



Education Quarterly Reviews

Liu, Lisha (2018), Educational Experimental Research Design: Investigating the Effect of “PAD + Microlectures” EAP Teaching Model on Chinese Undergraduates’ Critical Thinking Development. In: *Education Quarterly Reviews*, Vol.1, No.2, 96-102.

ISSN 2621-5799

DOI: 10.31014/aior.1993.01.01.10

The online version of this article can be found at:
<https://www.asianinstituteofresearch.org/>

Published by:
The Asian Institute of Research

The *Education Quarterly Reviews* is an Open Access publication. It may be read, copied and distributed free of charge according to the conditions of the Creative Commons Attribution 4.0 International license.

The Asian Institute of Research *Education Quarterly Reviews* is a peer-reviewed International Journal. The journal covers scholarly articles in the fields of education, linguistics, literature, educational theory, research, and methodologies, curriculum, elementary and secondary education, higher education, foreign language education, teaching and learning, teacher education, education of special groups, and other fields of study related to education. As the journal is Open Access, it ensures high visibility and the increase of citations for all research articles published. The *Education Quarterly Reviews* aims to facilitate scholarly work on recent theoretical and practical aspects of Education.



ASIAN INSTITUTE OF RESEARCH
Connecting Scholars Worldwide

Educational Experimental Research Design: Investigating the Effect of “PAD + Microlectures” EAP Teaching Model on Chinese Undergraduates’ Critical Thinking Development

Lisha Liu¹

^{1,2} Lecturer, Center for Language Education, Southern University of Science and Technology, China, liuls@sustc.edu.cn

Abstract

This paper presents an educational, experimental research design, aiming at examining the effect of "PAD + microlectures" EAP teaching model on Chinese undergraduates' critical thinking development. It mainly analyzes this experiment from four aspects: research question and hypotheses, difficulties in key term definitions and selection of measuring instruments, potential risks of the design, as well as the challenge of ethics. Such a methodological analysis shows that educational experiments should follow the disciplines of objectivity, feasibility, maneuverability, effectiveness, and innovation.

Key Words: Critical thinking, PAD, Microlecture, English teaching, Academic English

1. Introduction

Critical thinking, a buzzword in both educational and academic fields (Fisher 2001, p.1), has been concerned by the Chinese higher education sector. Previous studies (Luo & Yang 2001; He 2005; Yao 2001) have indicated that western undergraduates usually achieve better results than Chinese ones in critical thinking tests. This finding reveals a dire threat to Chinese undergraduates' cognitive development. In order to reduce this threat, many Chinese universities have begun to embed critical thinking training in their subject courses, aimed at formulating effective learning strategies for their students. In contrast, Chinese undergraduates' English courses are still lack of critical thinking element (Sun, 2011). Under such a circumstance, with an increased interest in improving undergraduates' critical thinking in the English context, the curriculum of English for academic purpose (EAP) is developed. After all, high-level thoughts necessarily involve language. However, so far, little empirical research has been conducted on the teaching effects of EAP courses at Chinese universities.

In order to enhance undergraduates' engagement in English learning through computer-assisted language learning (CALL), an increasing number of Chinese college English teachers began to navigate new ICT models to obtain a better effect of their English teaching. Therefore, some newly established models like MOOCs and flipped class have been introduced and widespread over the last five years (Chen, Wang, and Jiang, 2015, p.67). Not only have the media announced them on a large scale, but also more educational research has been involved in them. Nevertheless, when these models are used as main teaching forms in practice, teachers and universities

are facing new difficulties in funding and infrastructure. Thus, the reform of college English teaching in China should consider the status quo of students, teachers, and universities.

In this case, a “PAD + microlectures” model was created, which aims at combining the PAD class with microlecture resources in order to improve Chinese non-English major undergraduates’ academic English. PAD class is a Chinese original teaching philosophy, created by Professor Zhang Xuexin at Fudan University. It separates the teaching process into three steps: presentation, assimilation, and discussion, which emphasizes the formative assessment and avoids cramming. Its essence is to divide the class time into halves, half for teachers’ teaching, the other half for students’ discussion. There is a week for students to assimilate the teaching contents between the aforementioned halves (Zhang, 2017). Although it is mainly student-centered, still retains the teachers’ guidance in class, which is an advantage of the traditional teaching method. The microlecture prototype was from the United States. In China, it is defined as video-recorded teaching activities concentrating on small and specific contents, aiming at helping students learn better (NUTNTC, 2014). It can be used as supplementary teaching and learning resource. To some extent, combining PAD and microlectures can not only develop students’ thinking effectively in class but also provide diverse materials for their language learning after class. Although some Chinese scholars (Zhang, 2016; Zhao, 2016; Wei, 2016) and college English teachers argue that this model is feasible because it has no class-size or workload restrictions, there still seems to be little empirical research on this model’s effect in the context of EAP teaching.

Given the lack of empirical investigations into the teaching effects of “PAD + microlectures” EAP courses on Chinese undergraduates’ critical thinking disposition, I attempt to evaluate such a topic experimentally. This paper seeks to identify the potential problems that might threaten the success of my experimental project and discuss how these problems might be minimised or averted. The findings of such a methodological analysis may provide some references for educational researchers.

2. The research question and hypotheses

Research questions and hypotheses are vital preconditions of research designs. Cresswell (2003) argues that the formulation of research purposes relies on those questions and hypotheses to be answered and tested. Only when these questions or assumptions are clearly outlined can researchers know how they should shape their research purposes. The main research question of my study is: ‘What is the impact of “PAD + microlectures” teaching model on undergraduates’ critical thinking ability?’ The purpose of my experiment is to measure the critical thinking dispositional change in undergraduates after being taught by “PAD + microlectures” in their academic English classes. Because ‘hypotheses are typically used in experiments in which investigators compare groups’ (Cresswell 2003, p.108), I specify the research question into the following pertinent hypotheses for testing.

H1: There is a significant difference in the critical thinking enhancement between the two treatment groups and two control groups.

H2: The critical thinking enhancement of the highest critical thinking students (20% of the sample) and the lowest ones are diverse (20% of the sample). The latter is more significant.

H3: The analytic and systematic thinking styles of the students in the treatment groups are significantly higher than those in the control groups.

3. Difficulties in the Definition of the Terms and the Selection of the Measuring Instruments

Providing definitions of the relevant terms ought to be considered as the first challenge in a piece of research because the readers need to understand their precise meanings. The most important term in my proposed study is ‘critical thinking.’ With respect to this term, John Dewey, the ‘father’ of the modern critical thinking tradition (Fisher 2001, p.2), defined ‘critical thinking’ as:

An active, persistent, and careful consideration of a belief or supposed form of knowledge in the light of the grounds which support it and the further conclusions to which it tends (Dewey 1909, p. 9).

Dewey thought that critical thinking should be regarded as ‘a form of purposeful judgment, specifically reflective judgment’ (Fisher 2001, p.3). We can also identify from this definition that critical thinking has huge

importance for reasoning, and 'skillful reasoning is a key element' (2001, p.3). It involves acquiring and assessing information to draw a well-justified conclusion. However, as a term definition, Dewey's explanation is abstract to the readers. The words 'active, persistent, and careful' are perceptual and difficult to measure. In order to make these adjectives clearer, Paul and Elder (2006) regard critical thinking as an art, arguing that it is a process of analyzing and evaluating thinking with a view to improving it. In other words, it requires a questioning approach to knowledge and perceived wisdom. To some extent, this definition is much more specific because the words are not abstract but familiar to readers.

Another difficulty is the selection of measuring instruments. Although a considerable number of measures are available for critical thinking, I need to consider the range, subject, and feasibility of each measure. The California Critical Thinking Disposition Inventory (CCTDI) is widely used for the assessment of university students' critical thinking capacities. There follows a description of it.

The inventory consists of 75 questions that represent 7 scales: truth-seeking, open-mindedness, analyticity, systematicity, self-confidence, inquisitiveness, and cognitive maturity. These 7 'habits of mind' can be thought of as the elements in our character that impel us toward using critical thinking skills. Each scale has subscales that are totaled for the student score. In addition, there is a total score from all 7 scales. For each subscale, a score below the cut score of 40 represents a general weakness in that area, while a score above the cut score of 50 indicates consistent strength in that area. A total score below 280 shows serious overall deficiencies in the student's disposition to think critically, while a score greater than 350 shows an overall strength (Phillips et al.2004, p.2).

It can be seen that the inventory measures both critical thinking skills and the disposition to think critically. The instruments of the CCTDI, including its Chinese version, have been shown to be highly valid and reliable (Luo &Yang 2001, p.51; Liu & Jin 2012, p.106). Therefore, I intend to use this inventory as an instrument for experimental pretest and posttest.

4. Potential Risks of the Experimental Design

The nature of educational experiments should be to confirm the correlation between educational phenomena through quantifying. The proposed experiment will adopt a randomized Solomon four-group design because it provides 'the best control of the threats to internal validity and will result in adequate statistical power' (Fraenkel & Wallen 1990, p.238). The measurements or observations are collected at the same time for all groups. The diagram of this design is as follows:

Table 1. The Solomon four-group design

Group		Pretest	Treatment	Posttest
1	R	O	X	O
2	R	O		O
3	R		X	O
4	R			O

Note. R=randomization, O=outcome measure, X=treatment.

A shortcoming of this design is that it requires a large sample, and conducting a study involving four groups at the same time requires a considerable amount of energy and effort on the part of the researcher (Campbell & Stanley 1963; Fraenkel & Wallen 1990). However, this EAP course will last for a whole year (first semester for treatment groups and second semester for control groups¹) and a lecturer of psychology will help me to monitor the experimental process and decode the data, so there must be enough time and people available to do that.

The second stage of the design is to determine the variables. According to the assumptions, the only optional variable (independent variable) of the experiment is the students involved in the critical thinking training course. The students in the control groups are individual learning students. They are in their natural state of self-

¹ 'The Challenge of Ethics' section will provide the details.

development and maturity, not having taken any critical thinking training. Dependent variables are the change in the overall disposition of the students' critical thinking and the changes in various components ('seven scales' of the CCTDI).

Nevertheless, some extraneous variables may prove to be the potential risks that threaten the final experimental results. In my proposed experiment, the extraneous variables will be considered from six aspects: social factors, family background, learning environment, course content, teachers and students' self-factors. The subjects are fresh second-year undergraduates. These students have experienced one year of campus life and have been exposed to the same environment. They follow the same educational system and teaching orientation. The concept of teaching and learning are almost consistent for them. Moreover, they have been living and studying together for a year, and they influence one another. All of these situations can keep these extraneous variables like social factors and family background constant. Furthermore, the experimental area will be located in a teaching building at the Southern University of Science and Technology in China. The treatment groups will attend the same EAP course developed by the Center for Language Education and be taught by the same teacher, so the learning environment, course content, and teachers are also constant. As for the students' own characteristics, such as their age and gender, I will create an inner group balance (Fraenkel & Wallen 1990, p.135).

After the determination of the variables, the sampling procedure becomes the next problem in the research, requiring special attention. Although drawing conclusions about a population after studying a sample is never completely satisfactory (Fraenkel & Wallen 1990, p.79), the diversity between the sample and the target population is likely to be insignificant if the sample is randomly selected and of sufficient size. On account of the quantitative nature of this research, probability sampling will be used to select students. Meanwhile, economic factors and practical principles should also be taken into account.

In order to guarantee the external validity of this study, the sample should be of sufficient size to satisfy the statistical demand and represent the population. The population consists of 1035 second-year undergraduates studying science and engineering subjects at Southern University of Science and Technology. A representative sample will be selected from the population as research subjects. One concern associated with the sampling method is that there are a few guidelines with regard to the minimum number of subjects needed. If the sample size of quantitative studies is calculated using a sample size calculator provided by the global panels of the GMI², the sample should consist of at least 340³ individuals. However, the characteristics of experimental research and the frequent interaction on the critical thinking course determine that the sample cannot be too large to avoid affecting the teaching effects. Fraenkel and Wallen (1990) recommend that the minimum number of subjects should be 30 individuals⁴ per group for experimental study. To solve the problem related to the sample size calculation, I believe that this recommendation can maintain the balance between effective treatment and statistical demand. In view of this, two alternative sampling strategies can be adopted.

Table 2. Alternative sampling strategies

Option A:	Option B:
1. Conduct norm-referenced tests among all 1035 2 nd -year undergraduates.	1. Adopt stratified sampling to select 120 students from 13 departments for the experiment.
2. Randomly select 150 students from them and make equal assignments to the treatment groups and control groups. ⁵ Match students according to certain variables to maintain the balance of group characteristics.	2. Match students according to certain variables and make equal assignments to the treatment groups and control groups.

The first option aims at controlling the effect of pretest on the experimental results, but the sample size is relatively large. 1035 students will receive the pretest and 150 will receive the posttest. The number of students

² GMI: Global Integrated Solutions for Market intelligence

³ It is with a confidence level of 95% and a confidence interval of 3.

⁴ In the statistics, a sample size of more than 30 individuals qualifies as a big sample (Wang 2001, p. 12).

⁵ It is a 125% selection to prevent subjects being lost and guarantee at least 30 individuals per group.

who receive the pretest is too large to manage, and this sampling method's cost is too high. The second option is easy to manipulate because of the small sample size and low cost, but it will lose subjects during the selection process, and the norm cannot be established. Thus, it cannot be regarded as a sampling technique that offers high external validity.

To overcome these weaknesses and ensure that the sample complies with the statistical disciplines, I decided to use the following sampling procedure. First of all, pilot sampling was conducted to test its feasibility. If it is feasible, stratified random sampling will be adopted to select a sample of 150 as the norm from 1035 second-year undergraduates (the target population). The sampling rate depends on the percentage of individuals in each department. For example, there are 26 individuals at the Department of Physics and 61 in the Department of Biology, and the sampling rate will be calculated as follows:

$$\% \text{ Physics} = (26/1035) \times 100 \approx 3$$

$$\% \text{ Biology} = (61/1035) \times 100 \approx 6$$

Then, they will be randomly and equally assigned into four groups (two groups are the treatment groups while the other two are the control groups.) and balance them in terms of their gender and age. According to the results of the pilot sampling test, this sampling method is not only in line with the statistical principles, but also the sample will be more representative of the population and reduces the costs.

Another potential problem in this study is that the data analysis of the Solomon four-group design lacks certainty concerning the proper statistical treatment and is constantly disputed by researchers. This forms another problem for my study. Campbell and Stanley (1963) made some preliminary suggestions based on its statistics but still neglected many details. In terms of my research, all possibilities that may arise in my data analysis should be carefully considered.

If the effect of pretest or the interaction between pretest and treatment could be neglected, the data will be analyzed by a one-way analysis of variance (ANOVA) to test and compare the posttest means of the four groups. If the effect of pretest cannot be confirmed, the test will be a two-group analysis of covariance (ANCOVA) in the posttest scores to compare the treatment group and the control group regarding pretests, covarying the pretest scores. At the same time, the T-test will be used for the treatment group and the control group without pretest. How to combine these two tests used to be a difficulty, but it has been solved by a meta-analytic approach (Glass 1978). If the ANCOVA and T-tests both reached the significant levels, the results of the experiments should be affirmed. Otherwise, I have to consider the effect of pretest and the interaction between pretest and treatment. The test for this would be a 2×2 ANOVA (Table 3). The means in the horizontal grids can be considered as the main effect of treatment. The means of the outcome measures in the column can be regarded as the main effect of the pretest. The means in the cross-lattice can be considered as the experimental interaction:

Table 3. 2×2 analysis of posttest scores

	Treatment groups	Control groups
With pre-test	O1	O3
Without pre-test	O2	O4

Note. O = outcome measure

5. The Challenge of Ethics

Many ethical issues arise during the stage of the data collection and analysis in an educational experiment. The bigger challenge facing any use of experimental designs in educational research may be 'an ethical rather than a technical one' (Gorard 2001, p. 144).

Above all, researchers need to respect the participants in research (Cohen *et al.* 2007). In this study, all of the groups of students must be guaranteed to participate voluntarily and have the right to withdraw at any time, so that 'the individual is not being coerced into participation' (Cresswell 2003, p.64). Nevertheless, this may lead to the insufficient samples and seriously affect the results of the experiment, which should be avoided to the best of the researcher's ability before and during the experiment. Therefore, maintaining the balance between the

experimental design and ethical consideration is crucial to the success of the experiment. In terms of the real situation, I carefully designed the sampling approaches and will conduct pilot sampling. In addition, testing the experimental results twice and giving the participants' rewards will also be used to control the withdrawal of the students. Participants also have the right to know the purpose, procedure, and benefits of my study, so that they can understand the nature of the experiment, and know what to anticipate and its likely impact on them.

At the stage of the data collection, I need to anticipate 'the possibility of harmful information being disclosed during the data collection process' (Cresswell 2003, p. 65). In this situation, the privacy of the individuals involved in my experiment ought to be protected. In the interpretation of the data, although I should provide an accurate account of the data, the anonymity of individuals, roles, and incidents in my project should also be considered. The language or words that are biased against persons because of their age, gender, racial group, or disability, should not appear in my description, although some of them may be vital to my research. Once the analysis is finished, the data from my experiment will be discarded in order to avoid it falling into others' hands for other purposes.

The final important issue relating to ethics is unique to experimental studies. It is that the experiment should not be discriminatory (Gorard 2001, p.146). The "PAD + microlectures" EAP teaching model is supposed to be helpful for Chinese undergraduate students' learning. I need to collect data and continue the treatment so that all of the participants, including both the treatment and control groups, ultimately receive the benefits of the EAP courses with such a model.

6. Conclusion

From the above analysis, it has been shown that the topic of the educational experiment should have both theoretical and practical value. This value enables the research to be significant in terms of generalization and future guidance. With respect to the hypotheses, all of them should employ certain testing methods. If researchers cannot find suitable methods for testing the hypotheses during the experimental process, those hypotheses must be laid aside and lose their significance. What is more, the most rigorous form of experimental research should be conducted using effective instruments.

In addition, the control of the extraneous variables is an integral part of the experimental procedure. Sampling strategies cannot be employed only for their convenient and easy manipulation, but should also be concentrated on the validity and reliability of the samples. In quantitative studies, the samples should be large enough to generalize the findings using statistical techniques. As for the data analysis, the methods and accuracy should be considered at the same time in order to guarantee the validity and reliability of the experimental results, so that they are ready for the hypotheses' testing. Then, any ethical issues arising from the process of the experiment ought to be immediately noticed and well handled. In conclusion, all of the experimental education designs should follow the five disciplines: 'objectivity, feasibility, maneuverability, effectiveness and innovation' (Dai 1986).

References

1. Campbell, D. & Stanley, J., (1963). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.
2. Chen, A., Wang, Y. & Jiang, J., (2015). Current problems and countermeasures of microlecture development, *Modern Educational Technology*, 25(10), pp. 67-72.
3. Cohen, L., Manion L. & Morrison, K., (2007). *Research methods in education*. 6th ed. London: Routledge.
4. Cresswell, J., (2003). *Research design: qualitative, quantitative, and mixed methods approaches*. 2nd ed. London: SAGE Publications.
5. Dai, Z., (1986). *Psychological and educational measurement*. Shanghai: Huadong Normal University Press.
6. Dewey, J., (1909). *How we think*. Boston: D.C.Heath and Co.
7. Fisher, A., (2001). *Critical thinking: an introduction*. Cambridge: Cambridge University Press.
8. Fraenkel, R. J. & Wallen, E.N., (1990). *How to design and evaluate research in education*. New York: McGraw-Hill.
9. Glass, V. G., (1978). Integrating findings: the meta-analysis of research. *Review of Research in Education*. 5, pp. 351-379.

10. Gorard, S., (2001). *Quantitative methods in educational research*. London: Continuum.
11. He, F., (2005). To foster college students' criticize thinking. *Journal of Education Institute of Taiyuan University*, 23(3), pp. 75-77.
12. Liu, H. & Jin, L. (2012). An empirical research on the correlation between English debate and Chinese undergraduates' critical thinking development. In Gong. Y. (Eds.), *Oral English teaching and critical thinking development* (pp. 101-112). Beijing: Foreign Language Teaching and Research Press.
13. Luo, Q. & Yang X., (2001). Revision for CCTDI (Chinese version). *Psychological Development and Education*, 3, pp. 47-51.
14. Paul, R & Elder, L., (2006). *The miniature guide to the art of asking essential questions*. Dillon Beach, CA: Foundation for Critical Thinking.
15. Phillips, R. C. *et al.*, (2004). The California critical thinking instruments for benchmarking, program assessment, and directing curricular change. *American Journal of Pharmaceutical Education*, 68(4), pp. 1-8.
16. Sun, Y., (2011). Focusing on developing English-majored undergraduates' critical thinking: deepening the teaching reform. *Foreign Languages in China*, 8(3), pp. 49-58.
17. The National University Teacher Network Training Center, (2014). *Research Report on Microlectures at Chinese universities*. Beijing: Chinese Ministry of Education.
18. Wang, X., (2001). *Educational statistics*. Shanghai: Huadong Normal University Press.
19. Wei, C., (2016). The application of PAD + microlectures model in college English teaching, *Language Research Journal*, 3, pp.92-93, 113.
20. Yao, L., (2001). Research on enhancing college students' critical thinking abroad. *Higher Education of Sciences*, 5, pp.18-21.
21. Zhao, H., (2016). An Investigation into PAD College English class, *Education Forum*, 13, pp.1-2.
22. Zhang, X., (2017). *PAD Class: The New Wisdom of Education in China*. Beijing: Science Publishing House.
23. Zhang, Y., (2016). Advantages and challenges of microlecture-assisted English PAD class. *Language, Art and PE Research*, 168(11), pp.347, 349.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).