neuroticism and conscientiousness at the within-and the between-person levels of analysis. *Journal of Research in Personality*, 44(5), 593-601. doi:10.1016/j.jrp.2010.07.004
- Bernales, C. (2016). Towards a comprehensive concept of willingness to communicate: Learners' predicted and self-reported participation in the foreign language classroom. *System*, 56, 1-12. doi:10.1016/j.system.2015.11.002
- Blaike, N. (2000). *Designing social research: The logic of anticipation*. Cambridge: Polity Press.
- de Bot, K., Lowie, W., & Verspoor, M. (2007). A dynamic systems theory approach to second language acquisition. *Bilingualism: Language and cognition*, 10(1), 7-21. doi:10.1017/S1366728906002732
- Bryman, A. (2012). *Social research methods* (5th ed.). Oxford: Oxford University Press.
- Buckingham, L., & Alpaslan, R. S. (2017). Promoting speaking proficiency and willingness to communicate in Turkish young learners of English through asynchronous computer-mediated practice. *System*, 65, 25-37. doi:10.1016/j.system.2016.12.016

- Cameron, D. (2013). Willingness to communicate in English as a second language as a stable trait or context-influenced variable: Case studies of Iranian migrants to New Zealand. *Australian Review of Applied Linguistics*, 36(2), 177-196. doi:10.1075/ara1.36.2.04cam
- Cao, Y. (2011). Investigating situational willingness to communicate within second language classrooms from an ecological perspective. *System*, 39(4), 468-479. doi:10.1016/j.system.2011.10.016
- Cao, Y. (2013). Exploring dynamism in willingness to communicate: A longitudinal case study. *Australian Review of Applied Linguistics*, 36(2), 160-176. doi:10.1075/ara1.36.2.03cao
- Cao, Y. (2014). A sociocognitive perspective on second language classroom willingness to communicate. *TESOL Quarterly*, 48(4), 789-814. doi:10.1002/tesq.155
- Cao, Y., & Philp, J. (2006). Interactional context and willingness to communicate: A comparison of behavior in whole class, group and dyadic interaction. *System*, 34(4), 480-493. doi:10.1016/j.system.2006.05.002
- Cetinkaya, Y. B. (2005). *Turkish college students' willingness to communicate in English as a foreign language*. Unpublished PhD dissertation, The Ohio State University. Retrieved from https://etd.ohiolink.edu/rws_etd/document/get/osu1133287531/inline
- Cheng, X. (2002). Chinese EFL students' cultures of learning. In C. Lee, & W. Littlewood (Eds.), *Culture, communication and language pedagogy* (pp. 103-116). Hong Kong: Hong Kong Baptist University Press.
- Chu, H. N. R. (2008). *Shyness and EFL learning in Taiwan: A study of shy and non-shy college students' use of strategies, foreign language anxiety, motivation, and willingness to communicate*. Unpublished PhD dissertation, University of Texas at Austin. Retrieved from <https://repositories.lib.utexas.edu/bitstream/handle/2152/3864/chuh41126.pdf?sequence=2>
- Church, A. T., Katigbak, M. S., Ching, C. M., Zhang, H., Shen, J., Arias, R. M., ... & Mastor, K. A. (2013). Within-individual variability in self-concepts and personality states: Applying density distribution and situation-behavior approaches across cultures. *Journal of Research in Personality*, 47(6), 922-935.

doi:10.1016/j.jrp.2013.09.002

- Clément, R., Baker, S. C., & MacIntyre, P. D. (2003). Willingness to communicate in a second language: The effects of context, norms, and vitality. *Journal of Language and Social Psychology, 22*(2), 190-209. doi:10.1177/0261927X03252758
- Cohen, L., Manion, L., & Morrison, K. (2011). *Research methods in education* (7th ed.). London; New York: Routledge.
- Coolican, H. (2014). *Research methods and statistics in psychology* (6th ed.). London; New York: Psychology Press.
- Cornish, F., Gillespie, A., & Zittoun, T. (2014). Collaborative analysis of qualitative data. In U. Flick (Ed.), *The SAGE handbook of qualitative data analysis* (pp. 79-93). London: SAGE.
- Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.
- Cousin, G. (2009). *Researching learning in higher education: An introduction to contemporary methods and approaches*. New York: Routledge.
- Craighead, C. W., Ketchen, D. J., Dunn, K. S., & Hult, G. T. M. (2011). Addressing common method variance: guidelines for survey research on information technology, operations, and supply chain management. *IEEE Transactions on Engineering Management, 58*(3), 578-588. doi:10.1109/TEM.2011.2136437
- Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Thousand Oaks; London: Sage Publication.
- Creswell, J. W., & Clark, V. P. (2011). *Designing and conducting mixed methods research* (2nd ed.). Los Angeles: SAGE.
- Crotty, M. (1998). *The foundations of social research: Meaning and perspective in the research process*. London: Sage Publications.
- Dai, W. (2008). 中国高校外语教育30年 [Foreign language education in Chinese higher education institutes in last three decades]. Retrieved from <http://edu.sina.com.cn/en/2008-09-02/103043805.shtml>
- Debusscher, J., Hofmans, J., & de Fruyt, F. (2016). The multiple face(t)s of state conscientiousness: Predicting task performance and organizational citizenship behavior. *Journal of Research in Personality*. doi:10.1016/j.jrp.2016.06.009

- Denscombe, M. (2008). Communities of practice: A research paradigm for the mixed methods approach. *Journal of mixed methods research*, 2(3), 270-283.
- Denscombe, M. (2017). *The good research guide: For small-scale social research projects* (6th ed.). London: Open University Press.
- Dewaele, J. M., & Furnham, A. (1999). Extraversion: The unloved variable in applied linguistic research. *Language Learning*, 49(3), 509-544. doi:10.1111/0023-8333.00098
- Dewey, J. (1957). *Reconstruction in philosophy*. Boston: Beacon Press
- Donovan, L.A. & MacIntyre, P. D. (2004). Age and sex differences in willingness to communicate, communication apprehension and self-perceived competence. *Communication Research Reports*, 21(4), 420-427. Retrieved from http://faculty.capebretonu.ca/pmacintyre/research_pages/journals/age_sex2004.pdf
- Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Dörnyei, Z. (2009, September). The antecedents of task behaviour: A dynamic systems account of task motivation. Paper presented at the *Third Biennial International Conference on Task-Based Language Teaching*, Lancaster, UK. Abstract retrieved from http://www.lancaster.ac.uk/fass/events/tblt2009/presentations/DORNYEI_handout
- Dörnyei, Z., & Kormos, J. (2000). The role of individual and social variables in oral task performance. *Language Teaching Research*, 4(3), 275-300. doi:10.1177/13621688000400305
- Dörnyei, Z., MacIntyre, P. D., & Henry, A. (2015). Introduction: Applying complex dynamic systems principles to empirical research on L2 motivation. In Z. Dörnyei, P. D. MacIntyre, & A. Henry (Eds.), *Motivational dynamics in language learning* (pp. 1-7). Bristol: Multilingual Matters.
- Eddy-U, M. (2015). Motivation for participation or non-participation in group tasks: A dynamic systems model of task-situated willingness to communicate. *System*, 50, 43-55. doi:10.1016/j.system.2015.03.005
- ELT Advisory Board under MOE. (2000). *高等学校英语专业教学大纲* [English curriculum for English majors in higher educational institutes]. Beijing: Foreign

Language Teaching and Research Press.

- Eng, W. E., & Kumar, V. (2009). Constructing meaning in a bilingual learning environment: Two primary classrooms in Malaysia (task 1). In R. Barnard, & M. E. Torres-Guzmán (Eds.), *Creating classroom communities of learning: International case studies and perspectives* (pp. 127-146). Bristol: Multilingual Matters.
- Epstein, S. (1994). Trait theory as personality theory: Can a part be as great as the whole?. *Psychological Inquiry*, 5(2), 120-122. doi:10.1207/s15327965pli0502_4
- Ezzy, D. (2002). *Qualitative analysis: Practice and innovation*. London: Routledge.
- Fallah, N. (2014). Willingness to communicate in English, communication self-confidence, motivation, shyness and teacher immediacy among Iranian English-major undergraduates: A structural equation modeling approach. *Learning and Individual Differences*, 30, 140-147. doi:10.1016/j.lindif.2013.12.006
- Fang, F. (2010). A discussion on developing students' communicative competence in college English teaching in China. *Journal of Language Teaching and Research*, 1(2), 111-116. doi:10.4304/jltr.1.2.111-116
- Finocchiaro, M., & Brumfit, C. (1983). *The functional-notional approach: From theory to practice*. Oxford: Oxford University Press.
- Fisher, C. D., & To, M. L. (2012). Using experience sampling methodology in organizational behavior. *Journal of Organizational Behavior*, 33(7), 865-877. doi:10.1002/job.1803
- Fleeson, W. (2001). Toward a structure- and process-integrated view of personality: Traits as density distributions of states. *Journal of Personality and Social Psychology*, 80(6), 1011-1027. doi:10.1037/0022-3514.80.6.1011
- Fleeson, W. (2017). The production mechanisms of traits: Reflections on two amazing decades. *Journal of Research in Personality*, 69, 4-12. doi:10.1016/j.jrp.2017.07.003
- Fleeson, W., & Jayawickreme, E. (2015). Whole trait theory. *Journal of Research in Personality*, 56, 82-92. doi:10.1016/j.jrp.2014.10.009
- Fleeson, W., & Leicht, C. (2006). On delineating and integrating the study of variability and stability in personality psychology: Interpersonal trust as illustration. *Journal of Research in Personality*, 40(1), 5-20. doi:10.1016/j.jrp.2005.08.004

- Flick, U. (2014). *An introduction to qualitative research* (5th ed.). Los Angeles; London: SAGE.
- Freiermuth, M., & Jarrell, D. (2006). Willingness to communicate: Can online chat help?. *International Journal of Applied Linguistics*, 16(2), 189-212. doi:10.1111/j.1473-4192.2006.00113.x
- Fu, L., Wang, X., & Wang, Y. (2012). The research on willingness to communicate in Chinese students' EFL study. *Journal of Language Teaching and Research*, 3(1), 112-117. doi:10.4304/jltr.3.1.112-117
- Ghanbarpour, M. (2016). Willingness to communicate, linguistic self-confidence, and language-use anxiety: The Iranian EFL context. *Theory and Practice in Language Studies*, 6(12), 2265-2271. doi:10.17507/tpls.0612.05
- Ghasemi, A. A., Kermanshahi, F. V., & Moharami, M. (2015). Conceptualizing willingness to communicate: A quantitative investigation of English-language major students. *Modern Journal of Language Teaching Methods*, 5(1), 103-110. Retrieved from <http://ezphost.dur.ac.uk/login?url=http://search.proquest.com/docview/1779884764?accountid=14533>
- Ghonsooly, B., Khajavy, G. H., & Asadpour, S. F. (2012). Willingness to communicate in English among Iranian non-English major university students. *Journal of Language and Social Psychology*, 31(2), 197-211. doi:10.1177/0261927X12438538
- Gil, J. (2016). English language education policies in the People's Republic of China. In R. Kirkpatrick (Ed.), *English language education policy in Asia* (pp. 49-90). Cham: Springer.
- Grix, J. (2001). *Demystifying postgraduate research: From MA to PhD*. Birmingham: University of Birmingham Press.
- Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, 4(1), 26-42. doi:10.1037/1040-3590.4.1.26
- Gorard, S. (2013). *Research design: Creating robust approaches for the social sciences*. London: SAGE.
- Gorard, S., & Taylor, C. (2004). *Combining methods in educational and social research*. Maidenhead: Open University Press.

- Han, J., & Yin, H. (2016). College English curriculum reform in Mainland China: Contexts, contents and changes. *Asian Education Studies*, 1(1), 1-10. doi:10.20849/aes.v1i1.9
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112. doi:10.3102/003465430298487
- Hay, C. (2002). *Political analysis: A critical introduction*. Basingstoke: Palgrave.
- Howitt, D., & Cramer, D. (2014). *Introduction to SPSS in psychology: For version 22 and earlier* (6th ed). Harlow: Pearson Education.
- Hu, G. (2002). Potential cultural resistance to pedagogical imports: The case of communicative language teaching in China. *Language Culture and Curriculum*, 15(2), 93-105. doi:10.1080/07908310208666636
- Hu, G. (2005). English language education in China: Policies, progress, and problems. *Language Policy*, 4(1), 5-24. doi:10.1007/s10993-004-6561-7
- Huang, K. (2010). Reconsideration on CLT in college English: Theory and practice. *Canadian Social Science*, 3(1), 87-90. Retrieved from <http://www.cscanada.net/index.php/css/article/view/j.css.1923669720070301.015/318>
- Huang, J. L., & Ryan, A. (2011). Beyond personality traits: A study of personality states and situational contingencies in customer service jobs. *Personnel Psychology*, 64(2), 451-488. doi:10.1111/j.1744-6570.2011.01216.x
- Hughes, J. A., & Sharrock, W. W. (1997). *The philosophy of social research* (3rd ed.). London: Longman.
- Hui, L. (2005). Chinese cultural schema of education: Implications for communication between Chinese students and Australian educators. *Issues in Educational Research*, 15(1), 17-36. Retrieved from <http://iier.waier.org.au/iier15/hui.pdf>
- James, W. (1975). *Pragmatism*. Cambridge, Mass: Harvard University Press.
- Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, 33(7), 14-26. Retrieved from <https://www.jstor.org/stable/pdf/3700093.pdf>
- Joe, H. K., Hiver, P., & Al-Hoorie, A. H. (2017). Classroom social climate, self-determined motivation, willingness to communicate, and achievement: A study of

- structural relationships in instructed second language settings. *Learning and Individual Differences*, 53, 133-144. doi:10.1016/j.lindif.2016.11.005
- Judge, T. A., Simon, L. S., Hurst, C., & Kelley, K. (2014). What I experienced yesterday is who I am today: Relationship of work motivations and behaviors to within-individual variation in the five-factor model of personality. *Journal of Applied Psychology*, 99(2), 199-221. doi:10.1037/a0034485
- Kang, S. J. (2005). Dynamic emergence of situational willingness to communicate in a second language. *System*, 33(2), 277-292. doi:10.1016/j.system.2004.10.004
- Kang, S. J. (2006). Individual and social-contextual factors affecting the learning and use of ESL: A case study of a visiting Korean physician. *TESL Canada Journal*, 24(1), 59-79. doi:10.18806/tesl.v24i1.28
- Kerlinger, F. N. (1970). A social attitude scale: Evidence on reliability and validity. *Psychological Reports*, 26(2), 379-383. doi:10.2466/pr0.1970.26.2.379
- Khajavy, G. H., Ghonsooly, B., Hosseini Fatemi, A., & Choi, C. W. (2014). Willingness to communicate in English: A microsystem model in the Iranian EFL classroom Context. *TESOL Quarterly*, 50(1), 154-180. doi:10.1002/tesq.204
- Khazaei, Z. M., Zadeh, A. M., & Ketabi, S. (2012). Willingness to communicate in Iranian EFL learners: The effect of class size. *English Language Teaching*, 5(11), 181-187. doi:10.5539/elt.v5n11p181
- King, N., & Horrocks, C. (2010). *Interviews in qualitative research*. London: SAGE.
- Larsen-Freeman, D., & Cameron, L. (2008). *Complex systems and applied linguistics*. Oxford: Oxford University Press.
- Lee, G. (2009). Speaking up: Six Korean students' oral participation in class discussions in US graduate seminars. *English for Specific Purposes*, 28(3), 142-156. doi:10.1016/j.esp.2009.01.007
- Lee, J. F. & VanPatten, B. (2003). *Making communicative language teaching happen* (2nd ed.). New York: McGraw-Hill Education.
- Li, J. (2012). *Cultural foundations of learning: East and west*. New York: Cambridge University Press.
- Liaoning Examination and Administration Service Centre. (2018). *2018 年普通高考指导* [2018 General College Entrance Examination Guide]. Liaoning: Liaoning

People's Publishing House.

- Liu, J. (2002). Negotiating silence in American classrooms: Three Chinese cases. *Language and Intercultural Communication*, 2(1), 37-54. doi:10.1080/14708470208668074
- Liu, M., & Jackson, J. (2008). An exploration of Chinese EFL learners' unwillingness to communicate and foreign language anxiety. *The Modern Language Journal*, 92(1), 71-86. doi:10.1111/j.1540-4781.2008.00687.x
- Liu, N. F., & Littlewood, W. (1997). Why do many students appear reluctant to participate in classroom learning discourse?. *System*, 25(3), 371-384. doi:10.1016/S0346-251X(97)00029-8
- Llanes, À., & Muñoz, C. (2013). Age effects in a study abroad context: Children and adults studying abroad and at home. *Language Learning*, 63(1), 63-90. doi:10.1111/j.1467-9922.2012.00731.x
- MacIntyre, P. D. (1994). Variables underlying willingness to communicate: A causal analysis. *Communication Research Reports*, 11, 135-142. doi:10.1080/08824099409359951
- MacIntyre, P. D. (2007). Willingness to communicate in the second language: Understanding the decision to speak as a volitional process. *The Modern Language Journal*, 91(4), 564-576. doi:10.1111/j.1540-4781.2007.00623.x
- MacIntyre, P. D. (2012). The idiodynamic method: A closer look at the dynamics of communication traits. *Communication Research Reports*, 29(4), 361-367. doi:10.1080/08824096.2012.723274
- MacIntyre, P. D., Baker, S. C., Clément, R., & Conrod, S. (2001). Willingness to communicate, social support, and language-learning orientations of immersion students. *Studies in Second Language Acquisition*, 23(3), 369-388. Retrieved from http://faculty.cbu.ca/pmacintyre/research_pages/journals/WTC_socialsupport_2001.pdf
- MacIntyre, P. D., Baker, S. C., Clément, R., & Donovan, L. A. (2002). Sex and age effects on willingness to communicate, anxiety, perceived competence, and L2 motivation among junior high school French immersion students. *Language Learning*, 52(3), 537-564. Retrieved from http://faculty.cbu.ca/pmacintyre/research_pages/journals/sex_age_WTC_2002.pdf

- MacIntyre, P. D., Baker, S. C., Clément, R., & Donovan, L. A. (2003). Talking in order to learn: Willingness to communicate and intensive language programs. *Canadian Modern Language Review*, 59(4), 589-608. doi:10.3138/cmlr.59.4.589
- MacIntyre, P. D., Burns, C., & Jessome, A. (2011). Ambivalence about communicating in a second language: A qualitative study of French immersion students' willingness to communicate. *The Modern Language Journal*, 95(1), 81-96. doi:10.1111/j.1540-4781.2010.01141.x
- MacIntyre, P. D., & Charos, C. (1996). Personality, attitudes, and affect as predictors of second language communication. *Journal of Language and Social Psychology*, 15(1), 3-26. doi:10.1177/0261927X960151001
- MacIntyre, P. D., Dörnyei, Z., Clément, R., & Noels, K. A. (1998). Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation. *The Modern Language Journal*, 82(4), 545-562. Retrieved from <http://www.jstor.org.ezphost.dur.ac.uk/stable/330224>
- MacIntyre, P. D., & Doucette, J. (2010). Willingness to communicate and action control. *System*, 38(2), 161-171. doi:10.1016/j.system.2009.12.013
- MacIntyre, P. D., & Gardner, R. C. (1991). Investigating language class anxiety using the focused essay technique. *The Modern Language Journal*, 75(3), 296-304. doi:10.2307/328723
- MacIntyre, P. D., & Legatto, J. J. (2011). A dynamic system approach to willingness to communicate: Developing an idiodynamic method to capture rapidly changing affect. *Applied Linguistics*, 32(2), 149-171. doi:10.1093/applin/amq037
- MacKinnon, D., & Manathunga, C. (2003). Going global with assessment: What to do when the dominant culture's literacy drives assessment. *Higher Education Research and Development*, 22(2), 131-144. doi:10.1080/07294360304110
- Mahmoodi, M. H., & Moazam, I. (2014). Willingness to communicate (WTC) and L2 achievement: The case of Arabic language learners. *Procedia-Social and Behavioral Sciences*, 98, 1069-1076. doi:10.1016/j.sbspro.2014.03.518
- Marlina, R. (2009). "I don't talk or I decide not to talk? Is it my culture?" – International students' experiences of tutorial participation. *International Journal of Educational Research*, 48(4), 235-244. doi:10.1016/j.ijer.2009.11.001

- Matthews, B., & Ross, L. (2010). *Research methods: A practical guide for the social sciences*. New York, NY: Pearson Longman.
- Mayring, P. (2004) Qualitative content analysis. In U. Flick, E. von Kardoff and I. Steinke (Eds.), *A companion to qualitative research* (pp. 266-269). London: SAGE.
- McCroskey, J. C., & Richmond, V. P. (1990). Willingness to communicate: A cognitive view. *Journal of Social Behavior and Personality*, 5(2), 19-37. Retrieved from <http://ezphost.dur.ac.uk/login?url=https://search-proquest-com.ezphost.dur.ac.uk/docview/1292258287?accountid=14533>
- Mertens, D. M. (2010). *Research and evaluation in education and psychology: Integrating diversity with quantitative, qualitative, and mixed methods* (3rd ed.). Thousand Oaks, CA: Sage publications.
- Mitchell, R. (1988). *Communicative language teaching in practice*. London: Centre for Information on Language Teaching and Research.
- Minbashian, A., Beckmann, N., & Wood, R. E. (2017). Emotional intelligence and individual differences in affective processes underlying task-contingent conscientiousness. *Journal of Organizational Behavior*. doi.org/10.1002/job.2233
- Minbashian, A., Wood, R. E., & Beckmann, N. (2010). Task-contingent conscientiousness as a unit of personality at work. *Journal of Applied Psychology*, 95(5), 793-806. doi:10.1037/a0020016
- Ministry of Education, Culture, Sports, Science, Technology (MEXT). (2011). *Five proposals and specific measures for developing proficiency in English for international communication*. Retrieved from http://www.mext.go.jp/component/english/_icsFiles/afieldfile/2012/07/09/1319707_1.pdf
- Ministry of Education of the People's Republic of China (MOE). (2001). *教育部关于积极推进小学开设英语课程的指导意见* [Guidelines for promoting English courses in primary schools]. Retrieved from http://www.moe.gov.cn/s78/A26/jces_left/moe_714/tnull_665.html
- Ministry of Education of the People's Republic of China (MOE). (2003). *普通高中英语课程标准* [English curriculum standards for senior high schools]. Retrieved from http://www.moe.gov.cn/srcsite/A08/s7056/200707/t20070710_110825.html

- Ministry of Education of the People's Republic of China (MOE). (2007). *大学英语课程教学要求* [College English curriculum requirements]. Beijing: People's Education Press.
- Mischel, W., & Shoda, Y. (1998). Reconciling processing dynamics and personality dispositions. *Annual Review of Psychology*, 49(1), 229-258. doi:10.1146/annurev.psych.49.1.229
- Munezane, Y. (2013). Attitudes, affect and ideal L2 self as predictors of willingness to communicate. In R. Leah, A. Ewert, M. Pawlak, & M. Wrembel (Eds.), *EUROSLA Yearbook: Volume 13* (pp. 176-198). Amsterdam: John Benjamins.
- Munezane, Y. (2015). Enhancing willingness to communicate: Relative effects of visualization and goal setting. *The Modern Language Journal*, 99(1), 175-191. doi:10.1111/modl.12193
- Mystkowska-Wiertelak, A. (2016). Dynamics of classroom WTC: Results of a semester study. *Studies in Second Language Learning and Teaching*, 6(4), 651-676. doi:10.14746/ssl.t.2016.6.4.5
- Mystkowska-Wiertelak, A., & Pawlak, M. (2014). Fluctuations in learners' willingness to communicate during communicative task performance: Conditions and tendencies. *Research in Language*, 12(3), 245-260. doi:10.2478/rela-2014-0019
- Mystkowska-Wiertelak, A., & Pawlak, M. (2017). *Willingness to communicate in instructed second language acquisition: Combining a macro- and micro-perspective*. Bristol: Multilingual Matters.
- Naderifar, A. A., & Esfandiari, R. (2016, July). Relationships between Iranian intermediate EFL learners' willingness to communicate, oral proficiency, autonomy, and ambiguity tolerance. *International Journal of Humanities and Cultural Studies*, 932-951. Retrieved from <http://ijhcs.com/index.php/ijhcs/article/view/2551>
- Nezlek, J. B. (2001). Multilevel random coefficient analyses of event-and interval-contingent data in social and personality psychology research. *Personality and Social Psychology Bulletin*, 27(7), 771-785. doi:10.1177/0146167201277001
- Nijhuis, J., Segers, M., & Gijssels, W. (2007). The interplay of perceptions of the learning environment, personality and learning strategies: A study amongst international business studies students. *Studies in Higher Education*, 32(1), 59-77. doi:10.1080/03075070601099457

- Ntalianis, F. (2010). Do personality and learning climate predict competence for learning? An investigation in a Greek academic setting. *Learning and Individual Differences*, 20(6), 664-668. doi:10.1016/j.lindif.2010.08.003
- Ortega, L. (2009). *Understanding second language acquisition*. London: Hodder Education.
- O'Sullivan, B. (2002). Learner acquaintanceship and oral proficiency test pair-task performance. *Language Testing*, 19(3), 277-295. doi:10.1191/0265532202lt205oa
- Oz, H. (2014). Big five personality traits and willingness to communicate among foreign language learners in Turkey. *Social Behavior and Personality: An International Journal*, 42(9), 1473-1482. doi:10.2224/sbp.2014.42.9.1473
- Öz, H., Demirezen, M., & Pourfeiz, J. (2015). Willingness to communicate of EFL learners in Turkish context. *Learning and Individual Differences*, 37, 269-275. doi:10.1016/j.lindif.2014.12.009
- Parrigon, S., Woo, S. E., Tay, L., & Wang, T. (2017). CAPTION-ing the situation: A lexically-derived taxonomy of psychological situation characteristics. *Journal of Personality and Social Psychology*, 112(4), 642-681. doi:10.1037/pspp0000111
- Pawlak, M., & Mystkowska-Wiertelak, A. (2015). Investigating the dynamic nature of L2 willingness to communicate. *System*, 50, 1-9. doi:10.1016/j.system.2015.02.001
- Pawlak, M., Mystkowska-Wiertelak, A., & Bielak, J. (2016). Investigating the nature of classroom willingness to communicate (WTC): A micro-perspective. *Language Teaching Research*, 20(5), 654-671. doi:10.1177/1362168815609615
- Pawłowska, D. K., Westerman, J. W., Bergman, S. M., & Huelsman, T. J. (2014). Student personality, classroom environment, and student outcomes: A person-environment fit analysis. *Learning and Individual Differences*, 36, 180-193. doi:10.1016/j.lindif.2014.10.005
- Peng, J. E. (2007a). Willingness to communicate in an L2 and integrative motivation among college students in an intensive English language program in China. *University of Sydney Papers in TESOL*, 2(1), 33-59. Retrieved from <https://pdfs.semanticscholar.org/4467/b8d102a7efa26cc1e13227016d4b6c7953e0.pdf>
- Peng, J. E. (2007b). Willingness to communicate in the Chinese EFL classroom: A cultural perspective. In J. Liu (Ed.), *English language teaching in China: New*

approaches, perspectives, and standards (pp. 250-269). London: Continuum.

- Peng, J. E. (2012). Towards an ecological understanding of willingness to communicate in EFL classrooms in China. *System*, 40(2), 203-213. doi:10.1016/j.system.2012.02.002
- Peng, J. E. (2014). *Willingness to communicate in the Chinese EFL university classroom: An ecological perspective*. Bristol: Multilingual Matters.
- Peng, J. E. (2015). L2 motivational self system, attitudes, and affect as predictors of L2 WTC: An imagined community perspective. *The Asia-Pacific Education Researcher*, 24(2), 433-443. doi:10.1007/s40299-014-0195-0
- Peng, J. E., & Woodrow, L. (2010). Willingness to communicate in English: A model in the Chinese EFL classroom context. *Language Learning*, 60(4), 834-876. doi:10.1111/j.1467-9922.2010.00576.x
- Peng, J. E., Zhang, L., & Chen, Y. (2017). The mediation of multimodal affordances on willingness to communicate in the English as a foreign language classroom. *TESOL Quarterly*, 51(2), 302-331. doi:10.1002/tesq.298
- Piechurska-Kuciel, E. (2018). Openness to experience as a predictor of L2 WTC. *System*, 72, 190-200. doi:10.1016/j.system.2018.01.001
- Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N.P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879-903. doi:10.1037/0021-9010.88.5.879
- Powell, R. G., & Powell, D. L. (2010). *Classroom communication and diversity: Enhancing instructional practice*. New York: Routledge.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: SAGE.
- Raudenbush, S. W., Bryk, A. S., Cheong, Y., & Congdon, R. T. (2000). *HLM 6: Hierarchical linear and nonlinear modelling*. Chicago: Scientific Software International.
- Rauthmann, J. F., Gallardo-Pujol, D., Guillaume, E. M., Todd, E., Nave, C. S., Sherman, R. A., ... Funder, D. C. (2014). The Situational Eight DIAMONDS: A taxonomy of major dimensions of situation characteristics. *Journal of Personality and Social*

Psychology, 107(4), 677-718. doi:10.1037/a0037250

- Rauthmann, J. F., & Sherman, R. A. (2016a). Situation change: Stability and change of situation variables between and within persons. *Frontiers in Psychology*, 6, 1938. doi:10.3389/fpsyg.2015.01938
- Rauthmann, J. F., & Sherman, R. A. (2016b). Ultra-brief measures for the Situational Eight DIAMONDS domains. *European Journal of Psychological Assessment*, 32(2), 165-174. doi:10.1027/1015-5759/a000245
- Rauthmann, J. F., Sherman, R. A., & Funder, D. C. (2015). Principles of situation research: Towards a better understanding of psychological situations. *European Journal of Personality*, 29, 363-381. doi:10.1002/per.1994
- Riasati, M. J. (2012). EFL learners' perception of factors influencing willingness to speak English in language classrooms: A qualitative study. *World Applied Sciences Journal*, 17(10), 1287-1297. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.390.1430&rep=rep1&type=pdf>
- Richards, L., & Morse, J. M. (2007). *Readme first for a user's guide to qualitative methods* (2nd ed.). London: SAGE.
- Ryan, G. (1999). Measuring the typicality of text: Using multiple coders for more than just reliability and validity checks. *Human Organization*, 58(3), 313-322. doi:10.17730/humo.58.3.g224147522545rln
- de Saint Léger, D., & Storch, N. (2009). Learners' perceptions and attitudes: Implications for willingness to communicate in an L2 classroom. *System*, 37(2), 269-285. doi:10.1016/j.system.2009.01.001
- Savignon, S. J. (2005). Communicative language teaching: Strategies and goals. In E. Hinkel. (Ed.), *Handbook of research in second language teaching and learning* (pp. 635-651). Mahwah, NJ: L. Erlbaum Associates.
- Schmitz, B. (2006). Advantages of studying processes in educational research. *Learning and Instruction*, 16(5), 433-449. doi:10.1016/j.learninstruc.2006.09.004
- Sherman, R. A., Rauthmann, J. F., Brown, N. A., Serfass, D. G., & Jones, A. B. (2015). The independent effects of personality and situations on real-time expressions of behavior and emotion. *Journal of Personality and Social Psychology*, 109(5), 872-888. doi:10.1037/pspp0000036

- Shi, L. (2006). The successors to Confucianism or a new generation? A questionnaire study on Chinese students' culture of learning English. *Language, Culture and Curriculum, 19*(1), 122-147. doi:10.1080/07908310608668758
- Shoda, Y., Mischel, W., & Wright, J. C. (1994). Intraindividual stability in the organization and patterning of behaviour: Incorporating psychological situations into the idiographic analysis of personality. *Journal of Personality and Social Psychology, 67*(4), 674-687. doi:10.1037/0022-3514.67.4.674
- Song, D. (2000). English teaching in China. *English Teacher: An International Journal, 4*(1), 31-39. Retrieved from http://www.aulibrary.au.edu/multim1/ABAC_Pub/The-English-Teacher-An-International-Journal/v4-n1-3.pdf
- Sprague, J. (1992). Expanding the research agenda for instructional communication: Raising some unasked questions. *Communication Education, 41*(1), 1-25. doi:10.1080/03634529209378867
- Sun, Y. F. (2008). *Motivation to speak: Perception and attitude of non-English major students in Taiwan*. Unpublished PhD dissertation, Indiana University. Retrieved from <https://search.proquest.com/docview/287953181?pq-origsite=gscholar>
- Tien, C. Y., & Barnard, R. (2009). Interaction in a Taiwanese primary school English classroom (task 1). In R. Barnard, & M. E. Torres-Guzmán (Eds.), *Creating classroom communities of learning: International case studies and perspectives* (pp. 88-108). Bristol: Multilingual Matters.
- Tsai, S. C. (2017). *East Asian Canadian graduated students, silent but engaged: A dialectical exploration of classroom learner speech in competing educational models*. Unpublished PhD dissertation, University of Toronto. Retrieved from https://tspace.library.utoronto.ca/bitstream/1807/79511/3/Tsai_Sh-Chen_201706_PhD_thesis.pdf
- de Vaus, D. A. (2001). *Research design in social research*. London: SAGE.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*(6), 1063-1070. doi:10.1037/0022-3514.54.6.1063

- Wen, W. P., & Clément, R. (2003). A Chinese conceptualisation of willingness to communicate in ESL. *Language Culture and Curriculum*, 16(1), 18-38. doi:10.1080/07908310308666654
- White, P. (2017). *Developing research questions* (2nd ed.). London: Palgrave.
- Wilson, R. E., Thompson, R. J., & Vazire, S. (2017). Are fluctuations in personality states more than fluctuations in affect?. *Journal of Research in Personality*, 69, 110-123. doi:10.1016/j.jrp.2016.06.006
- Wolf, J. P. (2013). Exploring and contrasting EFL learners' perceptions of textbook-assigned and self-selected discussion topics. *Language Teaching Research*, 17(1), 49-66. doi:10.1177/1362168812457535
- Wood, R. E., Beckmann, N., Birney, D. P., Beckmann, J. F., Minbashian, A., & Chau, R. (2018). Situation contingent units of personality at work. *Personality and Individual Differences*. doi:10.1016/j.paid.2018.01.026
- Xiao, L. (2011). Communicative language teaching in a Chinese university context: Beliefs and practice. *Chinese Journal of Applied Linguistics*, 34(2), 44-61. Retrieved from <http://www.celea.org.cn/teic/96/4.pdf>
- Yashima, T. (2002). Willingness to communicate in a second language: The Japanese EFL context. *The Modern Language Journal*, 86(1), 54-66. doi:10.1111/1540-4781.00136
- Yashima, T. (2009). International posture and the ideal L2 self in the Japanese EFL context. In Z. Dörnyei, & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 144-192). Bristol: Multilingual Matters.
- Yashima, T., & Zenuk-Nishide, L. (2008). The impact of learning contexts on proficiency, attitudes, and L2 communication: Creating an imagined international community. *System*, 36(4), 566-585. doi:10.1016/j.system.2008.03.006
- Yashima, T., Zenuk-Nishide, L., & Shimizu, K. (2004). The influence of attitudes and affect on willingness to communicate and second language communication. *Language Learning*, 54(1), 119-152. doi:10.1111/j.1467-9922.2004.00250.x
- Yu, L. (2001). Communicative language teaching in China: Progress and resistance. *TESOL Quarterly*, 35(1), 194-198. Retrieved from <http://www.jstor.org/stable/pdf/3587868.pdf?refreqid=excelsior%3Ad4d71116bb272a5ce93de2f7c131a51a>

- Yu, M. (2015). An examination of the dynamic feature of WTC through dyadic group interaction. *System*, 55, 11-20. doi:10.1016/j.system.2015.08.001
- Zarrinabadi, N. (2014). Communicating in a second language: Investigating the effect of teacher on learners' willingness to communicate. *System*, 42, 288-295. doi:10.1016/j.system.2013.12.014
- Zhang, J., Beckmann, N., & Beckmann, J. F. (2018). To talk or not to talk: A review of situational antecedents of willingness to communicate in the second language classroom. *System*, 72, 226-239. doi:10.1016/j.system.2018.01.003
- Zhong, Q. M. (2013). Understanding Chinese learners' willingness to communicate in a New Zealand ESL classroom: A multiple case study drawing on the theory of planned behaviour. *System*, 41(3), 740-751. doi:10.1016/j.system.2013.08.001

Appendix A An Overview of Relevant Studies on State WTC

Authors	Participants	Design	Measures	Relevant situational variables
Bernales (2016)	4 German-as-foreign-language learners in a Midwestern university in the US	Longitudinal (15 weeks)	Class observation/videotaping Stimulated recall interview	L2 speaking goals Confidence in L2 skills Activity and topic Teacher's expectations
Buckingham & Alpaslan (2017)	40 Turkish young learners of English	Experimental	Class observation	Asynchronous audio-visual speaking activities
Cameron (2013)	3 Iranian ESL learners in a New Zealand university	Cross-sectional	Questionnaire Interview Teacher report	Teaching methods & approaches Teacher support
Cao (2011)	12 ESL learners of various nationalities in a university-based language school in New Zealand	Longitudinal (20 weeks)		
Cao (2013)	12 ESL learners (mainly from China or Korea) in a university-based language school in New Zealand	Longitudinal (5 months)	Class observation Stimulated recall interview Reflective journal	Topic: content knowledge, familiarity, interest & sensitivity Task type: opportunities to talk & contribution to progress Interlocutor: familiarity, language proficiency, personality, nationality, participation & cooperation Teacher support and immediacy & teaching style Group size: dyadic, group or whole-class activity
Cao (2014)	6 Chinese ESL learners in a university-based language school in New Zealand	Longitudinal (5 months)		
Cao & Philp (2006)	8 ESL learners of various nationalities in a university-based language school in New Zealand	Longitudinal (1 month)	Questionnaire Class observation Audio record Interview	Group size: pair, group or whole-class activity Confidence Interlocutor: familiarity & participation Topic: familiarity

Eddy-U (2015)	25 Chinese EFL learners in two universities in Macau	Cross-sectional	Focus group interview	<p>Perception of the interlocutors</p> <p>Group members: familiarity, talkativeness, motivation, participation, L2 proficiency & gender</p> <p>Classroom atmosphere: the teacher & classmates</p> <p>Perception of the task: interest (topic & type of activity), effectiveness & difficulty</p> <p>State motivation: marks</p> <p>Confidence</p>
Fallah (2014)	252 Iranian English-major university students	Cross-sectional	Questionnaire	Teacher immediacy: students' motivation & security
Freiermuth & Jarrell (2006)	36 English learners in a university in Japan	Experimental	Questionnaire Task performance (discourse)	Online chatting vs. face-to-face mode: anxiety & attractive
Ghasemi et al. (2015)	137 English-major students in Iran	Cross-sectional	Questionnaire	Task type
Joe et al. (2017)	381 Korean secondary school EFL learners	Cross-sectional	Questionnaire	Classroom social climate: teacher emotional support, teacher academic support & classroom mutual respect
Kang (2005)	4 Korean ESL learners in an English Language Institute in the northeast of the US	Longitudinal (8 weeks)	Interview Video & audio record Stimulated recall	<p>Security</p> <ul style="list-style-type: none"> • Interlocutors: language proficiency, nationality, familiarity, number & support • Topic: background knowledge • Conversational context: stage in a conversation & when facing difficulties <p>Excitement</p> <ul style="list-style-type: none"> • Topic: interest, personal experiences & background knowledge • Interlocutors: nationality, appearance & support • Conversational context: when asked for additional information <p>Responsibility</p> <ul style="list-style-type: none"> • Topic: perceived usefulness and importance, background knowledge & sensitivity • Interlocutors: number & support • Conversational context: when misunderstood

Kang (2006)	1 Korean physician in the US	Longitudinal (13 months)	Observation in various situations Informal conversation	Interlocutor: native or non-native speaker Insecurity
Khajavy et al. (2014)	243 English-major students in Iran	Cross-sectional	Questionnaire	Teacher support Student cohesiveness Task orientation:
Khazaei et al. (2012)	30 adult Iranian EFL learners in the same institute in Iran	Experimental	Class observation (talk time & turn-taking)	Class size
Lee (2009)	6 Korean graduates students in a university in the southwest of the US	Longitudinal (1 semester)	Interview Class observation Informal conversation	Perception of teachers and classmates Whole-class or small group discussion
Liu (2002)	3 Chinese graduate students in a university in the Midwestern of the US	Longitudinal (1 year)	Interview Class observations Prolonged engagement with the participants	Security & self-protection
Liu & Littlewood (1997)	2,156 Chinese EFL learners & 437 lecturers in a university in Hong Kong	Cross-sectional	Questionnaire	Teaching style Confidence & anxiety
MacIntyre et al. (2011)	100 Canadian junior high school students in a French immersion program	Longitudinal (6 weeks)	Questionnaire Diary	Context Interlocutors
MacIntyre & Legatto (2011)	6 Canadian learners of French in universities	Short-term longitudinal (high-density repeated measurements within 8 tasks)	Self-rated WTC per second Stimulated recall Observation	Topic familiarity & vocabulary retrieval

Mystkowska-Wiertelak (2016)	12 English-major undergraduates in Poland	Longitudinal (high-density repeated measurements within different lessons throughout a semester)	Self-rated WTC every 5min Questionnaire Interview	Class-arrangement modes: pair, group or whole-class (security & pleasure) Interlocutor: familiarity, language proficiency, reaction & personality Topic interest & familiarity: vocabulary & knowledge Activity type & variety Stage of the class: the beginning, middle or end of the class (interest)
Mystkowska-Wiertelak & Pawlak (2014)	44 English-major undergraduates in two institutions of higher education in Poland	Short-term longitudinal (high-density repeated measurements within 2 tasks)	Self-rated WTC every 30s Questionnaire	Task type: monologue or dialogue Stage of the task: the trend from beginning to the end
Pawlak & Mystkowska-Wiertelak (2015)	8 English-major undergraduates in an institutions of higher education in Poland	Short-term longitudinal (high-density repeated measurements within a task)	Self-rated WTC every 30s Questionnaire Stimulated recall	Topic: interest, content knowledge & vocabulary retrieval Time for preparation Interlocutor: familiarity, involvement & cooperation Presence of the teacher
Pawlak et al. (2016)	60 English-major undergraduates in Poland	Short-term longitudinal (high-density repeated measurements within a lesson)	Self-rated WTC every 5min Questionnaire Lesson plan The teacher's comment	Pair, small group or whole-class activity Interlocutors: familiarity & proficiency The teacher: classroom arrangement, teaching style, personality, enthusiasm & rapport with the students Topic: personal experience & interest Activity type: game Stage of the class: beginning or end
Peng (2007b)	118 Chinese university students	Cross-sectional	Questionnaire Group interview Diary	Group cohesiveness Teacher support, teaching styles & classroom management
Peng (2012)	4 EFL learners in a university in southern China	Longitudinal (1.5 semesters)	Interview Class observation Learning journal	Classroom atmosphere Teacher support & teaching style Task: interest, usefulness & importance

Peng & Woodrow (2010)	579 non-English-major undergraduates in 8 universities in eastern China	Cross-sectional	Questionnaire	Teacher support Student cohesiveness Task orientation: importance & usefulness
Peng et al. (2017)	4 non-English-major students in a university in China	Two scenarios from a same class period	Stimulated recall Learning journal Scenarios transcription & annotation	The teacher's pedagogic discourse: language, gesture & gaze
Riasati (2012)	7 Iranian EFL learners in a private language institute	Cross-sectional	Interview	Interlocutor: gender, age, familiarity & participation Task type: individually, in pairs or groups Graded or not Confidence Topic: familiarity, interest & preparation Teacher attitude & teaching style Classroom atmosphere: students & the teacher
de Saint Léger & Storch (2009)	32 advanced learners of French in an Australian university	Longitudinal (12 weeks)	Questionnaire Focus group interview Teacher assessment	Source of difficulty: fluency & vocabulary Lack of confidence/anxiety Whole-class or small group discussion: communication opportunities
Wolf (2013)	101 EFL learners in a university in Japan	Cross-sectional	Questionnaire	Topic: interest & learner's knowledge (related to self-confidence)
Yu (2015)	18 English-major students in a university in China	Experimental	Questionnaire Task performance (number of words & turn-taking)	Interlocutor's WTC (in dyadic interactions)
Zarrinabadi (2014)	50 English-major undergraduates in Iran	Longitudinal (6 weeks)	Focused essay	Teacher: wait time, decision on topic, error correction & support
Zhong (2013)	5 Chinese ESL learners in a language school in New Zealand	Longitudinal (18 weeks)	Interview Learning log Class observation Stimulated recall	Teacher-fronted or collaborative learning activity Time for preparation

Appendix B Consent Form for Study 1



Shaped by the past, creating the future

30th January 2016

Participant Information Sheet

Title: Chinese EFL Learners' Willingness to Communicate

You are invited to take part in a research study of students' willingness to communicate in English. Please read this form carefully and ask any questions you may have before agreeing to participate in the study.

The study is conducted by Jiayi Zhang, as part of her doctoral studies at Durham University. This research project is supervised by Dr. Nadin Beckmann (nadin.beckmann@durham.ac.uk) and Prof. Jens F. Beckmann (j.beckmann@durham.ac.uk), School of Education, Durham University, UK.

The purpose of this study is to better understand learners' willingness to communicate in class, so that teachers can better support students in their learning.

There are two parts to this study. In part 1, you will be asked to fill in a questionnaire. It will take you approximately 10 minutes to complete this questionnaire. In part 2, you will be asked to answer a small set of questions during or at the end of an English class. It will take you about 5 minutes to complete this questionnaire each time.

You are free to decide whether or not to participate. If you decide to participate, you are free to withdraw at any time without any negative consequences for you.

All responses you give or other data collected will be kept confidential. The records of this study will be kept secure and private. All files containing any information you give are password protected. In any research report that may be published, no information will be included that will make it possible to identify you individually.

If you have any questions, requests or concerns regarding this research, please contact me via email at Jiayi Zhang, jiayi.zhang@durham.ac.uk or by telephone at +44 (0)7541022080.

This study has been reviewed and approved by the School of Education Ethics Sub-Committee at Durham University (date of approval: 15th December 2015)

Jiayi Zhang

Leazes Road
Durham City, DH1 1TA
Telephone +44 (0)191 334 2000 Fax +44 (0)191 334 8311
www.durham.ac.uk
Durham University is the trading name of the University of Durham

Declaration of Informed Consent

- I agree to participate in this study, the purpose of which is to better understand learners' willingness to communicate in class, so that teachers can better support students in their learning.
- I have read the participant information sheet and understand the information provided.
- I have been informed that I may decline to answer any questions or withdraw from the study without penalty of any kind.
- I have been informed that all of my responses will be kept confidential and secure, and that I will not be identified in any report or other publication resulting from this research.
- I have been informed that the investigator will answer any questions regarding the study and its procedures. Jiayi Zhang, School of Education, Durham University can be contacted via email: jiayi.zhang@durham.ac.uk or by telephone: +44 (0)7541022080.
- I will be provided with a copy of this form for my records.

Any concerns about this study should be addressed to the Ethics Sub-Committee of the School of Education, Durham University (email: ed.ethics@durham.ac.uk or telephone: +44 (0)1913348403).

Date

Participant Signature

I certify that I have presented the above information to the participant and secured his or her consent.

Date

Signature of Investigator

Appendix C Baseline Self-Report Questionnaire for Study 1

Introduction to the Study

The purpose of this study is to better understand learners' willingness to communicate in class, so that teachers can better support students in their learning. There are two parts to this study. In part 1, you will be asked to fill in a questionnaire. In this questionnaire, we are interested in how you see yourself as a learner. There is no right or wrong answer. It will take you approximately 10 minutes to complete this questionnaire.

In part 2, you will be asked to answer a small set of questions either during or at the end of an English class. You will be asked to answer the questions for multiple times. It will take you about 5 minutes to complete this questionnaire each time. Here we are interested in how you think and feel about yourself and your experience IN CLASS AT THIS POINT IN TIME.

Please keep in mind your responses will be absolutely confidential. In order to maintain your anonymity, we will use a confidential code instead of your name.

To create a code, please think of the first two letters of your mother's first name, your date of birth, the first two letters of your father's first name, and your gender (F for female and M for Male). For example, if my mother's FIRST NAME is Feng, my date of birth is 13th, my father's FIRST NAME is Zhenqiu, and I am a female, then my confidential code will be FE13ZHF. Please write down your confidential code in the following boxes and keep it in mind.

□□□□□□□□

Demographics

Instructions: In the following section, we will ask you to provide some demographic information about yourself.

Your programme/subject of study (i.e. major)

Your current year of study on this programme

Year 1 Year 2 Year 3 Year 4

Your age

Your level of class

Advanced class Regular class

Your current English score (i.e. final English score for last semester)

□□□□/100.

Describe Yourself

Instructions: In the following section, we are interested in how you would describe yourself. Describe yourself as you **GENERALLY** are now, not as you wish to be in the future. Describe yourself as you honestly see yourself, in relation to other people you know of the same sex as you are, and roughly your same age. Please keep in mind, your responses will be anonymous, so you can describe yourself in an honest manner. For each of the items, please indicate the level of accuracy that describes you as a person and mark the box.

Very Inaccurate	Inaccurate	Moderately Inaccurate	Neither Accurate nor Inaccurate	Moderately Accurate	Accurate	Very Accurate
---------------------------	-------------------	---------------------------------	---	------------------------	----------	------------------

I ...

	Very Inaccurate						Very Accurate
am the life of the party.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
feel little concern for others.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
am always prepared.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
get stressed out easily.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
have a rich vocabulary.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
don't talk a lot.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
am interested in people.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
leave my belongings around.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
am relaxed most of the time.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
have difficulty understanding abstract ideas.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
feel comfortable around people.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
insult people.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
pay attention to details.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
worry about things.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
have a vivid imagination.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
keep in the background.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
sympathise with others' feelings.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
make a mess of things.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
seldom feel blue.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
am not interested in abstract ideas.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
start conversations.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
am not interested in other people's problems.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

	Very Inaccurate					Very Accurate	
get chores done right away.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
am easily disturbed.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
have excellent ideas.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
have little to say.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
have a soft heart.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
often forget to put things back in their proper place.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
get upset easily.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
do not have a good imagination.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
talk to a lot of different people at parties.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
am not really interested in others.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
like order.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
change my mood a lot.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
am quick to understand things.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
don't like to draw attention to myself.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
take time out for others.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
shirk my duties.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
have frequent mood swings.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
use difficult words.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
don't mind being the centre of attention.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
feel others' emotions.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
follow a schedule.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
get irritated easily.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
spend time reflecting on things.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
am quiet around strangers.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
make people feel at ease.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
am exacting in my work.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
often feel blue.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
am full of ideas.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Willingness to Communicate in Mandarin

Instructions: Below are some situations in which a person might choose to communicate. Presume that you have completely free choice. Please indicate how willing you would be to communicate in each type of situation. For each of the items, please indicate the level of accuracy that describes your response and mark the box. Here we are interested in how willing you GENERALLY are to communicate.

Stranger

I AM WILLING TO ...

	Very Inaccurate				Very Accurate			
talk to a shop assistant.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
speak in public to a group of strangers (about 30 people).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
talk with a stranger on campus.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
talk in a small group of strangers (about five people).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
talk with a waiter/waitress in a restaurant.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
talk with a stranger while standing in line.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
talk in a large meeting of strangers (about 10 people).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Acquaintance

I AM WILLING TO ...

	Very Inaccurate				Very Accurate			
talk when I happen to meet an acquaintance.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
talk in a large meeting of acquaintances (about 10 people).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
talk to a teacher after class.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
talk with an acquaintance while standing in line.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
talk in a small group of acquaintances (about five people).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
talk with support staff (e.g. tutor, admin, librarian, porter, etc.).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
speak in public to a group of acquaintances (about 30 people).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Friend

I AM WILLING TO ...

	Very Inaccurate				Very Accurate			
talk in a large meeting of friends (about 10 people).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
talk with a friend while standing in line.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
talk with a fellow student when engaging in extracurricular activities.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
speak in public to a group of friends (about 30 people).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

talk with one of my roommates.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
talk in a small group of friends (about five people).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
talk with a fellow student sitting next to me in class.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Classroom activity

I AM WILLING TO ...

	Very In accurate						Very Accurate
volunteer an answer when the teacher asks a question in class.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
ask a question in class.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
present my own opinions in class.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
participate in group discussions in class.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
help others answer a question in class.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

If you have any other comments, please feel free to write them down in the following box.

Thank you for your participation!

Willingness to Communicate in English

Instructions: The following statements describe some communicative situations during as well as outside an English class. Please indicate how willing you are to engage in these communication activities USING ENGLISH. For each of the items, please indicate the level of accuracy that describes your response and mark the box. Here we are interested in how willing you GENERALLY are to communicate IN ENGLISH during as well as outside the English class that you have experienced in this university.

Very Inaccurate	Inaccurate	Moderately Inaccurate	Neither Accurate nor Inaccurate	Moderately Accurate	Accurate	Very Accurate
---------------------------	-------------------	---------------------------------	---	------------------------	----------	------------------

Speaking

I AM WILLING TO ...

	Very Inaccurate						Very Accurate
participate in a dialogue in English at my desk with my neighbour.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
ask the teacher a question in English.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
do a short presentation in English to the class with notes.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
do a role-play standing in front of the class in English.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
ask my neighbour in English how to pronounce a word in English.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
ask my neighbour in English how to express my thoughts in English.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
ask my neighbour in English the meaning of an English word.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
give a short self-introduction in English to the class without notes.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
volunteer an answer in English when the teacher asks a question.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
help others answer a question in English.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
participate in group discussions in English.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
present my own opinions in English to the class.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
participate in the English activities outside the classroom (e.g. English speaking contest, English corner, English imitation show, etc.).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
read out a paragraph in English to the class.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
translate a spoken utterance from Chinese into English.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Writing

I AM WILLING TO ...

	Very Inaccurate				Very Accurate			
write a CV or personal statement in English (e.g. to apply for an internship online).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
do a structured writing task in English from the textbook.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
write a short report in English on an article or book I read.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
write a story in English.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
write a piece of status or a comment in English on SNS (e.g. Weibo, WeChat, Renren, QQ, etc.).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
write a greeting card or short message in English.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
write down a list in English of homework I must do.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
write answers in English to the exercises from the textbook.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
write a narration in English (e.g. about a Chinese event, my hometown, university life in China, etc.).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
write an argumentation in English (e.g. agreeing or disagreeing with a person's point of view, describing the cause and effect of something, etc.).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
write a diary about my daily life in English.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
write a self-introduction in English.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
translate a piece of writing from Chinese into English.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

Reading

I AM WILLING TO ...

	Very Inaccurate				Very Accurate			
read a novel in English.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
read a newspaper article in English.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
read a piece of status or a comment in English on SNS (e.g. Weibo, WeChat, Renren, QQ, etc.).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
read an article in English from the textbook.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
read an advertisement in English (e.g. to find an internship opportunity online).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
read reviews in English for popular movies.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
watch a movie/TV series in English.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
change the language settings on some of my mobile devices into English (e.g. mobile phone, pad, laptop etc).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

Classroom Support

Instructions: The following statements describe some characteristics of English classes. Please indicate how you think and feel about your experience of the College English classes at this university. For each of the item, please indicate the level of accuracy that describes your response and mark the box. Here we are interested in how the College English classes that you have experienced at this university GENERALLY are.

Teacher

	Very Inaccurate				Very Accurate			
The teacher provides a timely response to students' concerns.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The teacher is patient in teaching.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The teacher is supportive during the class.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The teacher asks questions that solicit viewpoints or opinions.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The teacher provides students with enough time to prepare for their responses or reactions.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The teacher interrupts students to correct their errors when students are speaking or writing in English.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The teacher negotiates with students about the topic for the next session.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Students

	Very Inaccurate				Very Accurate			
In my English class, I work well with students.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
In my English class, I am friendly to students.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
In my English class, I help students who are having difficulties with their work.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
In my English class, I trust students.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
In my English class, the classroom climate is active.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
In my English class, students are supportive.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
In my English class, I make friends with students.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Tasks

	Very Inaccurate				Very Accurate			
In my English class, tasks designed are useful.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
In my English class, tasks designed are interesting.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
In my English class, I know what I am trying to accomplish.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
In my English class, activities are carefully planned.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
In my English class, instructions for activities are clear so everyone knows what to do.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

In my English class, I am familiar with the topics discussed.

In my English class, tasks designed are challenging.

If you have any other comments, please feel free to write them down in the following box.

Thank you for your participation!

Appendix D Momentary Self-Report Questionnaire for Study 1

Code:

In this part, we would like to know how you think and feel about yourself and your experience IN THIS CLASS AT THIS POINT IN TIME. When answering the questions, please think about an activity when you were given the chance to communicate IN ENGLISH with your teacher or peers, either orally or in writing. Please mark the box that best describes your response in relation to THAT SPECIFIC ACTIVITY.

Not at all	A little	Somewhat	Moderately	Quite	Very	Extremely
------------	----------	----------	------------	-------	------	-----------

The communication activity

	Not at all						Extremely
The activity was important for me.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The activity was difficult for me.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The activity was interesting for me.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The class was active.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
My teacher was supportive.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
My classmates were supportive.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I was familiar with the people I interacted with.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The English proficiency of the people I interacted with was higher than mine.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The people I interacted with were cooperative.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
There were sufficient opportunities to communicate.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I performed well during the activity.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

The class situation

	Not at all						Extremely
Work had to be done.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Deep thinking was required.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Somebody was being threatened, accused, or criticised.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Potential romantic partners were present.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The situation was pleasant.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The situation contained negative feelings (e.g. stress, anxiety, shame).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Social interactions were expected.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Your willingness to communicate in English

	Not at all						Extremely
I was willing to communicate in English in the activity.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I did communicate in English in the activity.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Your thoughts and feelings about yourself during the activity

I WAS ...

	Not at all				Extremely			
talkative.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
cooperative.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
organised.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
intelligent.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
energetic.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
trustful.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
insecure.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
assertive.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
hardworking.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
optimistic.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
inquisitive.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
warm.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
responsible.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
vulnerable.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
creative.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
excited.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
afraid.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
inspired.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
upset.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
proud.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
nervous.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
attentive.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
ashamed.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
interested.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
irritable.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
confident.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

If you have any other comments, please feel free to write them down in the following box.

Thank you for your participation!

Appendix E Consent Form for Study 2



Shaped by the past, creating the future

Participant Information Sheet

Title: Chinese EFL Learners' Willingness to Communicate

You are invited to take part in a research study of students' willingness to communicate in English. Please read this form carefully and ask any questions you may have before agreeing to be in the study.

The study is conducted by Jiayi Zhang as part of her PG studies at Durham University.

* This research project is supervised by Dr. Nadin Beckmann (nadin.beckmann@durham.ac.uk) and Prof. Jens F. Beckmann (j.beckmann@durham.ac.uk) from the School of Education at Durham University.

The purpose of this study is (a) raise the awareness of fluctuations in learners' willingness to communicate in class, as well as (b) enable teachers to create supportive learning environments.

There are two parts to this study. If you agree to be in this study, in part 1, you will be asked to fill in a brief questionnaire. It will take you no more than 10 minutes to complete this questionnaire. In part 2, you will have your classes as usual, and a few sessions will be recorded. After class, you will be invited to go through the recording, reflect on your experience during the session and answer a small set of questions. It will take you about an hour to complete this part of the study. The interview will be recorded.

You are free to decide whether or not to participate. If you decide to participate, you are free to withdraw at any time without any negative consequences for you.

All responses you give or other data collected will be kept confidential. The records of this study will be kept secure and private. All files containing any information you give will be password protected. In any research report that may be published, no information will be included that will make it possible to identify you individually. There will be no way to connect your name to your responses at any time during or after the study.

If you have any questions, requests or concerns regarding this research, please contact me via email at Jiayi Zhang, jiayi.zhang@durham.ac.uk or by telephone at +44 (0)7541022080.

This study has been reviewed and approved by the School of Education Ethics Sub-Committee at Durham University (date of approval: December, 2016)

Jiayi Zhang

Leazes Road
Durham City, DH1 1TA
Telephone +44 (0)191 334 2000 Fax +44 (0)191 334 8311
www.durham.ac.uk
Durham University is the trading name of the University of Durham

Declaration of Informed Consent

- I agree to participate in this study, the purpose of which is to (a) raise the awareness of fluctuations in learners' willingness to communicate in class, as well as (b) enable teachers to create supportive learning environments.
- I have read the participant information sheet and understand the information provided.
- I have been informed that I may decline to answer any questions or withdraw from the study without penalty of any kind.
- I have been informed that data collection will involve the use of recording devices.
- I have been informed that all of my responses will be kept confidential and secure, and that I will not be identified in any report or other publication resulting from this research.
- I have been informed that the investigator will answer any questions regarding the study and its procedures. Jiayi Zhang, School of Education, Durham University can be contacted via email: jiayi.zhang@durham.ac.uk or telephone: +44 (0)7541022080.
- I will be provided with a copy of this form for my records.

Any concerns about this study should be addressed to the School of Education Ethics Sub-Committee, Durham University via email to ed.ethics@durham.ac.uk.

Date	Participant Name (please print)	Participant Signature
------	---------------------------------	-----------------------

I certify that I have presented the above information to the participant and secured his or her consent.

Date	Signature of Investigator
------	---------------------------

Appendix F Baseline Questionnaire for Teachers in Study 2

In the following section, we are interested in how you would describe each student in this class. Please mark the box that accurately describes his or her **communication behaviour** in general, and assess each student's language proficiency.

Communication behaviour refers to students' participation in activities when they are given the chance to talk in English with either the teacher or their peers (e.g. participate in group discussions, volunteer an answer when the teacher asks a question, ask the teacher a question, etc.).

Not at all	A little	Somewhat	Moderately	Quite	Very	Extremely
------------	----------	----------	------------	-------	------	-----------

Name	Communication Behaviour							English Proficiency (0–100)	Other Comments
	Not at all						Extremely		
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		

Thank you for your participation!

Appendix G Momentary Self-Report Questionnaire for Study 2

Code:

In this part, we would like to know how you think and feel about yourself and your experience IN THIS SPECIFIC SESSION. When answering the questions, please think about the activities when you were given the chance to talk IN ENGLISH with your teacher or peers. Please mark the box that best describes your response in relation to THESE SPECIFIC ACTIVITIES.

Not at all	A little	Somewhat	Moderately	Quite	Very	Extremely
------------	----------	----------	------------	-------	------	-----------

Activity 1

	Not at all						Extremely
I was willing to communicate in English in this activity.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I did communicate in English in this activity.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I performed well during this activity.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Activity 2

	Not at all						Extremely
I was willing to communicate in English in this activity.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I did communicate in English in this activity.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I performed well during this activity.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Activity 3

	Not at all						Extremely
I was willing to communicate in English in this activity.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I did communicate in English in this activity.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I performed well during this activity.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Activity 4

	Not at all						Extremely
I was willing to communicate in English in this activity.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I did communicate in English in this activity.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I performed well during this activity.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

If you have any other comments, please feel free to write them down in the following box.

Thank you for your participation!

Appendix H Momentary Questionnaire for Teachers in Study 2

In this part, we would like to know how you think about _____'s communication behaviour IN THIS SPECIFIC SESSION. Please mark the box that best describes your response in relation to EACH SPECIFIC ACTIVITY.

Not at all	A little	Somewhat	Moderately	Quite	Very	Extremely
------------	----------	----------	------------	-------	------	-----------

Activity 1

	Not at all						Extremely
He/She did communicate in English in this activity.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
He/She performed well during this activity.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Activity 2

	Not at all						Extremely
He/She did communicate in English in this activity.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
He/She performed well during this activity.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Activity 3

	Not at all						Extremely
He/She did communicate in English in this activity.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
He/She performed well during this activity.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Activity 4

	Not at all						Extremely
He/She did communicate in English in this activity.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
He/She performed well during this activity.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

If you have any other comments, please feel free to write them down in the following box.

Thank you for your participation!

Appendix I Interview Schedule

General questions

- How did you feel/think about today's class? What did you feel happy/unhappy with?
- Did you feel like talking in today's class? When did you feel most/least willing to communicate? Why/Why not?
- Did you work well with your classmates/group members/partner during this activity? Did you feel supported by the teacher?
- Did you think this activity was useful?
- Did you enjoy this activity?
- Did you think you were supposed to talk? Were there enough opportunities for you to talk during this activity?
- Did you feel competitive, anxious or embarrassed during this activity?
- Was there any point in time when you felt like to communicate but you did not talk? Why?
- Comparing the activities you experienced today, which one did you engage most? Why?
- How will you assess your language performance in today's class? Why?

Probing questions

- I was wondering if I could ask you something. I am just curious. Can you remember what you were feeling/thinking when XYZ said that/those words?
- I saw you were nodding/head shaking/looking down/looking around/chatting with your neighbour/laughing, can you please tell me what you were feeling/thinking at that moment?
- I noticed that you mentioned... quite a lot. Is that what you were most concerned about when you were speaking? Can you say a bit more about this?

Appendix J Observation Schedule

Date: _____ Class: _____ Activity: _____

Communication Behaviour		Individuals				
		Student A	Student B	Student C	Student D	Student E
Nominated Speaking	Presenting in front of the class					
	Answering a question					
Voluntary Speaking	Presenting in front of the class					
	Volunteering an answer					
	Presenting an opinion					
	Responding to an opinion					
	Asking a question					
	Asking for clarification					
	Response in chorus					
	Non-public response					
Non-Verbal Reaction	Chatting					
	Listening/ reading					
	Taking notes/ writing					
	Hand raising					
	Nodding/ head shaking					
	Looking around					
	Looking down					
	Laughing					
Off-task activity with neighbours						

Other Notes: