

# Using Burke's Dramatistic Pentad to Interpret Chinese "Gao-Kao" High Stakes Testing and Stressing – Paralleled Testing in the U.S. as Cross-Cultural Context

KOME – An International Journal of Pure  
Communication Inquiry  
Volume 1 Issue 2 pp 55-63.  
© The Author(s) 2013  
Reprints and Permission:  
kome@komejournal.com  
Published by the Hungarian Communication  
Studies Association

Lei Wang <sup>a</sup>, Xiaoting Huang <sup>b</sup>, Jim Schnell <sup>c</sup>

<sup>a</sup> National Education Examinations Authority, Beijing, China

<sup>b</sup> China Institute for Educational Finance Research, Peking University, Beijing, China

<sup>c</sup> U.S. Army Culture & Foreign Language Directorate, Fort Eustis, Virginia, USA

**Abstract:** The Chinese College Entrance Examination (“*Gao-Kao*”) is the most high stakes assessment in China and parallels the most competitive examinations globally. Although it can provide Chinese educators and policy makers with an enormous pool of information about student achievement growth, school efficiency, etc., the current use of the test is mainly limited to ranking students by their raw scores. In this study, we tried two modifications to the traditional test to connect the assessment outcomes with school accountability. First, we linked the Gao-Kao English tests from 2010 and 2011 and aligned them on a Rasch scale. Secondly, we collected background information of the examinees via a background survey.

The result showed that students from Hainan province improved a little in 2011 overall. In addition, school level reports were generated to show the school's growth as well as the county and province averages. By implementing test equating and background survey measures, this study demonstrated that Gao-Kao data can be used to construct a longitudinal data source as an initial step to build a value-added school accountability system.

The aforementioned findings and how they are communicated help to frame global use of such high stakes testing. Kenneth Burke's Dramatistic Pentad is used as foundation for communicative interpretation of these findings. The international context provides backdrop within which the findings are nested. Contrast with testing in the U.S. serves to highlight unique features of the Gao-Kao examination approach.

**Keywords:** Gao-Kao, test equating, school accountability, Dramatistic Pentad

## Introduction

The College Entrance Examination in China, called “*Gao-Kao*” in Chinese, is the most high stakes assessment in China and parallels the most competitive examinations globally. Two days each year, millions of high school graduates and people with equivalent educational qualifications take the test. Students with higher Gao-Kao scores get into better universities. They can get better jobs after graduation, and eventually, become winners in the thriving economy. Hence, the test is considered to be the most critical turning point in every student’s life, and studying for it can never be over emphasized.

This type of high stakes examination has parallels with other national college entrance examinations, such as the ACT and SAT in the U.S., but the U.S. examinations do not carry the weight that Gao-Kao does. The U.S. system considers other factors within the college placement process.

Under such high pressure, over use of the Gao-Kao test score seems inevitable. In many places, high schools are ranked by their average Gao-Kao scores and teachers are rewarded by their class averages. As a result, many people blame Gao-Kao for causing bad educational practices such as teaching to the test and social problems such as creating students who are test-taking “machines” with limited creativity. Thus, the Gao-Kao examination manifests symbolic meanings that correlate with stress and intense competition.

Ironically, the Gao-Kao data has seldom been used for important educational policy making decisions. China’s NEEA (National Education Examinations Authority) is directly under the MOE (Ministry of Education). The NEEA exerts great effort to ensure the quality of the test questions and the reliability of the tests. There is not examination continuity from year to year. That is, unlike the SAT or ACT exams in the U.S., Gao-Kao scores from different years cannot be compared directly. It is not possible to tell whether the difference in Gao-Kao scores from one year to another is the result of changes in student proficiency or a shift in item difficulty. Findings from investigations on trends regarding education quality in individual schools, regionally, or on national levels remain untapped.

Modern score equating techniques (Kolen & Brennan, 2004) provide the NEEA with a tool to make better use of the Gao-Kao data. The goal of equating is to produce a linkage between different test forms so that the scores from each test have the same meaning and can be compared directly. Analysis of student growth becomes possible when Gao-Kao tests over successive years can be equated.

Moreover, it has been widely acknowledged that one-time assessment scores are not a fair way to compare teachers or schools since students come to school with different backgrounds (Doran, 2003). This is significantly different, in contrast with the U.S., where multiple factors are considered for college entrance decisions. Such factors include aptitude testing, high school GPA (grade point average), extra-curricular activities and unique life experiences.

Over the last decade, value-added analysis in China has become the most promising tool to evaluate school effectiveness. The idea behind the value-added approach is simple. School quality is determined by the increase in student knowledge and skills, extracting the impact of non-school factors such as the student’s family SES (socio-economic status), etc. (Ballou & Sanders, 2004). It is strongly desired that the Gao-Kao data be used in this way since value-added accountability models can greatly motivate teachers and schools (Doran & Fleischman, 2005).

In this study, we took data from Hainan province as an example and implemented two technical modifications to the Gao-Kao English test. First, we linked the tests from 2010 and 2011, and aligned them onto a Rasch scale (Kolen & Brennan, 2004). Secondly, we collected background information from the examinees via a background survey. The results were used

to examine student growth and were further applied to construct a value-added school accountability model.

All of the aforementioned, taken together, offer relevant themes for interpretation using Kenneth Burkes Dramatistic Pentad (Golden, Berquist, Coleman & Sproule, 2011) as a foundation for framing the varied communicative elements. The Dramatistic Pentad stresses the act, agents, scene, agency and purpose (which will be explained later in this report).

## Methods

### Test Equating

Test equating seeks to produce comparable scores for examinees who take different editions of the same test. Researchers have developed many data collection designs, such as the single group design, the equivalent groups design, and the anchor test design (Holland & Dorans, 2006). An anchor test design allows for a new test to be administered to a sample of examinees from each test-taking population. It is most appropriate in high-stakes situations when item reuse leads to test security problems. So we chose to use the Nonequivalent Groups with External Anchor Test (*NEAT*) design (von Davier, Holland, & Thayer, 2004) in this study. Specifically, a 28-item anchor test was administered to two groups of examinees one month before Gao-Kao in 2010 and 2011. The details are shown in table 1.

	2010 English test	2011 English test	28-item anchor test
2010 Gao-Kao participants	X		
Sample of 2010 participants	X		X
2011 Gao-Kao participants		X	
Sample of 2010 participants		X	X

**Table 1: The data collection design for equating 2010 and 2011 Gao-Kao English test**

The NEAT design poses strict requirements on the quality of the anchor test as it greatly impacts the accuracy of equating. It needs to measure the same construct of the full tests. Even though it is usually shorter and less reliable than the full tests, it is desirable because of the high quality variables. In this study a team of professional item developers, working for NEEA, were hired to construct representative common-item sets. The anchor test was built to the same test specifications as Gao-Kao, except that it did not contain Listening and Writing sections. More than 50 items were developed and administered originally. Items that displayed undesirable psychometric properties were excluded from the analysis and only 28 items were used in the final equating procedure. The number of anchor items exceeded 20% of the total length of Gao-Kao, meeting the rule of thumb proposed by Kolen and Brennan (2004). One thousand-four hundred-seventeen examinees in 2010 and 580 in 2011 were sampled using the anchor test.

The next step of test equating is to produce comparable scores for the 2010 and 2011 tests. Different procedures to convert scores are available (Kolen & Brennan, 2004). In this study, we chose to align the two tests onto a Rasch scale mainly for two reasons. First, the Rasch scale is considered to be an objective scale because the difficulty of an item is independent of student abilities and the ability estimates are independent of the items (Wilson, 2005). Secondly, data from both years can be analyzed simultaneously via a concurrent estimation

procedure (Kolen & Brennan, 2004). Scale scores produced by the software are directly comparable. Computer software *ConQuest* (Wu, Adams, & Wilson, 1998) was used to scale the data in this investigation. Finally, the Rasch scores were converted to follow the NEEA process of multiplying by ten then adding 50 ( $Y_{\text{final score}} = 10 * X_{\text{Rasch score}} + 50$ ).

### Background Survey

To extract non-school impacts on Gao-Kao outcomes it is necessary to collect student background information, including students' family socio-economic status, parents' education level, occupation, and students' after school learning activities (Strand, 2011).

So we conducted a 25-question online survey in both years. After filling out the Gao-Kao online registration, examinees were prompted to decide whether they would participate in the survey. The response rate reached a high of 80% because we informed the students beforehand and asked teachers to encourage their students to take it.

Correlations between the non-school factors with student achievement scores were examined. The background information also enabled further discussion regarding trends linked to school quality and education equity.

### **Results**

Altogether, our focus on Hainan revealed that 54,100 students participated with the Gao-Kao English test in 2010 and 53,755 students did so in 2011. The overall reliability in the concurrent estimation is as high as 0.94. The average final score increased slightly from 54.2 in 2010 to 55.1 in 2011. The change is not statistically significant ( $p < 0.01$ ). Figure 1 shows the Wright Map (Wilson, 2005) for the two successive years aligned on the logit scale. The right hand side is the distribution of the examinee abilities estimations, while the left hand side is the distribution of item difficulties.

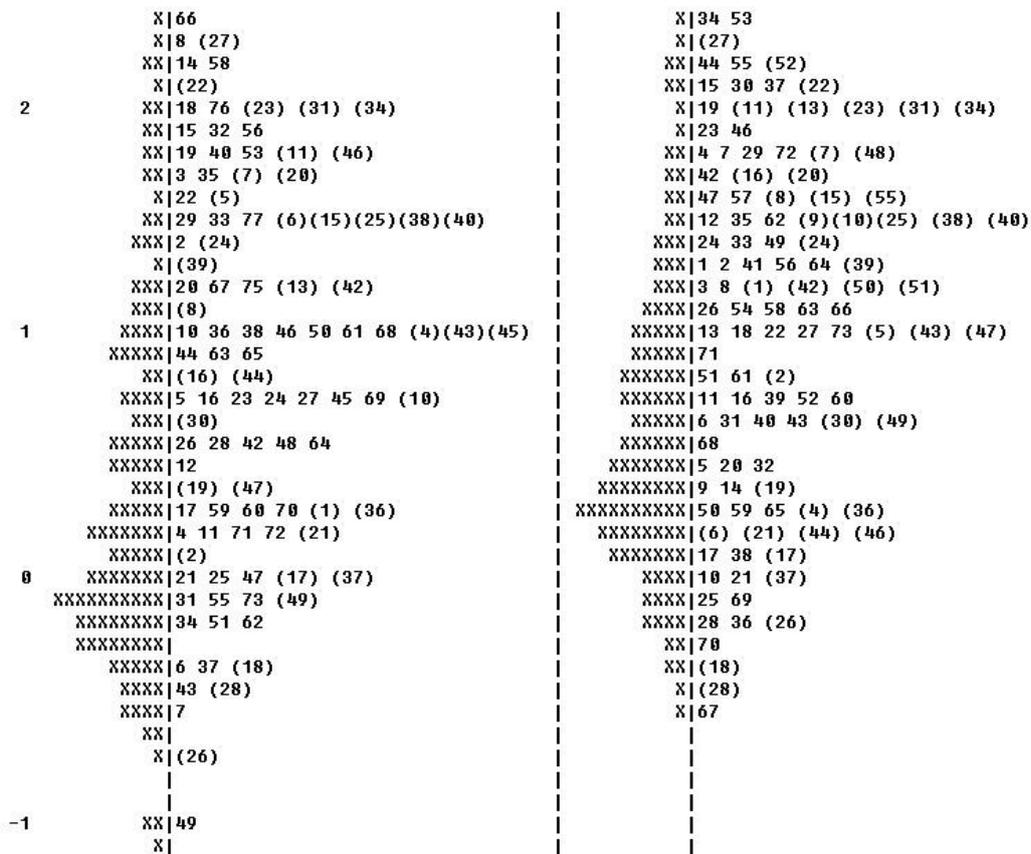
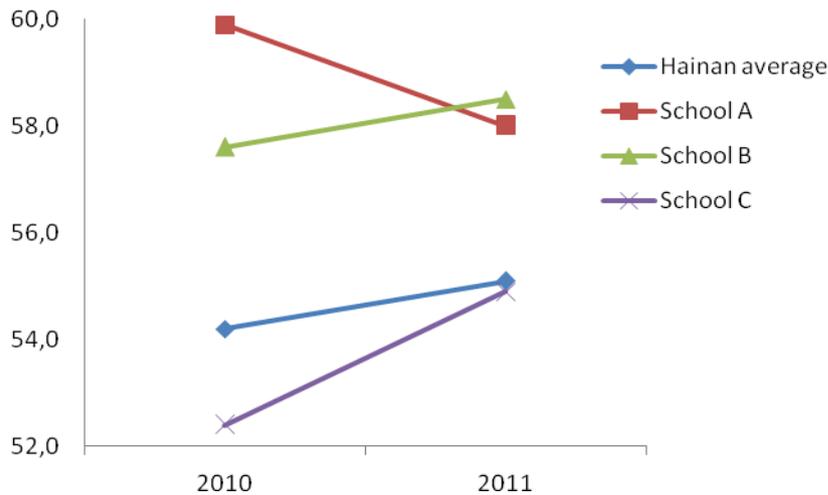


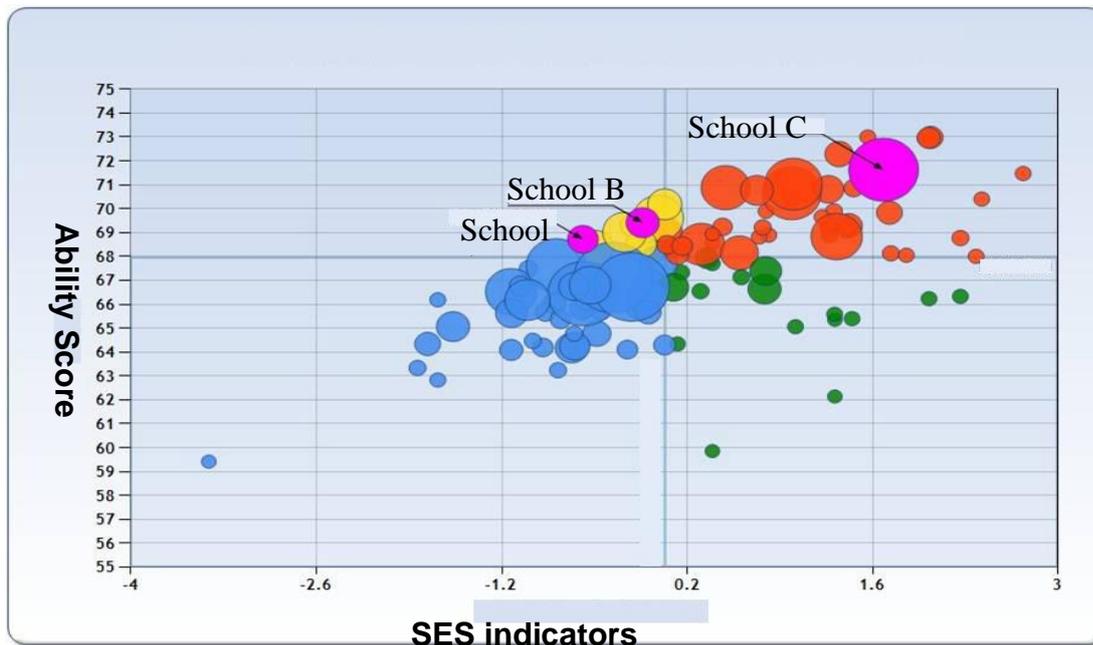
Figure 1: Wright Map of the 2010 and 2011 Gao-Kao English Test including anchor items

Schools can easily compare their performance trend with the province averages. Each school had access to review straightforward reports and graphical representations of its own performance as well as the district and province averages on the project’s website. Evaluating school performance longitudinally can affect educators and impact relevant policy decisions. For example, Schools A, B and C had average scores of 59.9, 57.6 and 52.4 in 2010, respectively. In 2011, they scored 58.0, 58.5 and 54.9. When ranked by one-time scores, School A was number one in 2010, number two in 2011, whereas School C was the worst in both years. However, when we looked at the schools longitudinally, we found that School C made the greatest improvement among the three, but School A actually did worse in that regard. Figure 2 demonstrates the trend of student learning for the three schools compared with provincial averages.



**Figure 2: Measuring school performance longitudinally**

Finally, student background information was analyzed. The family SES indicator was found to be the most important factor associated with student English achievement. Specifically, aggregated at school level, lower performing schools tended to have lower average SES. Figure 3 showed how school average scores increased with the SES index. (In the graph, each bubble represents one school. The size of the bubble is proportional to the number of examinees in the school.) This indicates that lower performing schools are not necessarily less effective since non-school factors play a significant role in influencing student learning outcomes.



**Figure 3: The relationship between SES and the scaled English score at school level**

### Discussion

Ranking schools or teachers by a single year’s Gao-Kao scores often leads to negative or damaging impacts on school culture and the instructional programs. As first steps to promote

better usage of the Gao-Kao data, we equated tests from 2010 and 2011 and showed how longitudinal analysis might be carried out to inform educators and policy makers.

In addition, we collected and analyzed student background data. In many well-known large-scale assessments, such as PISA (Programme for International Student Assessment) and TIMSS (Trends in International Mathematics and Science Study), background information is always collected, and statistical analyses are routinely carried out to examine factors that significantly impact students' learning outcomes. However, this part has been largely neglected in China in the past.

In this study, we found that SES was the most prominent non-school factor that correlated with students' English achievement. This finding is in accordance with previous research (Perry & McConney, 2010). This provides Chinese researchers and policy makers with more information on education equity and quality.

The test equating practice and survey are our initial attempts to build a value-added school accountability system. However, it is important to note that it would be necessary to build a vertical scale (Briggs & Weeks, 2009) that aligns student ability from their entrance point to Gao-Kao in order to construct such a model. We started to collect data from cohorts that entered high school in 2010 and 2011. But it has not been long enough for us to follow them until their Gao-Kao examination time. The vertical scale is still under construction. As a result, this report is limited with regard to discussions about equating results between the two Gao-Kao English tests, which is a very important initial step in the big picture.

In addition, student growth in one subject is a very narrow angle to investigate school quality. Equating tests on other subjects is our next mission. It is more complicated to equate math and science tests and correspondingly build vertical scales as these subjects consist of several sub-areas. More efforts to develop high-quality anchor tests are called for. And multidimensional equating procedures may be applied to tackle the problem (Oshima, Davey, & Lee, 2000).

Moreover, because the accuracy of equating is critical in providing valid information for high-stakes policy decisions (Peterson, 2007), it will be worthwhile to equate with different models and evaluate whether differences in equating functions have practical significance. To construct statistically and theoretically sound value-added models, additional information on teacher, school and district characteristics is needed. Teacher and school principal surveys are being developed.

This high stakes testing scenario, as exemplified via the Gao-Kao examination phenomena in China, is ripe for interpretation using Kenneth Burke's Dramatistic Pentad (Golden, Berquist, Coleman & Sproule, 2011) that is commonly used for framing communicative elements. The Dramatistic Pentad is composed of five parts: 1) the act, 2) agents, 3) scene, 4) agency and 5) purpose.

The act is the focal center within which the other parts function. The agents are the individuals involved. These agents operate in a scene or situation. They employ an agency, or means, so they can accomplish a specific purpose.

The act, in this high stakes testing scenario, can be understood as the high drama associated with the completion of a single exam that will have crucial significance on the direction the examinees lives will take (regarding opportunities for advancement in life). The agents are the students primarily and their families secondarily. The scene, or situation, involves the interactive dynamics that result in an environment that is exceedingly high stress and very relevant for students and their corresponding relationships with their families and the larger social order.

The agency, or means, through which the agents will progress centers upon student performance on this single Gao-Kao examination. The ultimate purpose being to score well

on the Gao-Kao examination and use this positioning as foundation for advancing toward a higher quality of life than would be available if a lower score is achieved.

## Summary

With careful test equating and background surveys, we set out to build a longitudinal data source with China's most important test, the Gao-Kao. This marks the starting point where China's testing practitioners and researchers can seek to utilize high-stakes assessment data to ensure school accountability.

The aforementioned, placed within a global context, reveals how such high stakes testing can come to be symbolically represented via various communication channels. The factual clarifications and delineations in this report serve to demystify such high stakes testing so as to make it understandable and, thus, minimize confusion, resentment and despair by the general public.

In contrast with paralleled testing in the U.S., such as the SAT and ACT examinations, we find similar kinds of testing but less emphasis being given to the examination itself due to consideration for other (non-examination) factors. Thus, there is considerably less anxiety associated with such examinations in the U.S. and less anxiety conveyed in that regard via relevant communication channels.

Kenneth Burke's Dramatistic Pentad is offered as a framework for understanding this high stakes testing scenario. Such emphasis on understanding of act, agents, scene, agency and purpose serve to frame this scenario so it can best be understood via the communicative variables that give it meaning.

We do not believe one approach, either Chinese or U.S., is to be preferred over another. Ironically, at the present time, we observe that the U.S. is seeking means to give such high stakes testing more emphasis as a college entrance consideration and the Chinese are considering measures for redefining the role of high stakes testing such as Gao-Kao.

We do believe communicative representations of high stakes testing are key in molding public perceptions of such testing over time. How the public interprets such representations impacts their responses to the examination process. Cross-cultural understanding of such impacts will serve to augment the fund of information that can be considered regarding improvements. This report is intended to be a contribution to that fund.

## References

- Ballou, D., & Sanders, W. (2004). Controlling for student background in value-added assessment of teachers. *Journal of Educational and Behavioral Statistics*, 29(1), 37-69.
- Briggs, D. C., & Weeks, J. P. (2009). The impact of vertical scaling decisions on growth interpretations. *Educational Measurement: Issues and Practice*, 28(4), 3-14.
- Doran, H. C. (2003). Adding value to accountability. *Educational Leadership*, 61(3), 55-59.
- Doran, H. C., & Fleischman, S. (2005). Challenges of value-added assessment. *Educational Leadership*, 63(3), 85-87.
- Golden, J.L., Berquist, G.W., Coleman, W.E. & Sproule, J.M. (2011). *The Rhetoric of Western Thought* (10th ed.) Dubuque, Iowa: Kendall-Hunt.
- Holland, P. W., & Dorans, N. J. (2006). Linking and equating. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 187-220). Westport, CT: Praeger.

Kolen, M. J., & Brennan, R. L. (2004). *Test equating, scaling, and linking: Methods and practices* (2nd ed.). New York, NY: Springer.

Oshima, T. C., Davey, T. C., & Lee, K. (2000). Multidimensional equating: Four practical approaches. *Educational Measurement, 37*, 357-373.

Perry, L. B., & McConney, A. (2010). Does the SES of the school Matter? An examination of socioeconomic status and student achievement using PISA 2003. *Teachers College Record, 112*(4), 1137-1162.

Peterson, N. S. (2007). Equating: Best practices and challenges to best practices. In N. J. Dorans, M. Pommerich & P. W. Holland (Eds.), *Linking and aligning scores and scales*. New York, NY: Springer.

Strand, S. (2011). *Ethnic, gender and socio-economic gaps in achievement: The perils of 'main effects'*. Paper presented at the Annual Conference of the American Educational Research Association (AERA), New Orleans, USA.

von Davier, A. A., Holland, P. W., & Thayer, D. T. (2004). *The kernel method of test equating*. New York, NY: Springer.

Wilson, M. (2005). *Constructing Measures: An Item Response Modeling Approach*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Wu, M., Adams, R. J., & Wilson (1998). ConQuest. Hawthorn, Australia: ACER Press.