

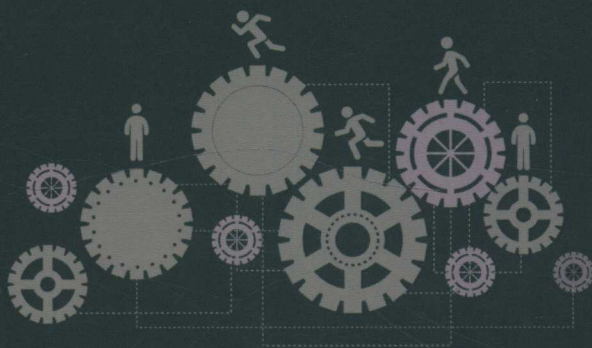


外语·文化·教学论丛

An Investigation into the Relationship Between
Test-takers' Personalities and Their Interactional
Competence: A Mixed Approach

受试者性格与其英语互动交际 能力的关系：基于混合法的研究

楼荷英 著



ZHEJIANG UNIVERSITY PRESS

浙江大学出版社



外语·文化·教学论丛

受试者性格与其英语互动交际能力的关系：基于混合法的研究

楼荷英 著

贵州师范学院内部使用



ZHEJIANG UNIVERSITY PRESS

浙江大学出版社

图书在版编目 (CIP) 数据

受试者性格与其英语互动交际能力的关系: 基于混
合法的研究 / 楼荷英著. — 杭州: 浙江大学出版社,
2019.3

ISBN 978-7-308-18607-0

I. ①受… II. ①楼… III. ①性格—关系—英语—口
语—研究 IV. ①H319.9

中国版本图书馆CIP数据核字 (2018) 第209188号

受试者性格与其英语互动交际能力的关系: 基于混合法的研究 楼荷英 著

策 划 张 琛 董 唯

责任编辑 田 慧

责任校对 胡建素

封面设计 周 灵

出版发行 浙江大学出版社

(杭州市天目山路148号 邮政编码310007)

(网址: <http://www.zjupress.com>)

排 版 杭州兴邦电子印务有限公司

印 刷 虎彩印艺股份有限公司

开 本 710mm×1000mm 1/16

印 张 13.75

字 数 311千

版 次 2019年3月第1版 2019年3月第1次印刷

书 号 ISBN 978-7-308-18607-0

定 价 45.00元

版权所有 翻印必究 印装差错 负责调换

浙江大学出版社市场运营中心联系方式: (0571) 88925591; <http://zjdxcs.tmall.com>

序 言

口语测试是直接反映考生语言交际能力的一种测试,形式多样。近年来,配对(小组)口语测试因省时实用、任务逼真、反拨积极等优势,而越来越多地被用于课堂测试和大规模高风险测试中。因其使用日益广泛,配对(小组)口语测试的效度开始受到语言测试界和测试分数使用者的广泛关注。语言测试研究者逐步开展了针对此类测试效度验证的实证性研究,考生特质对其口试表现的影响也成了研究焦点之一。《受试者性格与其英语互动交际能力的关系:基于混合法的研究》正是针对考生性格与其在小组口语测试中互动能力的发挥之间的关系而展开的多维度研究。

本研究采用自然数据,基于 Bachman 和 Palmer(2010)的评估使用论据(Assessment Use Argument, 简称 AUA)框架,采用定量和定性相结合的方法,探讨大规模口语测试环境下考生性格对其互动能力的发挥的影响。

本研究主要聚焦 AUA 中的一个重要主张(claim),即考试分数富有意义,并具有概括性,且考试对于考生互动能力的评估是有效的。主要的理据(warrant)是考试设计有助于考生的互动能力的发挥,因此考试结果可用来推断考生的互动能力,实现了考试大纲所定义的构想。主要的反驳(rebuttal)有三个:一是考生性格影响其互动能力的发挥,二是小组成员的性格影响个体考生互动能力的发挥,三是在小组讨论中的互动与考试设计有差异。

定量和定性研究结果表明:考生性格未对考生互动能力的发挥产生显著影响,说明考生性格不构成反驳。但定量和定性分析均发现考生在小组讨论中互动能力特征表现不充分,小组活动中考生之间互动不够,说明反驳三成立。

本研究在研究视角、理论框架及研究方法方面具有一定的创新性。尝试使用 Bachman 和 Palmer 提出的 AUA 框架开展效度验证,对传统的效度验证框架是一个很好的补充。同时,本研究的发现对我国大学英语教学、口语测试开发及评分量表的设计具有一定的借鉴意义。

何莲珍
2019年1月

Preface

This research is an attempt to investigate the effect of test-takers' personalities on their display of interactional competence (IC) in paired (group) oral tests. Specifically, the study examines: (1) The degree to which extroversion/introversion as measured by Gong Yaoxian's version of Eysenck Personality Questionnaire (also known as Eysenck Personality Questionnaire-Revised for Chinese, EPQ-RSC) affects an individual test-taker's display of IC; and (2) the degree to which the levels of extroversion of the group test-takers are assigned to affect their display of IC.

The study included 255 CET-SET (College English Test-Spoken English Test) test-takers. Their personalities were determined by the *T* value obtained from EPQ-RSC. Their group discussions in CET-SET were video-recorded and transcribed. Their IC sub-scores were compared. The IC features displayed by each test-taker during the discussion were coded and analyzed both quantitatively and qualitatively.

The results of descriptive statistics and MANOVA (Multivariate Analysis of Variance) showed no significant difference between the IC scores of the individual extroverts and introverts. Additional descriptive statistics indicated the degree of extroversion did not affect the display of the spoken discourse features. Frequency count showed slight differences in the display of socio-cultural competence and strategic competence between the extroverts and introverts. Results of conversation analysis confirmed that active participation in the discussion depended on test-takers' oral English proficiency rather than their personality.

Descriptive statistics and MANOVA analysis also found group members' personalities didn't significantly affect individual test-taker's IC scores. Conversation analysis further revealed that the individual test-taker's performance largely depended on his/her own oral English proficiency. However, frequency counting found no consistent results in the display of socio-cultural



competence and strategic competence among different personality groups.

On the whole, this study found test-takers' personalities were not the rebuttal to the warrants of CET-SET. However, the shared salient interaction features of all the discussion groups should arouse the test developers' attention.

The whole research bears substantial theoretical significance for language testing research, and carries practical implications for the administration of oral tests and the development of interactional competence in English language teaching.

The completion of the research and this book would have never been possible but for the support of my teachers, friends and my family. I would like to express my sincere appreciation to many individuals who have stayed with me along the long and winding road and made the completion of this book.

First of all, I'm deeply grateful to Prof. He Lianzhen for her excellent counsel and constant support. I couldn't have completed the project without her insights, guidance and encouragement.

My sincere thanks also go to Prof. Shao Yongzhen and Prof. Ying Huilan who helped me collect the data and encouraged me throughout the project.

I would also like to express my gratitude to Prof. Lyle Bachman, Prof. Pang Jixian and Prof. Wu Zongjie who gave me so much advice and inspiration during my study in the School of International Studies, Zhejiang University.

I deeply appreciate the financial support from Zhejiang Association of Philosophy and Social Sciences and Graduate School of Zhejiang University of Technology.

Finally, I would dedicate this book to my husband and my daughter for their enduring love and support.

Contents

Chapter 1 Introduction

1.1	Context of the Study	003
1.2	Purpose of the Study	006
1.3	Significance of the Study	006
1.4	Research Questions	007
1.5	Definitions of the Key Terms	008
1.6	Organization of the Book	009

Chapter 2 Literature Review

2.1	The Paired (Group) Oral Test	013
2.2	Empirical Researches on the Paired (Group) Oral Test	014
2.2.1	Researches on the Reliability of the Paired (Group) Oral Test	015
2.2.2	Researches on the Validity of the Paired (Group) Oral Test	016
2.2.2.1	Interlocutor effect of language proficiency	017
2.2.2.2	Interlocutor effect of gender	019
2.2.2.3	Interlocutor effect of acquaintanceship	019
2.2.2.4	Interlocutor effect of personality	020
2.3	CA in Language Testing Researches	026
2.3.1	CA	026



2.3.2	Quantification of the Talk	027
2.3.3	CA in Language Testing Researches	029
2.4	Bachman & Palmer's AUA.....	031
2.5	Summary	035

Chapter 3 Methodology

3.1	Data	039
3.1.1	CET-SET	039
3.1.1.1	Task features	040
3.1.1.2	Assessment	040
3.2	Instruments	042
3.2.1	EPQ	042
3.2.2	Gong Yaoxian's Version of EPQ	044
3.2.3	MANOVA	045
3.2.4	CA	045
3.3	Procedures	046
3.3.1	Administration of Gong Yaoxian's Version of EPQ	046
3.3.2	Scoring of the EPQ.....	047
3.3.3	Determining the Extroversion/Introversion	047
3.3.4	Determining the Oral Proficiency Levels	049
3.3.5	Data Transcription	049
3.3.6	Coding	050
3.3.6.1	Development of the coding scheme	050
3.3.6.2	Procedures for coding	059
3.3.7	Summary	064

**Chapter 4 The Relationship Between Test-Takers’ Personalities and
 Their Interactional Competence**

4.1 The Relationship Between an Individual Test-Taker’s Personality
 and IC067

 4.1.1 The Effect of an Individual Test-Taker’s Personality
 on His/Her IC Sub-scores068

 4.1.2 The Effect of an Individual Test-Taker’s Personality
 on His/Her Own Spoken Discourse070

 4.1.3 The Effect of an Individual Test-Taker’s Personality
 on His/Her Display of Other IC Features071

4.2 The Relationship Between a Test-Taker’s Discourse, Sub-Scores
 and His/Her Group Members’ Personalities072

 4.2.1 Group Classification073

 4.2.2 Quantitative Analysis073

 4.2.2.1 MANOVA analysis073

 4.2.2.2 Between-subjects test of a test-taker’s IC sub-scores
 and his/her group members’ personalities074

 4.2.3 Results of CA081

 4.2.3.1 The performance of the introverted when grouped
 with the extroverted082

 4.2.3.2 The performance of the groups with all extroverts095

4.3 The Display of the IC of All the Test-Takers Examined111

 4.3.1 Quantitative Analysis111

 4.3.2 Shared Salient Distinguishing Features in the Discussions of
 All the Groups113

4.4 Summary125



Chapter 5 Discussion and Conclusion

5.1	Summary and Discussion of the Findings	129
5.1.1	Summary of the Findings	129
5.1.2	Discussion of the Findings	132
5.1.2.1	Framing	132
5.1.2.2	Performativity theory	134
5.1.2.3	Rating scale	135
5.1.2.4	The eligibility for sitting CET-SET	136
5.1.2.5	Articulating the meaningfulness of the IC scores	138
5.2	Implications	141
5.2.1	The Teaching of IC to EFL Learners	141
5.2.2	The Revision of the CET-SET Rating Scale	144
5.3	Limitations of the Study	145
5.4	Suggestions for Future Research	146
5.5	Conclusion	146
	References	148
	Appendixes	163
	Appendix 1 CET-SET Sample Test (Examiner's Material)	163
	Appendix 2 CET-SET Rating Scale	166
	Appendix 3 Score Conversion and Band Grade Description	168
	Appendix 4 The Four Group Discussion Topics of CET-SET	169
	Appendix 5 Transcription Conventions	170
	Appendix 6 Transcription Excerpts	171
	Appendix 7 EPQ-RSC Questionnaire	200
	Index	204

Table of Figures & Diagrams

Figure 4.1	Box plot for Sub-score 1 average of the three personality groups	075
Figure 4.2	Box plot for Sub-score 2 average of the three personality groups	076
Figure 4.3	Group mean frequency distribution of the main socio-cultural competence features.....	080
Figure 4.4	Group mean frequency distribution of the strategic competence features	081
Figure 5.1	Articulation of the argument for the interpretations of IC to be assessed	139
Diagram 3.1	The two supertraits and the four quadrants divided (Eysenck & Eysenck, 1958: 15)	048



Table of Tables

Table 3.1	Coding scheme for IC	057
Table 4.1	Descriptive statistics for the IC sub-scores of the introverts	068
Table 4.2	Descriptive statistics for the IC sub-scores of the extroverts	069
Table 4.3	Multivariate tests of personality and the IC sub-scores	069
Table 4.4	Tests of between-subjects effects of personality and the IC sub-scores	069
Table 4.5	Descriptive statistics for the discourse features produced by the introverts	070
Table 4.6	Descriptive statistics for the discourse features produced by the extroverts	071
Table 4.7	Mean frequency of the use of other IC features	072
Table 4.8	Descriptive statistics for personality groups	074
Table 4.9	Multivariate tests for the group effect	077
Table 4.10	Tests of between-subjects effects of test-takers' IC and their group members' personalities	077
Table 4.11	Multivariate tests of the grouping effect on the IC sub-scores	078
Table 4.12	Tests of between-subjects effects of personality	078
Table 4.13	Multivariate tests of personality and spoken discourse	079
Table 4.14	Tests of between-subjects effects of personality and the spoken discourse features	079
Table 4.15	Frequency distribution of the main socio-cultural competence features	111
Table 4.16	Frequency distribution of the strategic competence features (1)	112
Table 4.17	Frequency distribution of the strategic competence features (2)	112

Abbreviations

AUA	Assessment Use Argument
CA	conversation analysis
CET	College English Test
CET-SET	College English Test-Spoken English Test
EPQ	Eysenck Personality Questionnaire
EPQ-RSC	Eysenck Personality Questionnaire-Revised for Chinese
IC	interactional competence
ICS	Initial Coding Scheme
MANOVA	Multivariate Analysis of Variance
OPI	Oral Proficiency Interview
PENL	Psychoticism, Extroversion, Neuroticism, and Lie



Chapter 1

Introduction

贵州师范学院内部使用

1.1

Context of the Study

The theoretical advances in applied linguistics in the last few decades regarding “communicative competence” (Canale, 1983; Canale & Swain, 1980; Celce-Murcia, Dornyei & Thurrel, 1995; Hymes, 1972) and “communicative language ability” (Bachman, 1990; Bachman & Palmer, 1996) have exerted long-lasting impact on both language teaching and testing community. While the notion of communicative language teaching changes the classroom teaching models, the concept of communicative competence also becomes the impetus for the evolution of theory and practice in language testing field. With the construct of language ability being reexamined, communicative language testing has been advocated since the second half of the 20th century. In the assessment of speaking abilities of language learners, formats have continuously been changing. Recently, paired (group) tests have been adopted more and more in both high-stakes tests and low-stakes tests. And increased attention has turned to the investigation of interaction and its role in the speaking construct (Plough, Banerjee & Iwashita, 2018). Thus, the construct of interactional competence (IC) and its operationalization in the speaking tests have become the language testing professionals’ research interest (Galaczi, 2008, 2014; Nakatsuhara, 2011; Plough, Banerjee & Iwashita, 2018).

Since they elicit a wide range of interaction skills and are more authentic, paired (group) oral tests have become one of the most widely-used test formats for assessing the speaking abilities of language learners, and the discourse yielded from which is more natural than Oral Proficiency Interviews (OPIs). With more authentic test tasks, the score interpretations of the tests are more valid (Ockey, 2009). Theoretically speaking, compared with OPIs, more valid inferences can be made about the test-takers’ oral proficiency from the scores obtained in the paired (group) oral discussions. However, problems of paired (group) oral test format surface with the increasing popularity of its use. A small but growing number of studies have indicated some concerns about the threat to



test-takers' score-based inferences (Davis, 2009; Folland & Robertson, 1976; Foot, 1999; He & Dai, 2006; Iwashita, 1998; Nevo & Shohamy, 1984; Norton, 2005; Ockey, 2009; O'Sullivan, 2002; Shohamy, Reves & Bejarano, 1986; Plough et al, Banerjee & Iwashita 2018). As many scholars have pointed out, challenges for testing L2 oral proficiency include matters such as topic familiarity, interlocutor proficiency, test-taker relations, and task understandings (Brooks, 2009; Davis, 2009; Galaczi, 2008; Iwashita, 2001; Lazaraton & Davis, 2008; May, 2011). Oral tests present additional challenges to standardization, as social interaction is a joint achievement and external conditions are difficult to control. The joint construction of the performance in such kind of tests is still to be accounted for in both construct definitions, rating scales and other facets (Brooks, 2009; Plough, Banerjee & Iwashita, 2018).

While the potential threats to score-based inference have received some attention, empirical studies to collect validity evidences are scanty (Davis, 2009; Galaczi, 2008, 2014; Fulcher, 1996; He & Dai, 2006; Iwashita, 1998; Lazaraton, 2008; Nevo & Shohamy, 1984; Norton, 2005; Ockey, 2009; O'Sullivan, 2002; Shohamy, Reves & Bejarano, 1986; Nakatsuhara, 2011; Lam et al., 2018). It has been argued that research on L2 oral proficiency is limited (Moeller & Theiler, 2014) and research on L2 oral tests is even more limited (Ducass & Brown, 2009; Plough, Banerjee & Iwashta, 2018). To understand the whole interaction process of group oral test and obtain a deeper understanding of what a test actually measures, multiple perspectives with a postpositivist outlook to examine the validity are essential (multiple O'Loughlin, 2002). As McNamara suggests, "What language-testing researchers need to do now is to identify every possible effect of different facets of the test task on the inference made from the scores" (McNamara, 2001). To achieve the aforementioned goal, both well-designed qualitative inquiries and quantitative researches with advanced technologies are indispensable.

For the researches of spoken tests, especially those on the accountability of the inferential scores, qualitative inquiries, or discourse-based studies to be exact, can provide rich evidences. Young & Milanovic (1992) convincingly argue that researchers need to go into much greater detail in describing the nature of contingency as it exists between each set of turns to understand how discourse is created (420-421). Since in a social space like group oral tests, members engage