

李清华 著

*The purpose of the present study is to investigate whether and how PBWA, a low-stakes small scale writing assessment, impacts on Chinese EFL learners.*

*The conceptual framework was designed based on Hughes' (1993) trichotomy and Bailey's (1996, 1999) reorganization within the framework, learners'*

*affective factor (motivation and attitude),*

# Consequence of EFL Writing Formative Assessment

# 外语写作形成性 评估的 后效研究



李清华 著

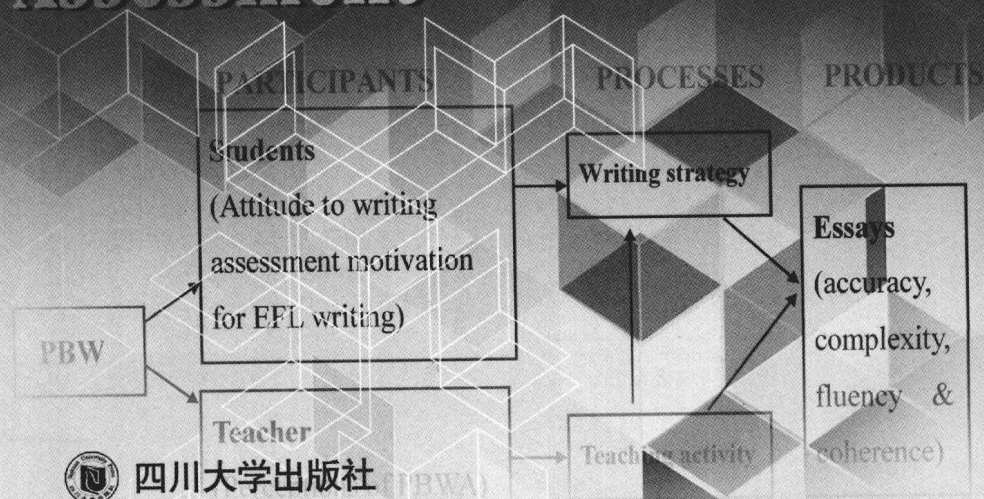
*The purpose of the present study is to investigate whether and how PBWA, a low-stakes small scale writing assessment, impacts on Chinese EFL learners.*

*The conceptual framework was designed based on Hughes' (1993) trichotomy and Bailey's (1996, 1999) reorganization within the framework, learners'*

*ffective factor (motivation and attitude),*

# Consequence of EFL Writing Formative Assessment

# 外语写作形成性 评估的 后效研究



四川大学出版社

责任编辑:黄新路  
责任校对:敬铃凌  
封面设计:米茄设计工作室  
责任印制:李 平

### 图书在版编目(CIP)数据

外语写作形成性评估的后效研究 / 李清华著. —成都:  
四川大学出版社, 2008.2

ISBN 978-7-5614-3960-9

I. 外… II. 李… III. 英语-写作-研究 IV. H315

中国版本图书馆 CIP 数据核字 (2008) 第 015559 号

### 书名 外语写作形成性评估的后效研究

---

著 者 李清华  
出 版 四川大学出版社  
地 址 成都市一环路南一段 24 号 (610065)  
发 行 四川大学出版社  
书 号 ISBN 978-7-5614-3960-9/H·256  
印 刷 郫县犀浦印刷厂  
成品尺寸 140 mm×202 mm  
印 张 9.75  
字 数 248 千字  
版 次 2008 年 3 月第 1 版  
印 次 2008 年 3 月第 1 次印刷  
印 数 0 001~1 500 册  
定 价 28.00 元

◆读者邮购本书,请与本社发行科  
联系。电 话:85408408/85401670/  
85408023 邮政编码:610065  
◆本社图书如有印装质量问题,请  
寄回出版社调换。

◆网址:www.scupress.com.cn

---

版权所有◆侵权必究

## 序

清华的博士论文正式出版了，作为导师，我为之感到高兴。清华在广外攻读博士学位期间，表现出较深厚的学术素养和较强的研究能力，在国内权威学术期刊上发表过多篇关于语言测试的学术论文。

清华的博士论文选题时，我正在美国加利福尼亚大学洛杉矶校区访学，了解到美国的教育评估有很多成功的经验值得我们借鉴学习，其中之一就是形成性评估。由此，我建议清华关注这一领域的研究。他当时已有了较成熟的开题思路，但还是阅读了大量有关形成性评估的文献，决定把博士论文的研究方向转为外语的形成性评估。

教育评估大致可分为三种：(1)对学习的评估；(2)为学习的评估；(3)作为学习的评估。对学习的评估，即，在单元、课程或学期结束时进行的终结性评估，目的是根据评估结果（考试分数或等级）划分学生的“等级”，目前流行的考试大都属于这类评估；为学习的评估则把重心从终结性转移到形成性，目的是跟踪学生学习发展的过程，对学生的下阶段学习提供必要的诊断性信息，这种诊断性信息是基于学生的个人“产品”(artifacts)，特别是学习过程中产出的 portfolio；在作为学习的评估中，学生既是评估的对象又是评估的主人，使评估真正融入教学与学习之中。

随着语言学习观的发展与变化,语言学习与语言测试都开始把重心从重结果向重过程转移。形成性评估越来越受到语言测试研究者及语言教师的重视。比如,在写作评估中,档案式评估(Portfolio-based Writing Assessment, PBWA)受到了普遍的关注。PBWA 比传统写作测试具有很多优势:(1)提高教师课堂教学和评估的质量:Portfolio 包括学生作文的初稿、修改稿及对写作过程的反思,强调的是写作的发展过程,评估不再只是对学生评分数、分等级、排顺序的“外来的”活动,而是课堂活动的一部分。(2)促进学生的学习:这种评估方法能够 ①让学生选择自己的代表作品;②让学生对自己的作品自我反思;③让学生参与评估标准的制定;④要求学生进行自我评估和同伴评估;⑤让学生本人告诉教师他/她学会了什么、正在学什么,而不是教师告诉学生他/她认为学生学会了什么;⑥让学生为教师提供有关教学实践的反馈信息,这样,学生更多地参与评估过程,既是评估的对象,又是评估的执行者。(3)对学生的学习提供有效的测量:PBWA 的结果能够更有效地推断学生的写作能力,因为 a. 在课堂上进行的 PBWA 能够创造一种写作氛围而不是测试氛围,从而在真实的学习过程中达到评估目的;b. PBWA 是对学生在一段时间内所写的丰富多样的作文样本的评价,能够更全面地反映学生的写作过程;c. PBWA 本身是学习活动的一部分,学生在评估中学习。

PBWA 尽管有这么多潜在的优点,但实施 PBWA 仍面临诸

多挑战。比如, (1) 设计: 普通教师能否设计出合理可行的方案? 是否需要专家指导? 如果教师主持设计, 如何保证不同教师之间、不同学校之间的可比性? (2) 工作量: 实施 PBWA 会增加教师的工作量, 特别是 portfolio 的评分可能需要较长时间。(3) 成绩报告: PBWA 除了对学生作文(成品)评价, 还对写作过程评价(如, 修改的质量, 反思的质量等)。对习惯于根据一个分数来评价作文水平的学生、家长、管理者等有关人员, 如何让他们理解 PBWA 的评价结果? (4) 信度: 如何制定多体裁多篇作文的评分标准? 如何保证评分员之间、不同写作任务之间的一致性? (5) 效度: 如何控制剽窃现象? 如何保证学生享有同样的资源? 如何证明 PBWA 比传统的一次性限时写作测试有更高的构念效度?

清华的研究切入点是探索这种评估方法在中国外语教育体系中的可行性和有效性, 核心问题操作化为: PBWA 是否以及如何影响中国大学英语专业学生写作能力的发展? 在语言测试效度理论的基础上, 他根据研究语言测试反拨作用(washback)的三成分(trichotomy)模式, 提出了研究的理论框架, 即, 从三方面研究 PBWA 对学生写作能力发展的影响: ①学生(写作动机及对写作评估的态度)与教师(对 PBWA 的认识); ②学生的写作过程(写作策略的运用)与教师的教学过程(课堂活动); ③学生的写作结果(作文在准确度、复杂度、流利度和连贯性等方面的质量)。现有语言测试反拨作用的研究没有明确理论定

位, 主要关注对教师及其课堂教学的影响。清华的研究框架把语言测试反拨作用归入效度验证的证据之一——后果 (consequence) 证据, 并拓展了语言测试反拨作用的研究范围, 强调学生的学习过程和学习结果, 具有一定的创新性。为了探究 PBWA 对学生写作能力发展的影响, 清华采用对照实验结合嵌入式案例设计, 收集了从写作能力的客观测量、问卷调查和课堂观察所得到的量的数据及案例研究中从访谈记录、反思论文、课堂日记等途径获得的质的数据。经过缜密的分析, 得出了令人信服结论。

目前国内系统研究外语形成性评估的学术专著并不多。我相信, 该书的出版对我国形成性评估的研究及应用将起到积极的推动作用。

是为序。

**曾用强**

2007 年 10 月 28 日

广州白云山下

## Preface

The history of writing assessment falls into four periods: (1) direct testing (i.e., essay tests), represented by ancient China; (2) multiple-choice testing, represented by America; (3) portfolio-based assessment, represented by America and Britain; and (4) the emerging generation that will need to be technological, humanistic, political, and ethical (Hamp-Lyons, 2001, 2002). Portfolio is not a new invention. For decades, it is used as a standard form of assessment in fields related to the visual arts such as architecture, design, and photography (Brown & Hudson, 1998; Weigle, 2002). In English L1 writing assessment, portfolios have quite a long history. A body of evidence has indicated that portfolio-based writing assessment (PBWA) is now widely acknowledged as a promising alternative to the conventional timed impromptu essay test in English L1 context (Weigle, 2002; Hamp-Lyons & Condon, 2000). It is argued that the use of portfolios is more beneficial to ESL/ EFL students (Delett, Barnhardt, & Kevorkian, 2001; Hamp-Lyons & Condon, 2000; Song & August, 2002; among others). While there are numerous claims about the potential benefits of PBWA to ESL/EFL learners, a comprehensive review of literature on writing assessment has led to the conclusion that little evidence is available to confirm whether and how far these claims appear to be upheld. Furthermore, an overwhelming majority of the available literature on PBWA comes from English L1 context and not adequate attention has been drawn to the applications of PBWA in ESL/EFL context. Thus, the most pressing need in this still young field is to conduct more research to



### **Consequence of EFL Writing Formative Assessment**

validate the claims made by the advocates of PBWA, particularly, in ESL/EFL context. Given that far too little research on the practice and consequences of portfolio assessment with ESL/EFL learners, more empirical studies on the impact of PBWA on ESL/EFL learners' writing are called for.

Compared with portfolio use in North America, Europe, Taiwan district of China and other countries and districts, the implementation of portfolios in EFL writing assessment in mainland China is relatively rare. Evidently, the feasibility and effectiveness of PBWA for Chinese EFL learners should be investigated before this technology can be used to foster EFL writing teaching, learning and assessment in China's mainland. The present study focused on whether and how PBWA promoted EFL writing development of Chinese university English majors. The purpose was to explore the feasibility and effectiveness of PBWA in the educational system of mainland China. The key research question was: How does PBWA impact on Chinese undergraduates' EFL writing development? Based on the trichotomy model applied in washback research, the researcher produced the conceptual framework for the current study. That is, effects of PBWA should be investigated in terms of learners' motivation and attitude, their writing process (strategy use) and writing outputs (quality of essays). Consequently, the following six sub-questions were addressed in order to answer the overall question:

(1) Are there significant differences between the PBWA experimental class and the non-PBWA class in terms of their motivation to learn EFL writing at the end of the experimental semester?

(2) Are there significant differences between the PBWA

experimental class and the non-PBWA class in terms of their EFL writing strategies at the end of the experimental semester?

(3) Are there significant differences between the PBWA experimental class and the non-PBWA class in terms of writing products (accuracy, complexity, fluency and coherence) at the end of the experimental semester?

(4) Are there significant differences between the PBWA experimental class and the non-PBWA class in terms of writing class experiences?

(5) Are there significant differences between the PBWA experimental class and the non-PBWA class in terms of attitude towards writing assessment at the end of the semester?

(6) How do the students in the experimental class perceive PBWA during the PBWA project?

Pursuit of the effects of PBWA on writing development of Chinese EFL learners warranted a comparative study, in which two sophomore English major classes with similar writing motivation, strategy and proficiency, same number and sex distribution, taught by the same teacher were engaged. However, the results of the comparison could tell only part of the story because it provided little information about the detailed process of PBWA affecting EFL writing growth. A single embedded case study (the experimental class was the case, and the six selected students were the sub-units for analysis.), therefore, was employed to investigate the hidden transcripts. Multiple types of data were collected, compassing quantitative data from questionnaires and writing measurement, and qualitative data from transcribed interviews with the six selected students, journals, reflective essays in their portfolios as well as field

notes of classroom observation.

The major findings of the present study were as follows: First, EFL learners had the tendency to treat PBWA as a better means of assessment and an effective tool of learning, though they looked at it as a container of their essays upon entering the project. Compared with timed singular essay tests, PBWA was believed to be “more suitable and fairer” by the participants. Second, implementing PBWA was feasible in Chinese colleges, as it was welcomed by the English majors and their teacher. Third, PBWA aided the development of positive attitude towards EFL writing and its assessment, for students involving in the PBWA project were inspired to increase their writing motivation and to improve the writing strategy. Fourth, PBWA facilitated growth of EFL writing ability at least in some dimensions, specifically, accuracy and coherence.

This study tentatively concluded: On the whole, PBWA had profound positive impact on the growth of Chinese EFL learners’ writing ability.

The thesis project was a long journey during which I have got generous and cordial help from many people. First of all, I wish to express my sincere gratitude to my advisor, Professor Yongqiang Zeng. In my doctoral study, his guidance kept me on the right track and enabled me to go through difficulties; his rich knowledge and insightful perspectives often helped me to change confusion into clarity. Despite many other commitments, he was always accessible and ready to help me. His gracious mentorship, guidance and support not only helped me to complete my doctoral research but also taught me how to be a qualified teacher and researcher in my further academic career. My gratitude also goes to the external examiners of

### *Preface*

my dissertation, Prof. Yan Jin, Prof. Chuming Wang, Guangkeng He, Prof. Jianda Liu, and Dr. Xinping Chang, for their critical and constructive comments.

I would like to express my gratitude to professors in Guangdong University of Foreign Studies (GUFS), especially, Prof. Shichun Gui, Prof. Ziran He, Prof. Xudong Wu, Prof. Jianda Liu, Prof. Binli Wen, Prof. Chuming Wang, Prof. Lunxia Qi, whose classes empowered me with professional knowledge and research method in linguistics and applied linguistics. In addition, my special appreciation goes to Dr. Zhong'en Xi. He is my colleague and friend. His devotion to academic research, his extensive knowledge of statistics and psychometrics forced, encouraged and helped me to overcome various difficulties in conducting research. He attentively reviewed the first draft of the dissertation and offered prompt and substantive suggestions. I also owe huge debt of gratitude to Dr. Maocheng Liang in Beijing Foreign Studies University and Dr. Zhiqiang Cai in Memphis University in America. Both of them offered me great help in application of coherence measurement.

Grateful thanks should be rewarded to my friends and fellow doctoral students in GUFS, Hong Zhang, Kaihong Liao, Jianfu Liu, Zhaochun Sun, Jie Cheng, Yusong Gao, Dongwei Chu, Chunshou Lan, Chen Zhao, Zhi gang Ma, Liang Li, Xinling Zhang, Nianning Chen, Jie Zhang, Xiaohong Jiang, Min Wang, Shuang Liang, Rong Zhao, to mention a few. The time of discussion with them, playing table tennis with them, playing basketball with them, and climbing Baiyun Mountain with them was the happiest period in the long tough trudge. Besides, thanks also go to Miss Wen Kong, a doctor candidate in Shanghai International Studies University, for her help and

encouragement.

This dissertation study was conducted with a project of EFL writing assessment reform in Shaoxing University. I owe a huge debt of gratitude to the participant teacher and students for their friendly cooperation and great help though I could not mention their names.

Finally, I would like to thank my family members. My beloved wife, Hui Kong, not only shouldered all the family responsibility but also helped me sort out and enter the data into computer though she has suffered from illness for a long time. My lovely daughter, Di Li, encouraged and inspired me when I encountered difficulties. My brother, Dr. Zhonghua Li, always made immediate response to my requirement of reference material by mail or email from America, though he was extremely busy with his study and work as well as household duties. They were always there, sharing my sadness, frustrations, excitement, and happiness. Without their love and support, I could not have completed my doctoral study.

Qinghua Li

July, 2007

## 前言

语言写作测试与评估的历史可分为四个阶段：(1) 以古代中国为代表的直接测试（论文写作）；(2) 以美国为代表的多项选择式客观测试；(3) 以欧美等国和台湾地区为代表的档案式评估（portfolio-based assessment）；(4) 初露端倪的以高科技和人性化为主要特点的新型评估（Hamp-Lyons, 2001, 2002）。大量证据表明，档案式写作评估（portfolio-based writing assessment, PBWA）作为传统限时写作测试的替代方法在以英语为第一语言的环境下已得到广泛认可（Hamp-Lyons & Condon, 2000; Weigle, 2002）。据称，PBWA 更有利于以英语为第二语言或外语（ESL/EFL）的学生（如，Delett, Barnhardt, & Kevorkian, 2001; Hamp-Lyons & Condon, 2000; Song & August, 2002）。自 20 世纪 80 年代以来，PBWA 的应用越来越多，但相关的实证研究却很少，而且多数 PBWA 研究都是在以英语为第一语言的环境中进行的。因此，对 PBWA 这一新兴的研究领域而言，当务之急是开展更多的实证研究以证明这种评估方法的效度。当代效度理论认为，效度验证需要多方面的证据，后果效度（consequence validity）是一个重要的方面。但是，PBWA 对 ESL/EFL 学生写作能力发展的影响，尚未得到足够的重视。

与欧美等国家和与中国台湾地区相比，PBWA 在中国大陆 EFL 写作评估中的应用几乎是空白。显然，其可行性和有效性应该得到充分研究论证后才能用于中国大陆 EFL 写作教学与评估。本研究探索这种评估方法在中国教育体系中的可行性和有效性，其核心问题是：PBWA 是否是以及如何影响中国大学英语专业学生写作能力的发展？根据研究语言测试反拨作用（washback）的三成分（trichotomy）模式，作者提出了本研究的理论框架，即，PBWA

对学生写作能力发展的影响应从以下三方面开展研究：学生（写作动机及对写作评估的态度）与教师（对 PBWA 的认识），学生的写作过程（写作策略的运用）与教师的教学过程（课堂活动），以及学生的写作结果（作文在准确度、复杂度、流利度和连贯性等方面的质量）。所以，本研究的核心问题可以通过下列六个次属问题来回答：

（1）在实验学期结束时 PBWA 实验班与对照班的学生在英语写作动机方面有无显著差异？

（2）在实验学期结束时 PBWA 实验班与对照班的学生在英语写作策略使用方面有无显著差异？

（3）在实验学期结束时 PBWA 实验班与对照班的学生在英语写作结果（准确度、复杂度、流利度和连贯性）使用方面有无显著差异？

（4）在实验学期结束时 PBWA 实验班与对照班的学生在英语写作课经历方面有无显著差异？

（5）在实验学期结束时 PBWA 实验班与对照班的学生在对英语写作评估的态度方面有无显著差异？

（6）PBWA 实验班的学生是如何认识和接受 PBWA 的？

为了考察 PBWA 对中国大学英语专业学生写作能力发展的影响，本研究采用对照实验设计。在实验前，实验班和对照班在人数、性别比例、英语水平、写作动机、写作策略、写作能力及教学条件等方面尽可能保证一致。在实验过程中，除评估方法外，两组的处理都保持一致。但这种对照研究的结果却无法揭示 PBWA 影响学生写作能力发展的过程，而案例研究（case study）可以提供这方面的信息。因此，本研究同时使用了嵌入式案例设计，即，把实验班作为个案，从中选取六个学生作为分析单位。本研究收集的证据包括从写作能力的客观测量、问卷调查和课堂观察中所得到的量的数据（quantitative data），及从案例研究

的访谈记录、反思论文(reflective essay)、课堂日记(journal)等获得的质的数据(qualitative data)。

历时一个学期的实验的主要发现包括以下方面：第一，尽管 PBWA 实验班的学生在实验初把 PBWA 作为一种新的作业形式，有的学生甚至有抵触情绪，但在实验末他们则普遍视其为一种更好的评估方法和有效的促学工具；第二，这种评估方式受到教师和学生的欢迎，因而在中国大学实施 PBWA 是可行的；第三，PBWA 有助于促进学生对 EFL 写作学习与评估采取积极态度，有助于促进学生写作动机的提高和写作策略的改进；第四，PBWA 至少在准确度和连贯性等方面有助于学生写作能力的发展。总之，本研究的初步结论是：总体而言，中国大学英语专业学生对 PBWA 持欢迎态度，PBWA 对他们写作能力的发展产生了显著的积极影响，在我国的 EFL 写作评估中是可行的。

本书是在作者博士论文的基础上修改而成的。本研究得到了多方慷慨帮助。作者首先感谢导师曾用强教授。在论文选题、实验、撰写和修改过程中，他丰富的专业知识和严谨的治学态度不仅帮助我顺利完成了学业，而且让我受益终生。曾教授还热情为本书作序，对笔者的研究多有褒奖。论文答辩委员会的金艳教授（主席）、王初明教授、何广铿教授、刘建达教授和常新平博士对本论文提出了许多批评与建议。作者的同事和朋友席仲恩博士在百忙中审阅了论文的初稿，提出了许多建设性的修改意见。北京外国语大学的梁茂成博士和美国孟菲斯大学的蔡志强博士为作者提供了写作连贯性测量方面的指导。上海外国语大学博士生孔文对本论文的修改提出了很多建议。在广东外语外贸大学读书期间，作者既得以聆听桂诗春、何自然、王初明、吴旭东、温宾利、刘建达、亓鲁霞等老师的教诲，还得到很多学友的启发与鼓励，特别是张宏、廖开洪、刘建福、孙肇春、程杰、高育松、兰春寿、李亮、张新玲、张洁、陈念宁、江小红、王敏、梁爽等等。对这



些师长和学长的关心，笔者深表谢意。

笔者感谢所有参与本实验的老师和同学。他们的积极配合和热情帮助使本研究得以顺利完成。

本书的面世还得益于很多方面的资助和帮助。绍兴文理学院学术出版基金对本书的出版提供了资助。四川大学出版社的黄新路先生为本书的编辑花了大量心血。在此，作者表示诚挚的谢意。

在攻读博士学位的三年里，笔者的爱人孔慧女士独自挑起家庭的重担，并时刻牵挂研究进展；爱女李迪不仅在自己的学习中不断进步，还经常给予鼓励；笔者的弟弟李中华博士在美国为我提供所需的资料。他们与笔者共享挫折的困惑与成功的喜悦。没有他们的关爱与支持，笔者不可能完成这项研究。对所有关心我的亲人，我衷心地道一声：“谢谢！”

本书基于当代语言测试效度理论和最新 EFL 写作能力理论，是系统研究我国 EFL 写作形成性评估的第一部著作，为相关研究提供了理论基础和方法论的参照。但由于作者水平有限，虽经多次修改，拙著仍难免讹误。敬请广大读者提出宝贵意见。

李清华

2007 年仲夏