Order Number 9402723

An empirical model for prediction of students' science achievement in the United States and the People's Republic of China

Wang, Jianjun, Ph.D.

Kansas State University, 1993

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

AN EMPIRICAL MODEL FOR PREDICTION OF STUDENTS' SCIENCE ACHIEVEMENT IN THE UNITED STATES AND THE PEOPLES' REPUBLIC OF CHINA

by

Jianjun Wang

B.S., Anhui Normal University
M. S., Beijing Normal University
M. S., Kansas State University

A DISSERTATION

submitted in partial fulfillment of the requirement

for the degree

DOCTOR OF PHILOSOPHY

Curriculum and Instruction
College of Education
KANSAS STATE UNIVERSITY
Manhattan, Kansas

1993

Approved by:

John R. Staver

Major Professor

ACKNOWLEDGEMENTS

I am most grateful to Professor John R. Staver, my major advisor, for inviting me to pursue the Ph.D. in science education. I knew Dr. Staver when I was the National Secretary of the International Association for the Evaluation of Educational Achievement (IEA). Dr. Staver had many publications in neo-Piagetian research and some of them were cited in Chinese graduate textbooks such as <u>Research in Education</u>. I applied for graduate study at Kansas State University because Dr. Staver is a member of the faculty.

I have experienced dramatic academic growth during the last three years. To support my doctoral study, I have taken more than sixty credit courses in education, physics, sociology and statistics beyond my Master degree. Based on the systematic course work, I obtained another MS degree in statistics and passed the Ph.D. Qualifying Examination in the Department of Statistics. It was Dr. Staver's insightful advice that guided me through the study process. My special gratitude is also due to other professors who served on my supervisory committee, Dr. Lawrence Scharmann, Dr. Leonard Bloomquist, Dr. Paul Burden and Dr. Dean Zollman, for their immeasurable assistance and steadfast support. Thanks to Professor Bob Smith, the outside chairperson of the final oral defense of my Ph.D. dissertation, for his valuable research suggestions and encouragement.

Other individuals who have been especially helpful to me and deserve special appreciation are: Ms. Marilyn Adams, Dr. Steve Oliver, Dr. Andrew Lumpe, Professor Li Ping, Professor Wang Yongzhen, Dr. Malcolm Rosier, Mr. Li Wei, Dr. Luo Ningyi, Dr. Wu Jianjun, Dr. Charles Heerman, Dr. Larry Enochs, Dr. Gail Shroyer and my wife, Rong Rong, daughter, Libing Wang and mother, Yangzhi Wang.

CONTENTS

LIST OF FIGURES

LISTS OF TABLES

CHAPTER 1

INTRODUCTION

The United States of America (USA) and the People's Republic of China (PRC) are two leaders in the world community. Almost equal in land area, China has nurtured the largest population in the world, and the United States has developed the most advanced industrial and technological civilization. In both countries, education is a top priority for national development. China is enforcing compulsory education up to the 9th grade level. The United States wants its students to be first in the world in science and mathematics achievement (President Bush's initiatives, January 31, 1990). A comparative study of students' science achievement may provide valuable information for improving Chinese and American students' achievement in science.

BACKGROUND

The most systematic international studies of school science education to date are the First IEA Science Study (FISS) and the Second IEA Science Study (SISS) (Keeves and Rosier, 1981). IEA, the International Association for the Evaluation of Educational Achievement, is an international research organization. Members of IEA are major educational research institutions from each participating country. The broad purpose of IEA research is to study the relationship between relevant input factors in social, economic, and pedagogical realm and output as measured by performance on international tests measuring both cognitive and non-cognitive outcomes (Postlethwaite, 1974). According to Husen (1987), IEA has coordinated one of the most influential research efforts in the history of educational research, and it has conducted the best-known international research on education.

FISS is a part of an IEA six-subject survey in the 1970s which includes science, reading comprehension, literature, French as a foreign language, English as a foreign language and civic education (Comber & Keeves, 1973; Walker, 1976). The science education segment of the survey was carried out in nineteen countries, including the United States. China, however, was involved in the turbulent Cultural Revolution (1966-1976) and did not participate in the IEA study at that time. Diplomatic relations between the United States and China were not fully restored until 1979. Hence, no comparative research between the two countries was conducted before the 1980s.

SISS, like FISS, was designed to provide an overview of science education across the world. The project started in 1981 and involved twenty-three countries. The survey (Jacobson & Doran, 1988) was conducted at three population levels: 5th grade; 9th grade; and 12th grade . The development of the international instruments for each population was a collaborative effort involving all participating countries to ensure the fairness of the cross-national comparisons. Considering that the time for testing was restricted in many school systems, the IEA staff decided to structure the test items into a core and a set of rotated tests (IEA, 1988). Each student was required to take the core test and two of the rotated tests. An Opportunity to Learn (OTL) questionnaire was developed to reflect the extent to which the students had had an opportunity to learn the content tested by each question. The Chinese version of the instruments for the 9th grade population is presented in Appendix 1, and the instrument titles are listed in English in Table 1. The English version of the instruments is available at the IEA Headquarters, 2517 GK the Hague, Netherlands.

Both China and the United States participated in the SISS project. The results are summarized in a three-volume IEA publication (Rosier & Keeves, 1991; Postlethwaite & Wiley, 1992; Keeves, 1992). However, the information concerning the United States and China has not been sufficiently presented in IEA publications.

In the United States, the SISS data were first collected in 1983 and had low degrees of response at the 12th grade level. To ensure the quality of the survey, the United States undertook a second SISS survey in 1986 (Phase II). The U.S. SISS advisory panel directed that only data from the Phase II survey be used as the student data (Postlethwaite and Wiley, 1992). The direction was submitted to the IEA data-processing team at the University of Hamburg.

Table 1: Names of SISS Instruments

STUDENT BOOKLET

1. Core Test
2. Rotated Tests
3. Student Questionnaire
4. Mathematics Test
5. Word Knowledge Test
6. Opinion Questionnaire
7. Description of Science Learning

TEACHER BOOKLET

1. OTL Questionnaire
2. Teacher Questionnaire

PRINCIPAL BOOKLET

1. School Questionnaire

The feedback from IEA was:

> This "direction" was followed but it created two problems: the first was that multivariate analyses for the United States became virtually impossible because many variables were not administered in the second round of testing; the second was that several items were dropped at the second round of testing, since no rotated tests were employed at Populations 1 and 2, and several items were dropped from the biology, chemistry, and physics tests. Hence, the "direction" could not be followed for all data analyses as it would have eliminated comparisons involving the United States. In the cases where these data were used, the purpose was not to estimate population means or proportions, but to explore variability and assess relationships of particular student and school characteristics to achievement (Postlethwaite & Wiley, 1992, p. 7).

On the other hand, SISS in China was named the "SISS Pilot Study" because the survey was conducted at the 9th grade level in three large cities, Beijing, Tianjin, and Taiyuan (Rosier & Keeves, 1991). However, China is a developing country, and about eighty percent of its population lives in the countryside. The primary objective of the SISS Pilot Study was to help Chinese researchers to understand the IEA methodology. The survey results did not reflect the nature of Chinese science education since the population in rural regions was not included in the pilot study.

In 1988, the China IEA Center modified the IEA instruments and launched the SISS Extension Study (SES) at the 9th grade level in seven provinces. The common instruments employed in SES and the SISS phase II survey in the United States are listed in Appendix 1. The instrument revision was based on research experience accumulated from the SISS Pilot Study in China and the Phase I SISS survey in the United States. The information covers students' attitude, gender, classroom experience, personal effort, home background, and test scores. As suggested by the representative of the National Science Foundation, the data collected from the revised instruments have the quality for the IEA international assessment (Postlethwaite & Wiley, 1992, p.185).

Nevertheless, SES was originally designed for an inter-province comparison or "province-other country" comparison. Each of the seven provinces in China was treated as an independent system. Hence, there is no legitimate method for integrating the survey over the seven provinces. To use the SES data for an international comparison, one of the provinces must be identified to represent the Chinese situation.

In China, the eastern and southern areas are more developed than the northern and western areas. On balance, a central province is more representative of the entire country than a boundary province. Among the SES seven provinces, Hubei province is the only one located in the central area of China. Thus, the data from Hubei province is chosen in this research to represent the Chinese situation in 1988.

In summary, according to the United States SISS advisory panel, the Phase II SISS survey is better than the Phase I survey. In China, compared to the SISS Pilot Study, the SES data set has at least two advantages: (1) The SES data were collected in both urban and rural regions; (2)

the population in each province is larger than the population of the SISS Pilot Study. Hence, SES in China and the Phase II SISS survey in the United States provides the best opportunity to compare school science education between the two countries.

STATEMENT OF THE PROBLEM

Because the information collected by SES and the Phase II SISS surveys is very extensive, many questions can be explored through the analysis of the two data sets. The question addressed in this research involves empirical models for prediction of students' science achievement.

Most prediction models in education postulate a linear relationship between student science achievement and related personal, school, or social predictors (Dryden, 1987a). At least two major reasons exist for using a linear model. First, a linear model is simple. Since no other models have been consistently supported by theories in science education, it is tempting to choose a linear model in a preliminary exploration. Second, an infinite number alternatives to a linear model exist. Each alternative is based on a non-linear function, and it is impossible to identify an appropriate model from all potential non-linear functions.

The dilemma between linear and non-linear models, however, is approached differently in this research. First, an empirical approach is taken to construct the model of prediction. Because neither linear nor non-linear models are supported by present theories, it is desirable to explore an empirical model based on high quality data sets. Second, a unified Taylor series (Ayres, 1964) is adopted as the mathematical function in the model exploration.

In mathematics, under the condition of the Taylor Theorem (Ayres, 1964), a function can be expressed as a Taylor polynomial series. For example,

$$ax + b = b + ax + 0 \cdot x^2 + 0 \cdot x^3 + \ldots\ldots \qquad \text{Linear function}$$

$$e^x = 1 + x + (2!)^{-1} \cdot x^2 + (3!)^{-1} \cdot x^3 - (4!)^{-1} \cdot x^4 + \ldots\ldots$$

$$\sin x = 0 + x - (3!)^{-1} \cdot x^3 + (5!)^{-1} \cdot x^5 - (7!)^{-1} \cdot x^7 + \ldots\ldots$$

$$\cos x = 1 - (2!)^{-1} \cdot x^2 + (4!)^{-1} \cdot x^4 - (6!)^{-1} \cdot x^6 - \ldots\ldots$$

$$\ln (1+x) = 0 + x - (2)^{-1} \cdot x^2 + (3)^{-1} \cdot x^3 - (4)^{-1} \cdot x^4 + \ldots\ldots$$

$$\arcsin x = 0 + x + 1 \cdot (2 \cdot 3)^{-1} \cdot x^3 + 1 \cdot 3 \cdot (2 \cdot 4 \cdot 5)^{-1} \cdot x^5 + \ldots\ldots$$

$$\arctan x = 0 + x - (3)^{-1} \cdot x^3 + (5)^{-1} \cdot x^5 - (7)^{-1} \cdot x^7 + \ldots\ldots$$

Hence, the infinite number of model explorations is simplified to the identification of a set of polynomial coefficients. In many cases, a truncated Taylor series provides a good approximation of the original function.

The validity of the empirical approach depends on two conditions. First, one needs well-designed data sets to construct an empirical model. Second, to obtain a close approximation, the model should include high degrees of polynomials, and the data set must contain a large number of observations. Because the IEA data sets are well-designed and contain thousands of observations, they fulfill the two conditions necessary for the construction of a potentially valid empirical model of prediction.

RESEARCH QUESTIONS

The IEA data sets contain student information in six areas, gender, attitude, home background, classroom experience, personal effort, and science achievement. Each of the first five aspects could have linear or nonlinear effects on students' achievement. The questions that guide this research are:

1. What are the linear or nonlinear factors and interactions constructed by variables from the first five aspects which have significant effects on students' science achievement?
2. Do differences exist between the United States and China in terms of the factor structures and interpretations?
3. Do differences in complexity exist between the Chinese and American models, and how can the differences be explained based on differing educational, political, social and cultural contexts in each country?

SIGNIFICANCE OF THE STUDY

The United States and the Peoples' Republic of China are two leaders in the world community with strikingly different cultural, economical, political and historical background. China has a unified culture nurtured through more than 2000 years of history. Chinese education is supported by a trickle down economy, and the central government has the authority to determine the school curriculum. The United States, on the other hand, has a short 200 year history, a diversified culture, and a free market economy. The authority to determine the school curriculum lies with individual school districts and states. Despite these differences, the author believes that the U.S. and China can learn from each other and that comparative research such as the study reported here can provide information to improve the American and Chinese students' achievement.

Walberg (1983) points out: "The best and perhaps only test data that permit reliable international comparisons of science achievement were obtained by the International Association for the Evaluation of Educational Achievement" (p. 6). A fundamental problem, prediction of student achievement in science education is addressed in this comparative research based on two IEA data sets, SES and Phase II of SISS, collected from the United States and the People's Republic of China, through an empirical approach with a sound mathematical basis (Taylor Theorem). This research explores the questions which can not be answered by present theory.

In methodology, most path analyses use either a Partial Least Square Method (PLS) developed by H. Wold (1975, 1979, 1982) or Analysis of Linear Structural Relationships by the Method of Maximum Likelihood (LISREL) developed by Karl Joreskog and Dag Sorbom (1984). However, based on statistical decision theory, the least squares method and maximum likelihood method are inadmissible when the number of parameters is larger than two (Stein,1956; James & Stein,1961; Casella & Berger, 1990; Hebel, et. al.,1993). Stein-type shrinkage estimators have proved better than the unbiased estimators and are widely accepted by statistical researchers (Efron and Morris, 1977). As Jennrich and Oman (1986) point out, "It thus comes as a surprise that Stein estimation is not routinely used in regression applications -- we know of no statistical packages with Stein regression routines, and when shrinkage estimators are used in multiple regression models, they are more often ridge-type estimates" (p. 113). As a part of this research, a computer program was edited to compute shrinkage coefficients for construction of the empirical model.

DELIMITATIONS OF THE STUDY

Two significant delimitations are identified for this study:

1. The SES in China was conducted at the ninth grade level. The SISS in the United States also contains a ninth grade population. Hence, the comparative research focuses on ninth grade science achievement in the two countries.
2. Construction of the empirical model is based only on information in the data sets.

LIMITATIONS OF THE STUDY

Three limitations are embedded in the study:

1. In essence, the two country comparison is an approximation because the information about China is estimated based on the SES survey in Hubei province.
2. Because SES in China and Phase II of SISS in the United States were conducted in different years, the comparative research is essentially built on the American information in 1986 and the Chinese information information in 1988.
3. Other predictors which have no observations in the data sets are not included in consideration.

ASSUMPTIONS OF THE STUDY

The comparative study is based on the following assumptions:

1. A prediction model can be approximated by a truncated Taylor polynomial function;
2. The English and Chinese versions of the IEA instruments are mutually equivalent;
3. There are no coding errors in the two data sets.

DEFINITION OF KEY TERMINOLOGY

Attitude scale: A set of questions which indicate students' attitudes to certain events. The responses, such as agree, uncertain, and disagree, are treated as interval scale in this study.

Parents' Education: The highest level of school that a student's father or mother has completed.

Personal Effort: The amount of work a student did for his/her study, e.g., the number of hours a student spent on homework in each week.

Classroom Experience: Daily events a student experienced in his/her classroom, e.g , lab activities, tests and selected teaching materials.

Science Achievement: Students' scores on an IEA international core test (see Appendix 1).

SUMMARY

Prediction of students' science achievement is a fundamental question in science education. Yet, no prediction model is uniformly supported by theories. The research presented here explores a possible empirical model for prediction of students' science achievement in China and the United States. Construction of the model is based on the ninth grade data sets from the SES survey in Hubei province of China and the Phase II of SISS survey in the United States. The prediction function is approximated by a truncated Taylor series. In Chapter 2, the related literature is reviewed to identify the research position for the present study.

CHAPTER 2

REVIEW OF THE RELATED LITERATURE

The literature reviewed in this chapter is presented in three parts: (1) prediction models; (2) the SISS and SES surveys; and (3) shrinkage regression. The chapter concludes with a description of the context of this doctoral research and its originality.

PREDICTION MODELS

Several researchers have reported nonlinear relations between measures of achievement and direct and/or proxy measures of classroom practices (e.g., Brophy & Evertson, 1974; Loucks, 1975; Rim & Coller, 1978; Soar, 1966, 1968, 1971, 1973; Soar & Soar, 1972). Flanders (1970) indicates that "the main credit for identifying and conceptualizing nonlinear (or curvilinear) relationships belongs to Soar" (p. 403). Soar and Soar (1976) claim:

> Although linear relationships have most often been used in studies of teaching effectiveness to identify relationships between classroom behavior and pupil gain, it seems clear that they are limited in the extent to which they can help us answer the question of what good teaching is. They are simplistic in implying that if some of a behavior is good, more is better and once the question is raised, it becomes difficult to imagine very many behaviors for which increasing amounts would be unqualifiedly good. (p. 265)
>
> Nonlinear relations were found between measures of achievement gain and measures of teacher behavior which appeared to represent teacher limitation of pupil freedom in the development of subject matter and thought. (p.263)

Walberg (1981) emphasizes the importance of interaction effects on students' achievement. He writes:

> The usual form of the regression equation for estimating learning is L = a + b A + c E, where a is a constant and b and c are coefficients or regression weights for the aptitude and environment terms.
>
> A possible problem of the equation itself, however, is that it does not allow for the possible "interaction" of aptitude and environment (Lindquist 1951, Cronbach 1957). For example, a student with high visual aptitude may benefit more than others from instruction making use of pictures. Such interactions may be tested by adding an arithmetic product term (not to be confused with educational or economic product) AE to the regression L = a + bA + cE + d AE. (p.84)

Jagodzinski, Weede and Tiefenbach (1981) maintain: "Polynomial regression is the standard procedure for testing curvilinear propositions, particularly if nonmonotonic relationships are being investigated" (p. 447). Cramer and Appelbaum (1978) have shown that a polynomial regression model can be applied to the random case as well as to the fixed case. They point out that the two models differ in terms of the standard errors of the estimates and power functions, but are identical in their estimation and hypothesis testing aspects.

Sockloff (1976) states:

> Recent works by Cohen (1968), Kelly, Beggs, McNeil, Eichelberger, and Lyon (1969), Kerlinger and Pedhazur (1973), McNeil (1970), Walberg (1971), and Bottenberg and Ward (1963) have attested to the flexibility of the General Linear Model. ··· The above publications, plus those of Digman (1966) and of McNeil and Spaner (1971), have shown the capabilities of the General Linear Model in handling the analysis of nonlinear data. This approach, with a history dating back to Court (1930), received modern impetus in education and psychology from Saunders' (1956) work on moderated regression, and much of its recent popularity can be attributed to the effects of Cohen's (1968) proselytic paper. (p.267-268)

Nevertheless, Dewit, Wister and Burch (1988) claim: "Higher-order polynomials are extremely rare using data from the social science" (p. 65). Bradley and Srivastava (1979) and Budescu (1980) have discussed multicolinearity of predictors in a polynomial model. They show that the degree of colinearity in a normal system rapidly increases as a direct function of the polynomial degree. Jagodzinski, Weede and Tiefenbach (1981) argue: "Even in second-order polynomial regression there are some problems; often there is extreme multicolinearity between simple and squared terms." (p. 447)

Cohen and Cohen (1983) suggest centering data as a partial solution to the problem of multicolinearity. However, Budescu (1980) contend that centering does not entirely solve the problem of multicolinearity when distribution is skewed. Liu (1981) writes:

> It is frequently suggested that centering variables prior to forming higher power regressors is essential. Although this may reduce the computation problem involved in calculating the ordinary least squares estimates, centering does nothing to reduce the effects of multicolinearity. When the centered model is expanded, the usual problems of multicolinearity still influence the individual estimates. (p. i)

Besides the problem of multicolinearity, polynomial regression is limited by its inability to provide clearly interpretable results in a curvilinear analysis (Curry, Roberts & Walling, 1986). Stimson, Carmines and Zeller (1978) write: "While polynomial regression is statistically sound, it produces awkward interpretational equations which use a series of linear slopes to describe a curve" (p. 515). Cohen and Cohen(1983) perceive interpretation and multicolinearity as the major difficulties of polynomial regression.

In summary, most previous research has treated polynomial regression as an alternative to a linear model. The approach adopted in this research is to view both linear and nonlinear models as special cases of a Taylor polynomial series. Possible solutions to multicolinearity are clarified in Chapter 3, and the problem of interpretation is discussed in Chapter 5.

THE SISS AND SES SURVEYS

A major feature of the Second IEA Science Study (SISS) is the international dimension which makes it possible for countries to learn from each other and to develop better science programs for their children and young people (IEA, 1988). IEA published its SISS reports in three volumes (Rosier & Keeves, 1991; Postlethwaite & Wiley, 1992; Keeves, 1992) on three topics: (1) science education and curricula in twenty-three countries; (2) science achievement in twenty-three countries; (3) changes in science education and achievement. The major results from the U.S. participation in SISS are summarized in Science Achievement in the United States and Sixteen Countries (Jacobson & Doran, 1988, p. 3). Jacobson & Doran (1991) write:

> Of special interest are findings related to such issues as the following:
> How did the science achievement of U.S. students compare with the science achievement of students in other countries?
> How did the science achievement of advanced science students who had studied a science for two or more years compare with advanced science students in selected other countries?
> To what extent was there growth in science achievement from Grade 5 through Grade 12?
> How did science achievement in the 1980s compare with science achievement in the 1970s?

How did the science achievement of girls compare with that of boys?
What factors in home, school, and community were associated with science achievement?
What approaches to teaching and learning were associated with science achievement? (p. 2-3)

The report is based primarily on ten dissertations (Chandavarkar, 1988; Chang, 1988; Clive, 1983; Dryden, 1987b; Ekeocha, 1986; Ferko, 1989; Humrich, 1988; Kanis, 1988; Micik, 1986; Miller, 1985) and an IEA preliminary report (IEA, 1988). Since then, at least three additional doctoral dissertations (Baker, 1989; Bayer,1990; O'Rafferty, 1991) have been conducted on the U.S. SISS project, and a number of monographs (Anderson, 1990; Ferko, Doran, & Jacobson, 1991; Helgeson, 1988; Humrich, 1991; Jacobson & Doran, 1988; 1991; Jacobson, Takemura, Doran, Kojima, Humrich & Miyake, 1986; Kanis, Doran & Jacobson, 1990; Miller, 1986;1988; Schneider, Muller, Doran & Jacobson, 1991; 1990) have been developed by SISS research associates. The American SISS team has also given many presentations at national or international conferences (Bell, 1989; Chang, 1990; 1989; Doran, 1992; Doran, et. al., 1991; Dryden, 1987a; Humrich, 1989a; 1989b; 1988; O'Rafferty, et. al., 1991; Jacobson, et. al., 1987; 1986; Tamir, et. al., 1989) and has published many papers in well-known research journals (Chandavarkar, Doran, & Jacobson, 1991; Crew, 1988; Doran, 1990; Doran & Jacobson,1984; Eichinger, 1990; Ferko, Doran & Jacobson, 1990; Humrich, Jacobson & Doran, 1990; Jacobson, 1990; 1988; Jacobson & Doran, 1985a, b; 1986; Jacobson, Doran & Schneider, 1991; Kanis, 1990; 1991; Krieger, 1984).

On the other hand, the SISS Extension Study (SES) in China was conducted by the China IEA Center. The author was a member of the Chinese IEA team and participated in the first two stages of the SES survey, population investigation and data collection. Based on the author's personal communication with the present staff at the IEA center, the SES data sets have not been released to the public yet, and therefore, no doctoral research has been conducted on the SES project. Apparently, the present study is the first report in which the SES data base is used in a comparative study between the United States and China. To facilitate the international comparison, the American SISS literature at the ninth grade level is reviewed in the rest of this section; the review focuses on the prediction of U.S. students' science achievement.

Ekeocha (1986) has studied students' correlates of science achievement based on the fifth grade SISS data set of the United States. According to the literature reviewed in his dissertation, three major constructs, the home, classroom experience, and student attitudes, potentially influence student science achievement. Ekeocha (1986) utilized these constructs to build a "general path analytic model" which assumes a linear structure of the construct relations for prediction of students' science achievement. He reports:

> The results from the causal models using the individual student as the unit of analysis indicate that the home and the student attitudes have a positive direct significant effect on science achievement. The effect of classroom construct on achievement was greater through mediation than by direct path. (p.1)

Usually, a construct has several indicators. Dryden (1987a) found that father's education and the number of books in the home are the best indicators of student home background variables. Ekeocha (1986) states: "A separate analysis of the home background component variables indicated that 'possession of books in the home' had a significant effect on science achievement" (p.1). Jacobson, Doran, Humrich and Kanis (1988) found the same result for the ninth grade population. According to Jacobson, Doran and Schneider (1992), "Books in the Home" may be a general surrogate for the overall level of culture in the home (p.393). Jacobson and Doran (1988) clarified the investigation of the home influence in SISS:

> The students reported on the amount of their parents' education, the nature of their parents' work, and the number of books in their homes. These variables have been viewed as indicators of family socio-economic status. ··· The parents of the advanced science students were more likely to have had a higher level of education than the parents of fifth and ninth grade students. (p. 86)

A second construct, classroom experience is reflected in SISS by students' responses to a science learning questionnaire that asked how they had studied science (Jacobson & Doran, 1988; p. 99). The questions were grouped into three categories, general teaching techniques, specific science procedures and approaches to homework. Jacobson and Doran summarize results of the responses as follows:

> A wide variety of teaching techniques were reported to have been used. Included were considerable laboratory work and teacher demonstrations. Students reported that their work in the classroom and the laboratory was mostly teacher directed. Tests were widely used at all grade levels. (p. 102)

In Dryden's dissertation (Dryden, 1987b), the effect of classroom experience on student achievement is further explored at the ninth grade level. Dryden (1987b) writes:

> It seems as if teacher and student variables are independent of each other, except for the process environment. This implies that teacher has no direct effect on science achievement or attitude toward science. Instead, the effect is mediated through the classroom process variables. All the school support, effort and education of the teacher means very little if the teacher cannot organize the classroom structure in a meaningful manner. The key is not the teacher, per se, but the teacher's ability to organize meaningful learning experiences for the student. (p. xii)

It should be noted that Dryden's research data were collected from Phase I of the SISS survey, and the models in that dissertation were constructed with Partial Least Square (PLS) and LISREL methods. According to statistical decision theory, the mean square error of prediction (MSEP) of these methods is larger than that of shrinkage regression (Casella and Berger, 1990), and hence, the prediction is less accurate.

A third construct, student attitude, is measured by a SISS instrument which elicits students' reactions to science and school (Jacobson & Doran, 1988; p. 109). The questions were chosen to represent several perspectives: science as a school subject, school, and the contribution of science to the country. The American results are reported by Jacobson and Doran (1988):

> U.S. students had positive attitudes toward science and school. Most students found school to be challenging, but some reported that school was not enjoyable. The students generally found studying science to be enjoyable and interesting. They indicated that studying science was not difficult when it involved handling apparatus, but that there were many facts to learn in science. (p. 115)

Gender is another factor which affects students' science achievement (Dryden, 1987; p. xii). Humrich (1988) has investigated gender effects based on data from Phase II of the SISS survey. She claims (1988):

> The major findings can be summarized as follows: Sex differences were found at every grade level and in every subject area in the written science achievement tests. This sex difference always favored males. Overall sex differences remained fairly constant, in the 5%-7% range, for all populations surveyed, with exception of biology 1 (3%) and non-science population (3.7%). (p. 1)

However, Micik conducted a case study at a New England junior high school using SISS instruments. He reports that, at the ninth grade level, the gender difference in science achievement (0.9%) is much less than the difference at the national level (Micik, 1986). This example seems to suggest that researchers should exercise caution when inferring gender differences at a specific school based on the national results.

Indeed, school equity is a potential assumption when applying national SISS results to a particular school. Gender differences, for example, may depend on whether a school is co-educational. Only if the school effects are negligible can national results be meaningful to a school. Otherwise, sampling errors and significance testing must be included in consideration. The technical difficulties are explained by Keeves (1992) in one of the IEA international reports:

> The highly stratified sample design with complex cluster sampling and differential losses at the student, school and strata levels make the task of calculating sampling errors and significance testing a very complex one. There are no simple procedures for the estimation of sampling errors, using variance ratios or formulae that are accurate or appropriate with such samples. This applies to all classes of statistics, whether means, correlation coefficients or regression coefficients. The only procedures that are considered appropriate are "jack-knifing" and "bootstrapping", and even here there is some controversy as to whether the latter is meaningful with large and complex samples. (p. 53)

In summary, the research reported herein is the first doctoral dissertation which utilizes the SES data base from China. Although many researchers have explored factors which affect students' science achievement in the United States, no prediction model has been developed empirically based on the Phase II of SISS survey.

SHRINKAGE REGRESSION

In the 1930s, Jerzy Neyman, Egon S. Pearson, and Abraham Wald undertook a mathematically more rigorous approach to statistical inference. The ideas they developed are part of what is now known as statistical decision theory. They discarded the requirement of unbiased estimation and examined all functions of the data that could serve as estimators of the unknown mean μ. These estimators are compared through a risk function, defined as the expected value of the squared error for every possible value of μ (Efron and Morris, 1977).

A criticism of the risk function focuses on its nonsymmetric penalty, i.e., underestimation has only a finite penalty while overestimation has an infinite penalty (Casella and Berger, 1990). Smith (1990) suggests that more than one criterion, such as variance, bias, or mean square error of prediction (MSEP), should be included in the risk function construction. However, other researchers (e.g., Bilodeau & Srivastava, 1988; James & Stein, 1961; Lemmer, 1988; Weigel, et. al., 1991) still believe that a sensible way of assessing the efficiency of an estimator is to calculate its mean square error. A summary in the Encyclopedia of Statistical Science (Kotz & Johnson, 1988) concludes: "It is generally accepted that minimum MSE is a highly desirable property, and it is therefore used as a criterion to compare different estimators with each other".

For a prediction model $y=X\beta+\varepsilon$, the vector y and the matrix X known, β a vector of unknown parameters, and ε the random error vector, the Gauss-Markov theorem assures that we cannot find linear unbiased estimators of the regression coefficients β which have smaller variances than the least square estimators $(X'X)^{-1}X'y$. Apart from the minimum variance property in the class of linear unbiased estimators, the least square estimators can be highly variable in certain dimensions (Liski, 1982; Whittemore, 1989; Stanley, 1990). Thus, biased alternatives to the ordinary least squares are recommended in order to obtain a substantial reduction in variance. According to statistical decision theory, an estimator will be judged good if it has a small, but probably nonzero, bias combined with a small variance (Casella and Berger, 1990). Hoerl and Kennnard (1970) have shown that such an estimator always exists.

One alternative to the least square estimator is the maximum likelihood estimator. However, under the condition of normality, the least square estimator is the same as the maximum likelihood estimator. James and Stein (1961) have developed a non-linear estimator with smaller mean square error of prediction (MSEP) than that of the maximum likelihood estimator throughout the parameter space when more than two uniquely estimable fixed effects are estimated in a normal linear model. Based on decision theory, maximum likelihood is inadmissible under the square error loss (Weigel, et. al., 1991). The inadmissibility means that at least another estimator exists that gives estimates with MSEP smaller than or equal to the MSEP of maximum likelihood method throughout the parameter space (Casella & Berger, 1990). Interestingly, the James-Stein estimator is not itself admissible (Draper & Norstrand, 1979). A detailed discussion on its inadmissibility is given by Sennetti and Kakar (1980). Efron and Morris (1977) have commented on Stein's ideas:

> *Sometimes a mathematical result is strikingly contrary to generally held belief even though an obviously valid proof is given. Charles Stein of Stanford University discovered such a paradox in statistics in 1955. His result undermined a century and a half of work on estimation theory, going back to Karl Friedrich Gauss and Adrien Marie Legendre. After a long period of resistance to Stein's ideas, punctuated by frequent and sometimes angry debate, the sense of the paradox has diminished and Stein's ideas are being incorporated into applied and theoretical statistics.*

The term "shrinkage" apparently originates with J. R. Thompson (1968) in connection with estimators that have been modified (Lemmer, 1988). According to Ralph (1976) and Matloff (1982), two popular shrinkage approaches to estimating regression coefficients are the ridge estimators of Hoerl and Kennard (1970) and the Stein-type estimators derived from the estimation methods given in the original papers by Stein (1956) and by James and Stein (1961). Ridge estimators are designed as a method to improve on the unsatisfactory characteristics of the least squares estimator when there are multicolinearities among the predictor variables. Stein-type estimators are frequently recommended because they reduce MSEP and they can be regarded as empirical Bayes estimators as in Efron and Morris (1973). However, as Jennrich and Oman (1986) point out, "It thus comes as a surprise that Stein estimation is not routinely used in regression applications -- we know of no statistical packages with Stein regression routines, and when shrinkage estimators are used in multiple regression models, they are more often ridge-type estimates" (p. 113).

Research in Stein estimation began in the mid-sixties. Its growth became phenomenal in the mid-seventies, and the topic is currently one of the most popular and active areas of research. Among the various Stein-type estimators, the estimator suggested by Copas (1983) has a good reputation with statisticians. The following comments about the Copas method are quoted from the Journal of Royal Statistics Society (Ser. B, 1983; pp. 335-348).

> *Using an ingeniously new and intuitively appealing notion of prediction shrinkage, Professor Copas has been able to give new life to the shrinkage estimator originally proposed by Charles Stein* (P. J. Brown, Imperial College, London).
> *This paper, I believe, will be seen as a very important one in that it ties together very neatly many of the ideas which have been tossed around over the past few years of prediction and shrinkage in the regression model* (I. R. Dunsmore, University of Sheffield).
> *Professor Copas is to be congratulated on an important and stimulating contribution to regression theory. His insights into the relationship between retrospective fit and prospective fit and subset selection are most illuminating.* (Professor J. A. Anderson, University of Newcastle-upon-Tyne).
> *I have read this paper with much admiration; it addresses the problem of regression prediction under the assumption that future data will be "rather similar" to past data; this will often be necessary and sensible when the aim is prediction* (A. J. Lawrance, Birmingham University).
> *This is a very fine paper, and an interesting pointer to the direction of future developments in regression analysis for both continuous and discrete data* (R. L. Plackett, University of Newcastle-upon-Tyne).

Research conducted by Hebel, et. al. (1993) has shown that the Copas shrinkage method is better than least square method in terms of minimizing mean square of prediction. Wang (1993) developed a program to compute the shrinkage estimator based on LS regression. The program is used in this research to construct a model for prediction of students' science achievement.

SUMMARY

The construction of a prediction model for student achievement is a fundamental research issue in education. Most previous research divides prediction models into linear vs. non-linear categories. However, as an empirical exploration, neither linear nor non-linear relations should be imposed as a pre-condition of the model construction. Instead, it is preferable to keep both linear and non-linear functions as possible options and to identify the optimal model by empirical data sets.

Many SISS researchers have investigated the effects of home, gender, classroom experience and student attitude on student science achievement. But, no comparative study has been conducted in the United States and China through the use of Phase II SISS and SES data sets. Neither has the shrinkage estimator been mentioned in the American SISS literature.

In this research, both linear and non-linear functions are treated as special cases of a Taylor polynomial series. The shrinkage method favored by Copas (1983) and Hebel, et. al. (1993) is employed to construct the polynomial coefficients in the truncated Taylor model. Compared to the methods of least squares and maximum likelihood, the shrinkage estimator has a smaller MSEP, and hence, provides better approximation to the empirical prediction function supported by the Phase II SISS and SES data sets. A detailed explanation of the construction of the model is presented in Chapter 3.

CHAPTER 3

METHODOLOGY

This research is designed to identify significant factors, linear or nonlinear, for prediction of students' science achievement. Construction of the model is based on students' information in five areas, gender, attitude, home background, classroom experience, and personal effort. A shrinkage method is employed to estimate regression coefficients in a truncated Taylor polynomial model. In this chapter, methodology for this study is delineated in four parts: (1) overview; (2) latent predictor construction; (3) significant factor selection; and (4) shrinkage estimation.

OVERVIEW

Predictors of students' achievement can be measured directly, such as gender, or indirectly, such as students' attitude, home background, classroom experience and personal effort. The predictors which are not directly measurable may be interpreted by students' responses to certain related questions. The responses are called indicators, and the predictors which are not directly measurable are called latent predictors. Ekeocha (1986), Dryden (1987a) and Keeves (1992) assert that both gender and latent predictors affect students' science achievement. The relations between the predictors and students' achievement is expressed in this research as a Taylor polynomial series which includes both linear and nonlinear functions as special cases.

It should be noted that a polynomial function, $y = \beta_0 + \Sigma \beta_i x_i^j$, is still linear in terms of the β_0 and β_i parameters. In other words, polynomial models belong to the statistical General Linear Model (GLM) family (Graybill, 1976; p.302). According to Sockloff (1976), "The General Linear Model is a name given to the family of models possessing a common characteristic, namely, linearity in the parameters of the equation specifying the model" (p.268). In the last two sections of this chapter, General Linear Model is applied to significant factor selection and parameter estimation.

Nevertheless, Graybill (1976) points out: "A problem that sometimes arises when a polynomial model is under consideration is that of determining the degree of the polynomial" (p. 303). Jagodzinski, Weede & Tiefenbach (1981) complain: "Even in second-order polynomial regression there are some problems; often there is extreme multicolinearity between simple and squared terms" (p. 447).

When multicolinearity exists, the observation matrix, $\{x_i\}$, is close to singular. In this case, a "very small" change in one or more observations produces a "significantly large" change in the estimation of β_0 and β_i parameters (Graybill, 1976; p. 230). Liu (1981) studied multicolinearity in her dissertation. She suggests principal component regression as a means of ameliorating the adverse effect of linear dependencies in a polynomial regression model (p.i).

Principal component analysis was originated by Pearson (1901) and later developed by Hotelling (1933). The dimension of latent prediction space can be explored by scree plot through a principal component analysis routine (PRINCOMP) in SAS (Johnson, 1992). By default, SAS treats principal components which have eigen values greater than 1 as information, and the remaining components as noise. Because principal components are orthogonal to one another, the number of principal components equals the number of dimensions in the latent space.

A disadvantage of the default option is that eigen values of some principal components may be so close to 1 that it is not appropriate to set the threshold among them. In a scree plot, eigenvalues are plotted for each principal component. Hence, one may select a clear-cut threshold to identify dimensions of the latent prediction space.

LATENT PREDICTOR CONSTRUCTION

The common instruments employed to collect information for prediction of students' science achievement in the SES and Phase II SISS surveys are listed in Table 2.

Table 2: Common Instruments of SES and Phase II SISS

Variable	Instrument
SEX	What is your sex? (A) male; (B) female.
FPOSTED	What is the highest level of school your father completed?
MPOSTED	What is the highest level of school your mother completed?
HOMEBOOK	How many books are there in your home?
HMWKALL	About how many hours a week do you usually spend on homework or other school work out of class for all subjects?
HMWKSCI	About how many hours a week do you usually spend on homework or other school work out of class for science subjects?
P_ATT05	Science is very important for a country's development. Agree; Disagree; Uncertain.
P_ATT06	School is not very enjoyable. Agree; Disagree; Uncertain.
P_ATT34	Science is an enjoyable school subject. Agree; Disagree; Uncertain.
P_ATT35	The science taught at school is interesting. Agree; Disagree; Uncertain.
P_ATT36	Science is a difficult subject. Agree; Disagree; Uncertain.
P2DES01	We use a textbook for our science lessons Often; Sometimes; Never.
P2DES02	We use books other than textbook for learning science. Often; Sometimes; Never.
P2DES08	We watch the teacher do experiments during our science lessons. Often; Sometimes; Never.
P2DES14	We have tests on what we learned in science. Often; Sometimes; Never.
P2DES18	We do experiments as part of the science lessons Often; Sometimes; Never.

Four adjustments are made on the SES and Phase II SISS data bases. First, the attitude scales (ATT05-36) are recoded as: agree = 1; uncertain = 0; and disagree = -1. Second, parents' education (P_ED) is defined as: P_ED = max (FPOSTED, MPOSTED). Third, the SEX predictor is recoded as: female = 0 and male = 1. Fourth, missing values are deleted.

The first two adjustments are consistent with the definition of terminology in chapter 1. The third adjustment, a recode of SEX, has two advantages: (1) SEX can be used as a dummy variable for regression; and (2) the polynomial model is simplified because $(SEX)^n=(SEX)$ for any integer n.

Table 3: Sample Sizes of SES and Phase II SISS Data Sets

Sample	Country	
	The United States	P. R. China
Designed Size	2519	3000
Achieved Size	2027	2871
Missing Value (%)	19	4

The fourth adjustment, deletion of missing values, is summarized in Table 3. The designed sample sizes are proportional to students' populations. The achieved samples are created by deleting cases which have missing values for the variables in Table 2. In both the U.S. and Chinese data sets, the percentage of missing values is less than 20%, and the achieved samples are larger than 2000. Thus, as an exploratory study, missing values are deleted from the SES and Phase II SISS data sets.

A possible method to construct latent predictors is to conduct a principal component analysis on indicators in Table 2 and then define the first four principal components as latent predictors of students' attitudes, home background, classroom experience and personal efforts respectively. Unfortunately, the latent predictors are difficult to interpret since they depend on information in several areas. To facilitate predictor interpretation, indicators are classified into meaningful groups and principal component analysis is applied to each group. Because the largest proportion of information in a group is accounted for by its first principal component, the latent predictor is represented by the first principal component for each group and the total number of predictors is chosen to be equal to the number of latent prediction dimensions identified by scree plots to avoid the problem of multicolinearity.

SIGNIFICANT FACTOR SELECTION

Prediction models explored in this research are mathematically expressed as polynomial functions with students' achievement as the dependent variable and all possible polynomials and interactions of predictors as independent factors.

Based on convergence property of a power series (Ayres, 1964), the first (k+1)th partial sum $S_{k+1}(x) = \Sigma c_j x^j$ with j = 0, 1, ..., k forms a kernel of the series and the remainder after (k+1) terms $R_{k+1}(x) = \Sigma c_i x^i$ with i = k+1, k + 2, ... is close to zero when k gets large. In terms of statistics, only the first (k+1) items have possibilities of significance. Thus, a truncated polynomial model provides a good approximation of the potential function which dominates the achievement prediction.

Graybill (1976) states: "We assume that the degree of the polynomial $\mu(x)$ is less than or equal to K, and the problem is to determine the exact degree"(p. 303). Based on the common variables of SES and Phase II SISS, factors of students' science achievement are the polynomials of variable SEX and the latent predictors and their interactions. In general, higher order polynomials and interactions are included in an exploratory model until the degree of (K + 1) is reached at which the factors are no longer significant. Then, K is the highest degree of the polynomials in the prediction model.

It should be mentioned that not all factors which have degree less than or equal to K are significant. For example, it was shown in chapter 1 that polynomial terms with even exponents are not significant for potential functions such as sin(x), arcsin(x) and arctan(x). Since the potential function for predicting students' achievement is unknown, it is desirable to select significant factors from all possible polynomials and interactions which have degree less than or equal to K.

Graybill (1976) points out that an appropriate routine for factor selection which takes all possible factors into consideration is the backward elimination procedure. The backward routine is available in Statistical Analysis System (SAS, 1982), and is adopted to select polynomial and interaction items for prediction of students' science achievement.

SHRINKAGE ESTIMATION

Copas shrinkage estimators have two properties: (1) the mean square error of prediction is smaller than the least square estimator (Hebel, 1989); and (2) the regression coefficients can be constructed based on least square estimators (Wang, 1993). The first property was reconfirmed by Wang (1993) using a forty year data base of Kansas wheat yield. This section is devoted to the explanation of the second property, the computation of shrinkage coefficients.

Let **b** be a vector of the least square estimator, and let **b**(c) be a vector of the shrinkage estimator, and R^2 be the coefficient of multiple determination for the least square model. Then for a general linear model $\mathbf{y}=X\cdot\mathbf{b}(c)+\mathbf{e}$, (**y** and X are known $n\cdot1$ and $n\cdot p$ matrices respectively, n is the number of observations and p is the number of parameters, **e** is distributed as normal with mean zero and covariance matrix $\sigma^2 I$), the two coefficients employed by Hebel, et. al. (1993) are:

$$c_1 = \frac{p-2}{n-p} \qquad (1)$$

$$c_2 = \frac{p-2}{n-p+2} \qquad (2)$$

where $p > 2$.

Either c_1 or c_2 can be used as c in the following formula:

$$k = 1 - c \frac{1 - R^2}{R^2} \qquad (3)$$

where $R^2 = \frac{\text{sum of squares due to LS regression}}{\text{total sum of squares}}$.

The shrinkage coefficient is:

$$\mathbf{b}(c) = K(c)\,\mathbf{b} \qquad (4)$$

where $K(c) = \text{Max}(\min.(k, 1), 0)$.

Two points should be noted for the shrinkage regression. First, since $0 < K(c) < 1$, the shrinkage coefficient K(c) shrinks the **b**(c) towards zero. Secondly, there are two values of **b**(c) corresponding to Copas c_1 and c_2 coefficients.

Hebel (1989) showed by computer simulation that c_1 and c_2 are equally good in parameter estimation. Wang (1993) compared Copas c_1 and c_2 coefficients in a real data analysis, and concluded that the mean square error of prediction is smaller for c_1. Accordingly, the Copas c_1 method is employed in this research to estimate regression coefficients for the model of prediction.

SUMMARY

Students' attitudes, home background, classroom experience, and personal effort are latent predictors of students' science achievement. The dimension of the latent prediction space is identified in this research by scree plots through principal component analysis. The common variables observed in the SES and Phase II SISS projects are classified into meaningful groups, and the first principal component is computed for each group. The factors of prediction are constructed by polynomials of the visible variable SEX, the latent principal components and their interactions. Significant factors are selected through the backward elimination procedure in SAS. Shrinkage regression is applied to estimate regression coefficients for the model of prediction. The results of the model construction are presented in Chapter 4.

CHAPTER 4

PRESENTATION OF RESULTS

The results are presented in four parts: (1) dimension of the latent prediction space; (2) structure of latent predictors; (3) significant factors in the prediction, and (4) shrinkage estimates of the causal effects on students' achievement.

DIMENSION OF LATENT PREDICTION SPACE

Among the variables in Table 2, only SEX is a visible predictor of students' achievement. The remaining variables are treated as indicators of several latent predictors. The dimension of the latent prediction space in the Phase II SISS and SES was determined by scree plots, which are presented in Figures 1 and 2 respectively. A scree plot is a plot of eigenvalues against the principal components constructed by the indicators in Table 2. Inspection of Figures 1 and 2 shows that the fourth, fifth and sixth principal components have eigenvalues around 1, and the difference of eigenvalues between the fourth and fifth principal components is larger than the difference between the fifth and sixth principal components. Hence, the dimension of the latent prediction space is four.

STRUCTURE OF LATENT PREDICTORS

Corresponding to the four dimensions of the latent prediction space, the indicators in Table 2 are grouped into four categories: (1) attitudes (P_ATT05-P_ATT36); (2) classroom experience (P2DES02-P2DES18); (3) home background (P_ED, HOMEBOOK); and (4) personal effort (HMWKALL, HMWKSCI). For each group, the first principal component is computed to represent the latent predictor of the dimension. The structure of latent predictors is expressed by the factor loadings of the first principal component. The factor loadings constructed by the U.S. and Chinese data are presented in Tables 4 and 5. It should be noted that the identical factor loadings for home background and personal effort factors in both the U.S. and China do not have meaningful interpretation. In a two-indicator case, the axis of a principal component is set at a direction of 45^{0} to each indicator axis. Thus, the factor loading of each

Figure 1: Scree Plot of Eigenvalues for U.S. Data

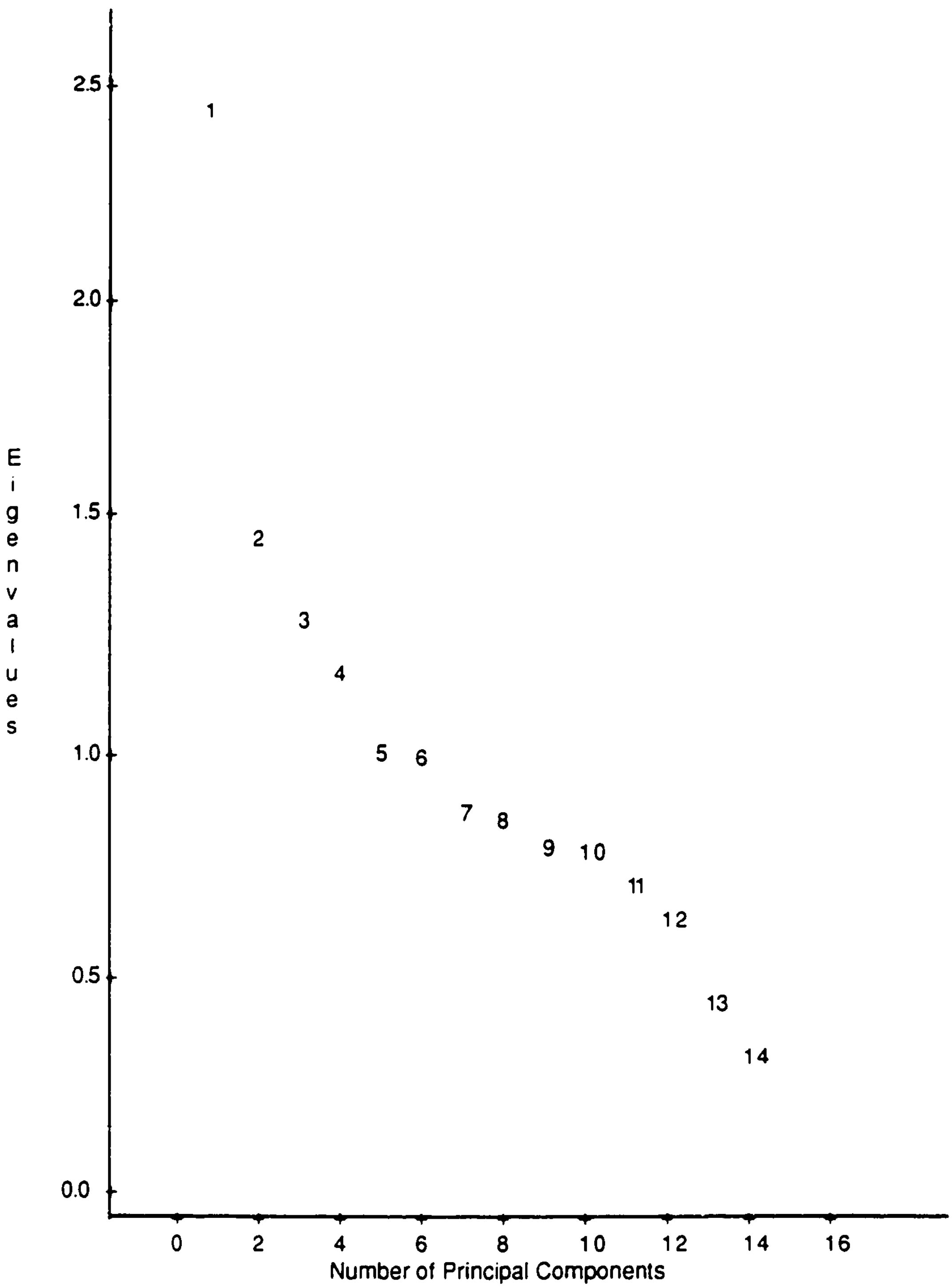


Figure 2: Scree Plot of Eigenvalues for China Data

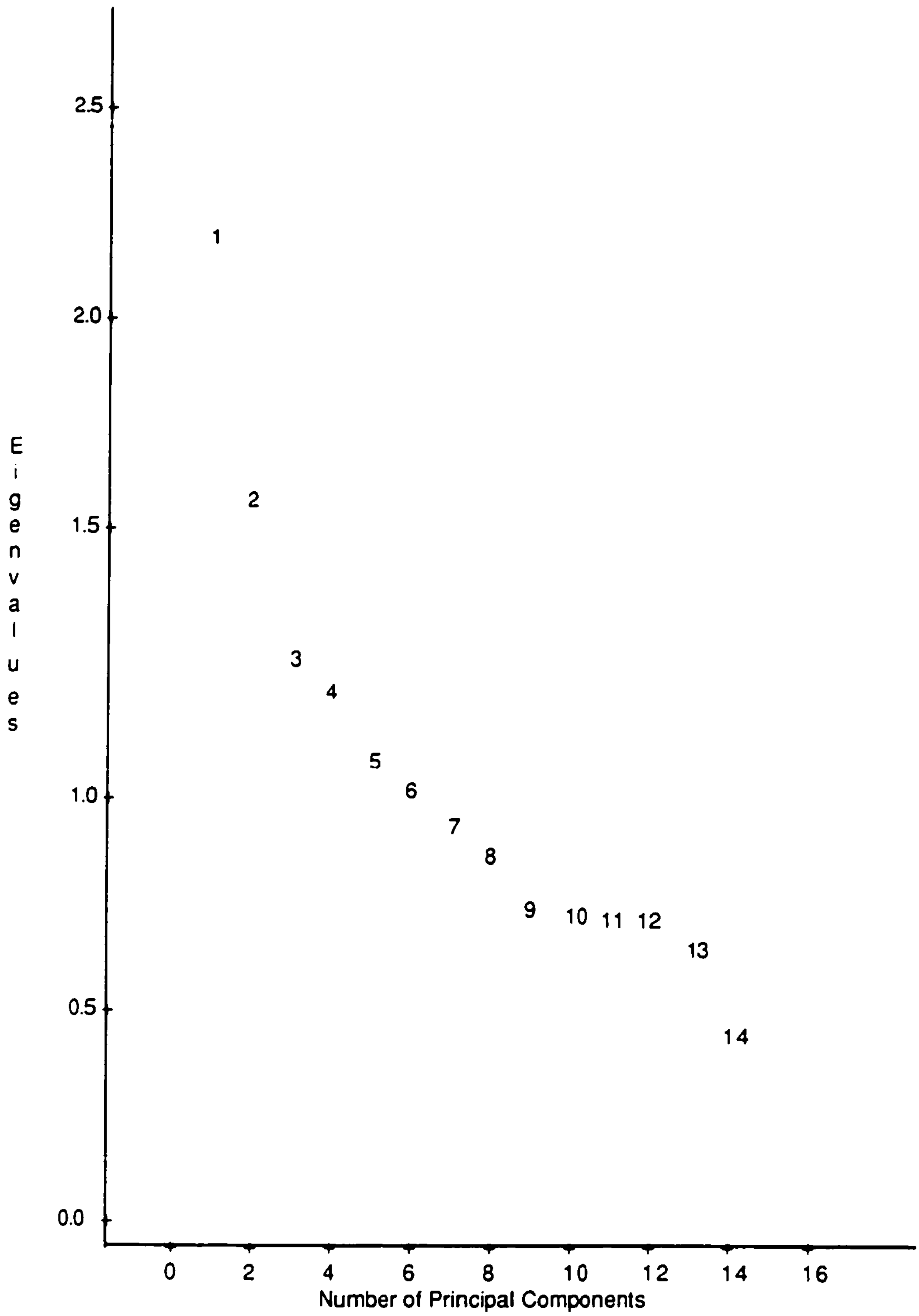


Table 4: Structure of the Latent Predictors in the U.S. Data

Predictor	Indicator	: Factor Loading
Attitude	P_ATT05	: 0.356644
	P_ATT06	: -0.298548
	P_ATT34	: 0.604701
	P_ATT35	: 0.573373
	P_ATT36	: -0.298753
Classroom Experience	P2DES01	: 0.074826
	P2DES02	: 0.237567
	P2DES08	: 0.653609
	P2DES14	: 0.229283
	P2DES18	: 0.653609
Home Background	P_ED	: 0.707107
	HOMEBOOK	: 0.707107
Personal Effort	HMWKALL	: 0.707107
	HMWKSCI	: 0.707107

Table 5: Structure of the Latent Predictors in the Chinese Data

Predictor	Indicator	: Factor Loading
Attitude	P_ATT05	: 0.275895
	P_ATT06	: -0.506237
	P_ATT34	: 0.562202
	P_ATT35	: 0.559620
	P_ATT36	: -0.195859
Classroom Experience	P2DES01	: 0.387042
	P2DES02	: 0.273908
	P2DES08	: 0.500908
	P2DES14	: 0.534701
	P2DES18	: 0.488220
Home Background	P_ED	: 0.707107
	HOMEBOOK	: 0.707107
Personal Effort	HMWKALL	: 0.707107
	HMWKSCI	: 0.707107

indicator is: Sin(45^o)=Cos(45^o)=0.707107, i.e., two indicators are equally weighted in construction of the factor. The factor loadings for the other two factors, attitude and classroom experience, reflect the relative contributions of each indicator to the latent factors. A detailed discussion of the two factor structures is given in Chapter 5.

SIGNIFICANT FACTORS IN THE PREDICTION

Significant factors in the prediction are selected from interactions and polynomials of variable SEX and the principal components through the SAS backward elimination procedure. The empirical model based on American data contains significant factors up to and including the fifth degree of polynomial. On the other hand, the empirical model based on Chinese data contains no significant factors beyond the fourth degree. Hence, the degree of prediction for the

Table 6: Significant Predictors of U.S. Students' Science Achievement

Variable*	Parameter Estimate (β_U)	P-Value
Intercept (β_0)	56.18340139	0.0001
S	6.02599883	0.0001
H	1.94039381	0.0001
E	2.86139405	0.0001
A	-2.27856193	0.0001
H•E	1.26561305	0.0019
H•H	-1.90926090	0.0001
C•C	-1.01071292	0.0001
S•A•A	-0.75133092	0.0073
S•C•C	0.92984590	0.0082
S•H•A	-1.12570535	0.0021
S•H•E	-1.44883757	0.0073
H•H•A	0.66445544	0.0079
A•A•C	-0.19754143	0.0416
E•C•C	-0.54054931	0.0131
E•E•E•C	-0.31713476	0.0227
E•E•E•E	-0.12420098	0.0012
E•E•E	-0.08416007	0.0003
S•H•H•A	-0.81375534	0.0108
E•E•E•E•E	0.04034700	0.0094
S•H•H•A•A	0.45744812	0.0034

* Note: S = Sex; C = Classroom Experience; E = Effort; H = Home Background; A = Attitude.

Chinese model is 4, and for the U.S. model is 5. The significant predictors are presented in Tables 6 and 7, respectively.

Table 7: Significant Predictors of Chinese Students' Science Achievement

Variable*	Parameter Estimate (β_C)	P-Value
Intercept (β_0)	52.70693632	0.0001
S	8.24791340	0.0001
E	1.96952871	0.0001
C	-0.66987121	0.0069
H•H	0.42823485	0.0040
E•E	-0.41561330	0.0032
S•H•C	0.45412394	0.0409
H•A•E	-0.25582318	0.0393
H•E•C	0.25386140	0.0495
A•C•C	0.14284998	0.0198
S•A•A•A	-0.05587136	0.0010
S•C•C•C	-0.12735585	0.0017

* Note: S = Sex; C = Classroom Experience; E = Effort; H = Home Background; A = Attitude.

SHRINKAGE ESTIMATES OF β's

Parameter β's are the regression coefficients of the significant predictors selected by the SAS backward procedure. The parameters estimated in Tables 6 and 7 are least squares estimates [**b**]. The calculation of shrinkage estimates [**b**(c)] using least squares estimates is illustrated in Table 8.

Table 8: Calculation of Shrinkage Estimates Based on Copas C_1 Coefficient

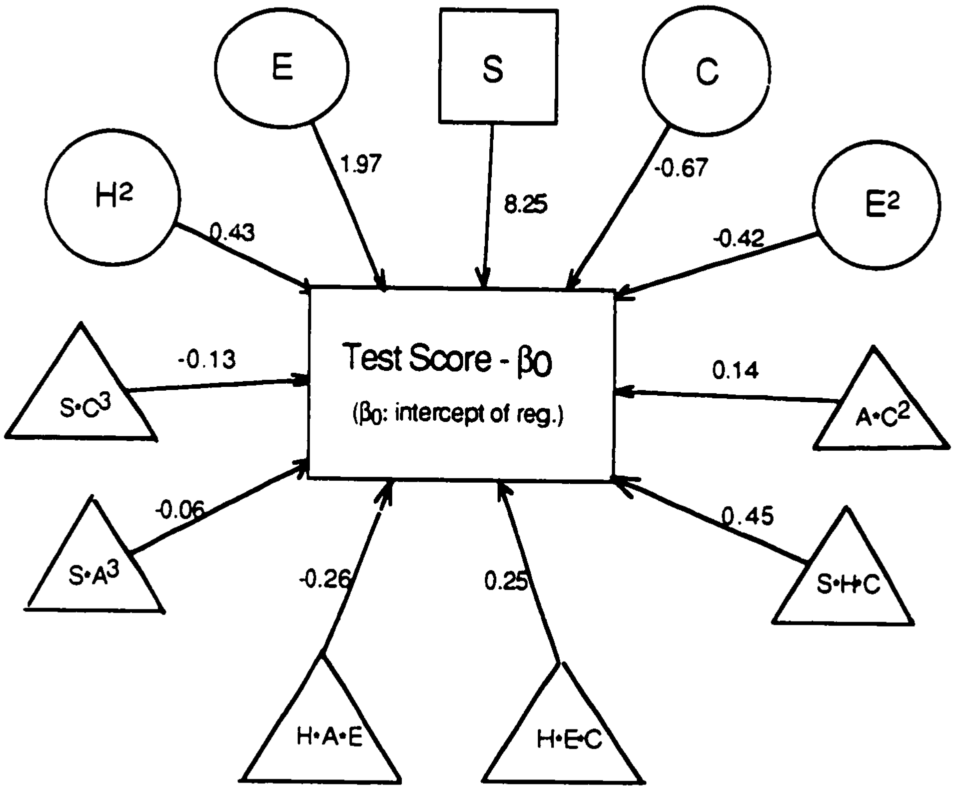
Data Base	n	p	C_1	k	K(c)	**b**(c)
SES	2871	12	$4.2 \cdot 10^{-3}$	0.2892	0.2892	0.2892•**b**
Phase II SISS	2027	21	$9.5 \cdot 10^{-3}$	0.4896	0.4896	0.4896•**b**

SUMMARY

The Chinese and American empirical models are depicted in Figures 3 and 4 respectively.

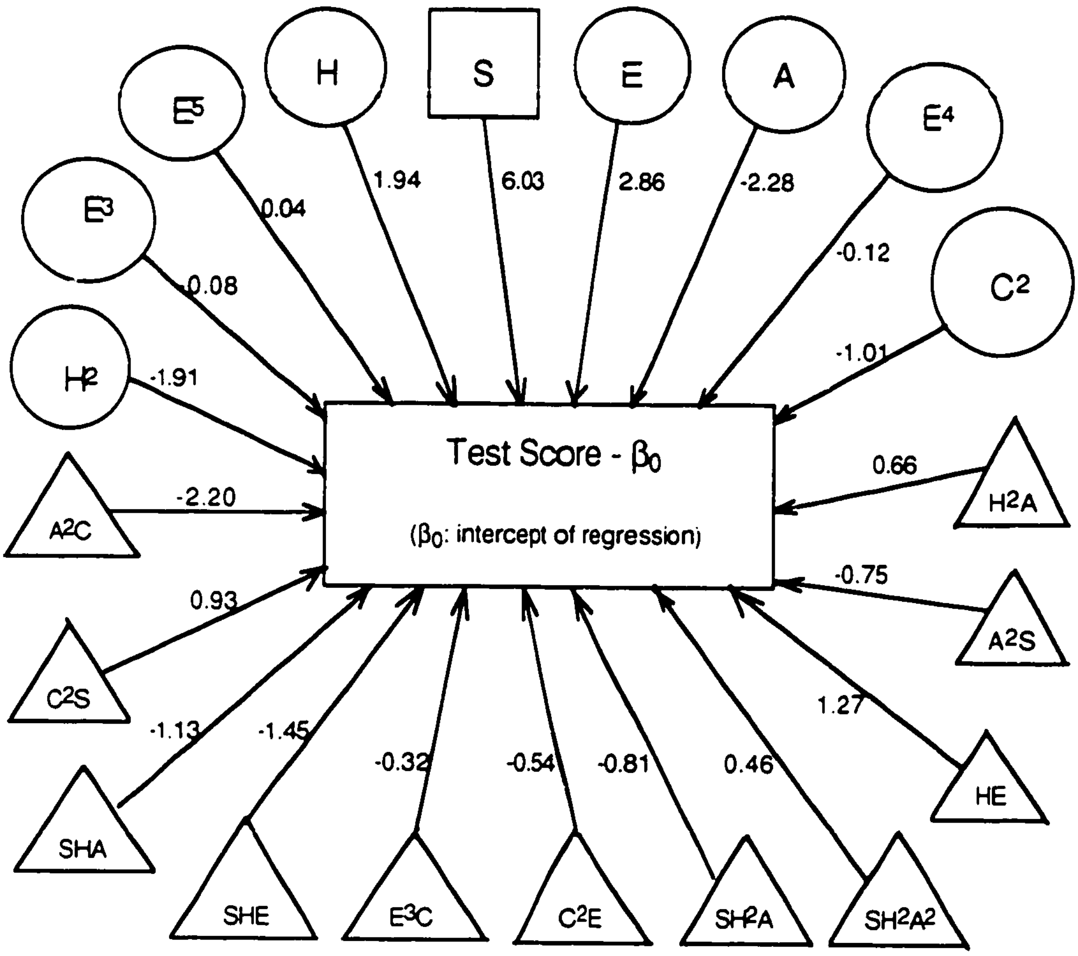
Visible predictor (SEX) is symbolized in a square; latent predictors are in circles; interactions are in triangles. Shrinkage estimates are depicted as path coefficients, and students' science achievement is adjusted by subtracting the intercept from test scores.

Figure 3*: Chinese Empirical Model



* Note: S = Sex; C = Classroom Experience; E = Effort; H = Home Background; A = Attitude.

Figure 4: American Empirical Model



* Note: S = Sex; C = Classroom Experience; E = Effort; H = Home Background; A = Attitude.

The number of factors and interactions which involve the five aspects, gender, attitude, classroom experience, personal effort and home background, are listed in Table 9. Discussion and interpretation of the empirical models are presented in Chapter 5.

Table 9: The Number of Predictors Constructed by Gender, Attitude, Classroom Experience, Personal Effort and Home Background

Country	Gender	Attitude	Classroom Experience	Personal Effort	Home Background
China	4	3	5	4	4
U.S.A	7	7	5	8	8

CHAPTER 5

SUMMARY AND DISCUSSION

The goal of this study was to investigate the effects of five factors, gender, attitude, home background, classroom experience and personal effort, on students' science achievement. To facilitate discussion of the factor effects and interactions, single factor effects are summarized under the condition that the other factors are fixed at constant levels. Interactions among the factors are analyzed at each polynomial level. The similarities and differences between the United States and China are discussed in terms of the educational, political, social and cultural contexts in each country. A philosophical recollection about the empirical approach and its value is presented in epilogue.

SINGLE FACTOR EFFECTS

The empirical results in Chapter 4 can be summarized in two regression equations, one for each country. Let y be the score of students' science test, then the regression equation for achievement of American ninth graders is:

$$y = \beta_{0U} + \beta_{1U}\cdot S + \beta_{2U}\cdot H + \beta_{3U}\cdot E + \beta_{4U}\cdot A + \beta_{5U}\cdot(H\cdot E) + \beta_{6U}\cdot(H^2) + \beta_{7U}\cdot(C^2) + \beta_{8U}\cdot(S\cdot A^2) + \beta_{9U}\cdot(S\cdot C^2) + \beta_{10U}\cdot(S\cdot H\cdot A) + \beta_{11U}\cdot(S\cdot H\cdot E) + \beta_{12U}\cdot(H^2\cdot A) + \beta_{13U}\cdot(A^2\cdot C) + \beta_{14U}\cdot(C^2\cdot E) + \beta_{15U}\cdot(E^3\cdot C) + \beta_{16U}\cdot(E^4) + \beta_{17U}\cdot(E^3) + \beta_{18U}\cdot(S\cdot H^2\cdot A) + \beta_{19U}\cdot(E^5) + \beta_{20U}\cdot(S\cdot H^2\cdot A^2) \quad (5 \text{ -- U.S.A.})$$

and the regression equation for Chinese ninth graders is:

$$y = \beta_{0C} + \beta_{1C}\cdot S + \beta_{2C}\cdot E + \beta_{3C}\cdot C + \beta_{4C}\cdot(H^2) + \beta_{5C}\cdot(E^2) + \beta_{6C}\cdot(S\cdot H\cdot C) + \beta_{7C}\cdot(H\cdot A\cdot E) + \beta_{8C}\cdot(H\cdot E\cdot C) + \beta_{9C}\cdot(C^2\cdot A) + \beta_{10C}\cdot(A^3\cdot S) + \beta_{11C}\cdot(C^3\cdot S) \quad (5 \text{ -- P.R.C.})$$

where S, H, E, A, β_{iU} and β_{jC} in both equations are defined in Table 6 and 7 respectively.

It is known in Euclidian geometry that two points determine a straight line and three points determine a plane. In other words, a unique curve on a plane can not be identified by a two-point scale. The visible factor, SEX, however, has only two values, 0 for female and 1 for male. Hence, the only gender effect that can be explored with an empirical model is linear in character. The information in Table 6 and 7 shows that the gender effect is significant in China and the United States.

The four latent factors, students' attitude, home background, classroom experience, and personal effort, are represented by their first principal components, respectively. It is interesting to note that some linear effects, such as the linear attitude and home background factors in China and the linear classroom experience factor in the U.S., are not significant in this study.

Quast, Cole, Sparks and Haubner (1963) have defined: " A line is a set of points that extends without end in two directions" (p.100). There is no direction change in a linear point extension. A curve, on the other hand, changes direction along a curvilinear extension. The rate of direction change is defined as curvature in calculus. A mathematical definition of curvature is:

> The curvature K of a curve *y = f(x)*, at any point P on it, is the rate of change in direction per unit of arc length s. Thus,
>
> $$K = (d\tau)/(ds) = \lim_{\Delta s \to 0} (\Delta\tau)/(\Delta s) = \frac{(d^2y)/(dx^2)}{\left\{1 + [(dy)/(dx)]^2\right\}^{3/2}} \qquad (6)$$
>
> (Ayres, 1964; p. 81).

Curvature is a unified approach for describing linear and curvilinear relations with a positive curvature corresponding to a convex curve, zero curvature corresponding to a straight line, and negative curvature corresponding to a concave curve.

The curvatures of the American empirical model (5 -- U.S.A.) are:

$$K_{HU} = \frac{2\beta_{6U} + 2\beta_{12U}\cdot A + 2\beta_{18U}\cdot S\cdot A + 2\beta_{20U}\cdot S\cdot A^2}{\{1 + [\beta_{2U} + \beta_{5U}\cdot E + 2\beta_{6U}\cdot H + \beta_{10U}\cdot S\cdot A + \beta_{11U}\cdot S\cdot E + 2\beta_{12U}\cdot H\cdot A + 2\beta_{18U}\cdot S\cdot H\cdot A + 2\beta_{20U}\cdot S\cdot H\cdot A^2]^2\}^{3/2}} \qquad (7 \text{ -- U.S.A.})$$

for the effect of home background,

$$K_{EU} = \frac{6\beta_{15U}\cdot E\cdot C + 12\beta_{16U}\cdot E^2 + 6\beta_{17U}\cdot E + 20\beta_{19U}\cdot E^3}{\{1 + [\beta_{3U} + \beta_{5U}\cdot H + \beta_{11U}\cdot S\cdot H + \beta_{14U}\cdot C^2 + 3\beta_{15U}\cdot E^2\cdot C + 4\beta_{16U}\cdot E^3 + 3\beta_{17U}\cdot E^2 + 5\beta_{19U}\cdot E^4]^2\}^{3/2}} \qquad (8 \text{ -- U.S.A.})$$

for the effect of personal effort,

$$K_{CU} = \frac{2\beta_{7U} + 2\beta_{9U}\cdot S + 2\beta_{14U}\cdot E}{\{1 + [2\beta_{7U}\cdot C + 2\beta_{9U}\cdot S\cdot C + \beta_{13U}\cdot A^2 + 2\beta_{14U}\cdot C\cdot E + \beta_{15U}\cdot E^3]^2\}^{3/2}}$$

(9 -- U.S.A.)

for the effect of classroom experience, and

$$K_{AU} = \frac{2\beta_{8U}\cdot S + 2\beta_{13U}\cdot C + 2\beta_{20U}\cdot S\cdot H^2}{\{1 + [\beta_{4U} + 2\beta_{8U}\cdot S\cdot A + \beta_{10U}\cdot S\cdot H + 2\beta_{1U}\cdot H^2 + 2\beta_{13U}\cdot A\cdot C + \beta_{18U}\cdot S\cdot H^2 + 2\beta_{20U}\cdot S\cdot H^2\cdot A]^2\}^{3/2}}$$

(10 -- U.S.A.)

for the effect of attitude.

The curvatures of the Chinese empirical model (5 -- P.R.C.) are:

$$K_{HC} = \frac{2\beta_{4C}}{\{1 + [2\beta_{4C}\cdot H + 2\beta_{6C}\cdot S\cdot C + \beta_{7C}\cdot A\cdot E + \beta_{8C}\cdot E\cdot C]^2\}^{3/2}}$$

(7 -- P.R.C.)

for the effect of home background,

$$K_{EC} = \frac{2\beta_{5C}}{\{1 + [\beta_{2C} + 2\beta_{5C}\cdot E + \beta_{7C}\cdot H\cdot A + \beta_{8C}\cdot H\cdot C]^2\}^{3/2}}$$

(8 -- P.R.C.)

for the effect of personal effort,

$$K_{CC} = \frac{2\beta_{9C}\cdot A + 6\beta_{11C}\cdot C\cdot S}{\{1 + [\beta_{3C} + \beta_{6C}\cdot S\cdot H + \beta_{8C}\cdot H\cdot E + 2\beta_{9C}\cdot C\cdot A + 3\beta_{11C}\cdot C^2\cdot S]^2\}^{3/2}}$$

(9 -- P.R.C.)

for the effect of classroom experience, and

$$K_{AC} = \frac{6\beta_{10C}\cdot A\cdot S}{\{1 + [\beta_{7C}\cdot H\cdot E + \beta_{9C}\cdot C^2 + 3\beta_{10C}\cdot A^2\cdot S\ ]^2\}^{3/2}}$$

(10 -- P.R.C.)

for the effect of attitude.

The single factor effect can be briefly summarized in terms of property of the curvatures:

1. Because the curvature of $y = f(S)$ is zero in equation (6) with $x = S$, gender has a linear effect on

students' achievement.

2. None of the curvatures in (7 -- U.S.A.) -- (10 -- P.R.C.) are equal to zero. Hence, the latent factors, students' attitude, home background, classroom experience and personal effort, have curvilinear effects on students' achievement in both countries.

3. In general, curvatures in (7 -- U.S.A.) -- (10 -- P.R.C.) depend on single factor effects and multi-factor interactions, i.e., a factor effect may change from positive to negative or from negative to positive over different levels of other factors.

4. The Chinese curvatures in (7 -- P.R.C.) -- (10 -- P.R.C.) are simpler than the corresponding American curvatures in (7 -- U.S.A.) -- (10 -- U.S.A.). The simplest curvatures are in (7 -- P.R.C.) and (8 -- P.R.C.). The general character of the two curvatures is shown in Figures 5 and 6. Home background has a convex effect ($K_{HC} > 0$), and personal effort has a concave effect ($K_{EC} < 0$) on students' science achievement.

Figure 5: Convexness of Y - H/K_{HC}>0

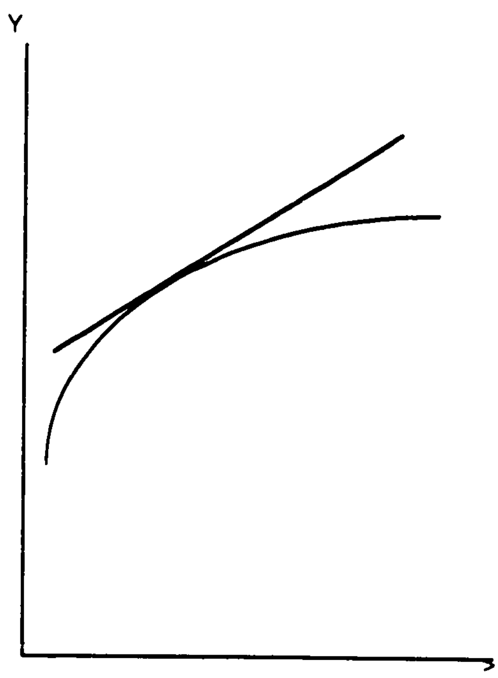


Figure 6: Concaveness of Y - H/K_{EC}<0

EFFECTS OF INTERACTIONS

In a multi-variate Taylor series, the level of an interaction is defined as the number of multiplications among the factors which construct the interaction. According to the results, the highest level of interaction is five in equation (5 -- U.S.A.) and four in equation (5 -- P.R.C.).

There are eleven interaction terms in the U.S. model equation (5 -- U.S.A.) and six interaction items in the Chinese model equation (5 -- P.R.C.). The interactions are sorted by their polynomial levels, and the U.S. results are shown in Figures 7 - 10. The Chinese results are presented in Figures 11 - 12.

Figure 7: The Second Order Interactions in U.S.A.

Factors		Interactions (◯)
Classroom Experience	(C)	
Personal Effort	(E)	E
Attitude	(A)	
Sex	(S)	
Home Background	(H)	H

Figure 8: The Third Order Interactions in U.S.A.

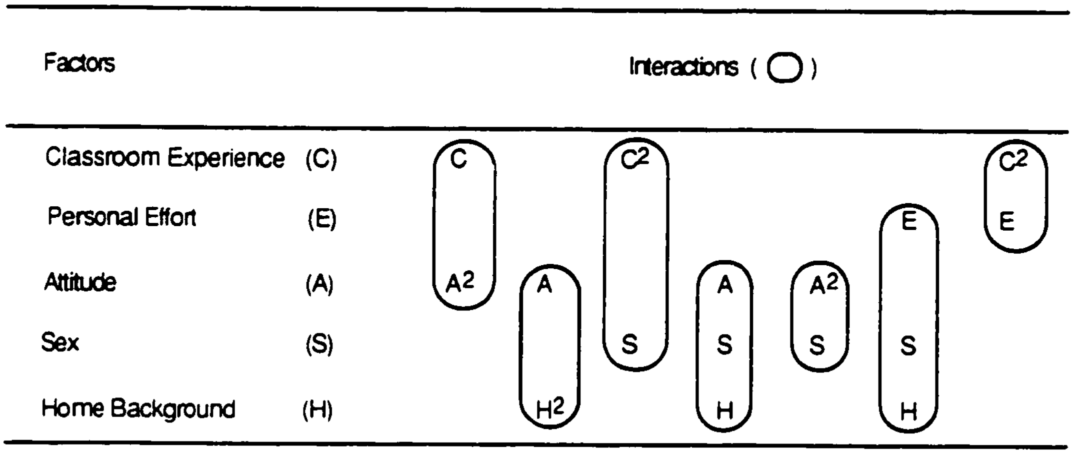


Figure 9: The Fourth Order Interactions in U.S.A.

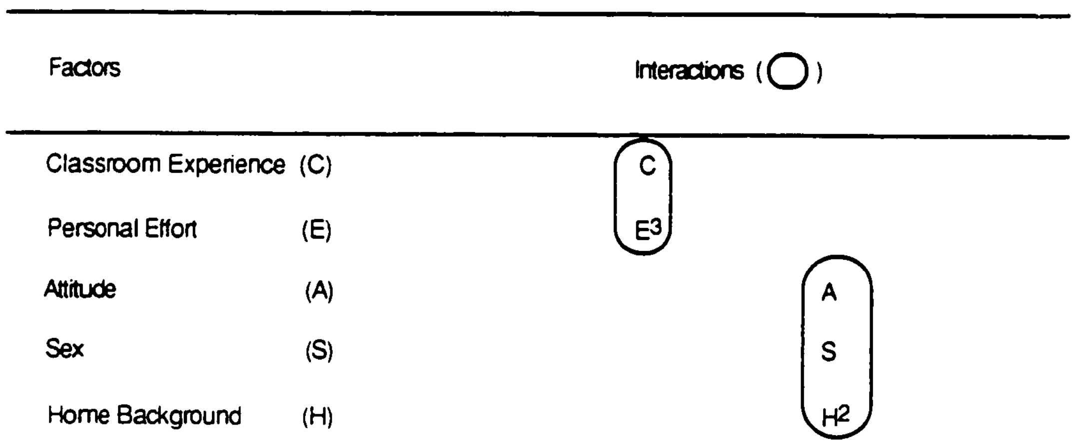


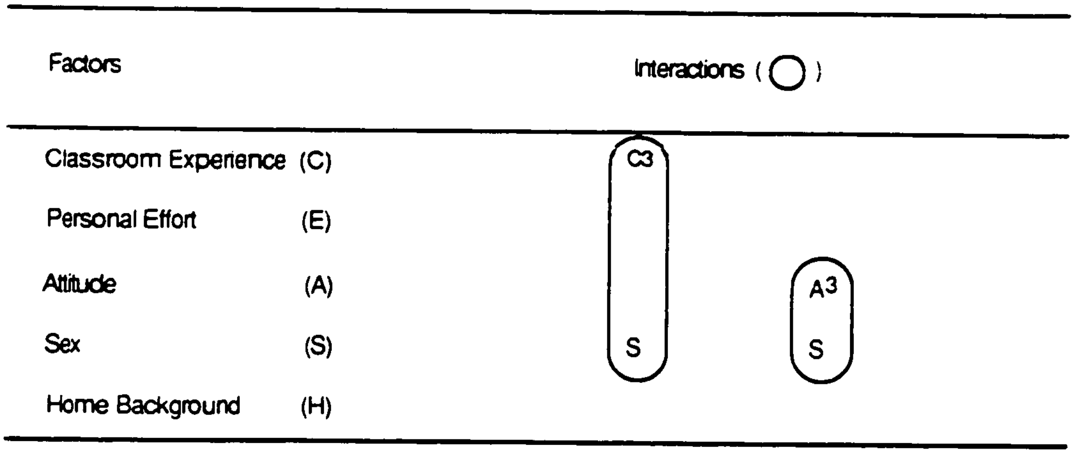
Figure 10: The Fifth Order Interactions in U.S.A.

Factors		Interactions (⬭)
Classroom Experience	(C)	
Personal Effort	(E)	
Attitude	(A)	A^2
Sex	(S)	S
Home Background	(H)	H^2

Figure 11: The Third Order Interactions in P.R.C.

Factors		Interactions (⬭)			
Classroom Experience	(C)	C	C		C^2
Personal Effort	(E)	E		E	
Attitude	(A)			A	A
Sex	(S)		S		
Home Background	(H)	H	H	H	

Figure 12: The Fourth Order Interactions in P.R.C.



The effects of interactions in Figures 7 - 12 can be summarized into four points:

1. There are more interaction items in the U.S. model than in the Chinese model. In both models, most interactions are at the third polynomial level.

2. No interactions between home background and classroom experience are significant in the U.S. model. The interactions are significant in the Chinese model through the effects of sex and personal effort.

3. No interactions between SEX and personal effort are significant in the Chinese model. In the United States, the interaction is significant through the effect of home background.

4. No interactions between attitude and personal effort are significant in the U.S. model. The interaction is significant in the Chinese model through the effect of home background.

THE INTERPRETATION OF THE RESULTS

Many interesting results have been obtained from the empirical study. The results can be classified into consistent vs. inconsistent categories. The consistent part is discussed in this section and the inconsistent part which needs further exploration is highlighted at the end of this section.

The most consistent aspects of this study are the latent factor construction and the complexity comparison of (5 -- U.S.A.) and (5 -- P.R.C.). The latent factor construction starts from identification of the latent prediction space. As shown in Figures 1 and 2, the dimension of the latent prediction space is four according to the results from principal component analyses. The complexity comparison is based on the number of single factor effects and multi-factor interactions. It has been elaborated in the previous two sections that there are more factor effects and interactions in the U.S. model than in the Chinese model. The exploratory interpretations in this section are based on differing educational, political, social and cultural contexts in the United States and the Peoples' Republic of China.

Latent Factor Construction

The structure of a factor is expressed by the factor loadings of indicators which compose the factor. The factor loadings of the four latent predictors, attitude, classroom experience, personal effort and home background, are listed in Tables 4 and 5. The corresponding elements in the two tables have the same positive or negative signs. Thus, the structural differences of the Chinese and American predictors are in their magnitudes rather than directions. The structures of two-indicator factors, home background and personal effort, have been discussed in Chapter 4. The interpretation of factor loadings in this section focuses on the factor of students' attitude and classroom experience.

In the U.S. data, indicators with the largest contributions to students' attitudes are students' interest in science (Table 4: P_ATT34) and their feeling of enjoying school science experience (Table 4: P_ATT35). This is also the case in the Chinese data (Table 5: P_ATT34 & P_ATT35). In addition, the attitudes of Chinese students are strongly affected by school pressure (Table 5: P_ATT06). In both countries, teachers' demonstrations and students' experiments are important science activities (Tables 4 & 5: P2DES08 & P2DES18). But the effects of the activities in China are not as strong as in the United States in determining the factor of classroom experience. A more important contribution in China comes from students' tests (Table 5: P2DES14).

Chinese schools are classified as key schools and general schools. A major criterion of the classification is the number of students in each school who have passed the National College

Entrance Examination. No matter what kind of pressure a school has, if many students in that school can pass the examination, the school will be promoted as a key school. This is in a line with what Deng Xiao-ping said, "It does not matter whether a cat is white or black as long as it catches rats".

In a centralized educational and political system, examination is a feasible way to avoid corruption in admission to higher education. The Chinese government takes great care to eliminate cheating on the examination. As a result, whatever power a student's parents may have, the only way for the student to pursue formal higher education is to pass the National College Entrance Examination. According to the wisdom of Chinese educators, the best way to cope with the examination is to make students take difficult tests in secondary education. Hence, test has the highest factor loading on the Chinese classroom experience.

In both the U.S. and China, teachers' demonstrations and students' experiments are important laboratory activities for science education. Nevertheless, in China, it is impossible to provide laboratories to simultaneously measure the experimental skills of millions of high school graduates in the College Entrance Examination. Thus, the National College Entrance Examination is a paper-pencil test and does not require experimental skills to achieve good scores. Moreover, China is a developing country and many schools, especially in rural areas, do not have well equipped teaching laboratories. The lack of equipment and the pressure of school examination appear to be the major reasons why laboratory activities have less effect on the Chinese classroom experience.

Complexity Comparison

It has been shown that the American model (5 -- U.S.A.) is more complicated than the Chinese model (5 -- P.R.C.). Interpretation of the difference is based on differing social and cultural contexts in each country.

The United States is a country populated by people from all over the world. The compulsory education enforced in the U.S. requires school-age children from various cultural backgrounds to complete their education at no less than the ninth grade level. Although the education is compulsory, not all U.S. students perceive an equal opportunity to learn. Coleman, Campbell, Hobson, McPartland, Mood, Weinfeld and York (1966) found that many minority students felt that somebody blocked them from success, even though they had the ability to learn. Also, the authority to determine school curricula is the responsibility of individual school districts or communities. In summary, the heterogeneous student population and diversified curricula are important factors which increase the complexity of predicting students' science achievement in the United States.

On the other hand, China has been a unified country since 221 B.C. About ninty-four percent of Chinese population is Han nationality. It is said that Chinese people are descendants of dragon and the Emperor is the son of God. "China" in Chinese means "middle kingdom". The national minorities in remote areas were ruled by courtiers of the Emperor and have been assimilated as members of the Chinese family after age-long cultural communications. The feudal system was not abandoned until the beginning of this century. The cultural foundation of the Chinese feudalism is Confucian philosophy. Confucious said: Those who study mentally should govern the people who work physically, and those who work physically should serve the people who study mentally. Thus, education is a vehicle to promote the social status of students' family.

Nevertheless, the improvement of education is closely tied to development of the Chinese economy. Because the feudal system impeded Chinese economic development for more than 2000 years, China is still a third world country in many respects. For example, compulsory education has not yet been fully enforced in China. In 1988, less than half of children finished their middle school education. Those who studied at the ninth grade were selected by regional Middle School Entrance Examinations. Thus, the Chinese model of prediction is based on the information from the selected middle school students.

The authority to determine Chinese school curriculum belongs to the central government. The national curriculum has an effect of standardizing science education over the country. Therefore, the prediction of students' science achievement should be less complicated in China given the unified culture, selected students, and standardized curriculum.

The inconsistent results between the two countries may be caused by empirical error or confounding factors which have not been included in the model construction. It should be noted that in Table 9 the number of predictors constructed by gender, attitude, classroom experience, personal effort are different in the Chinese and U.S. models. The contribution of each predictor on students' achievement depends on the path coefficient and the predictor scale. Because not all the predictors are constructed on the same scale range, larger path coefficients do not necessarily imply greater contributions to students' achievement. Questions identified by this research and subject to further exploration are:

1. Can the curvatures and interactions be further simplified to facilitate an appropriate comparative interpretation?
2. Do specific structures of the significant predictors re-appear in empirical studies in future?
3. Do the factor loadings and model complexity in the two countries have additional physical meaning?

EPILOGUE

According to constructivist epistemology, the natural world can not be known directly in any absolute sense, but must be interpreted through phenomena. Empirical inquiry is a research methodology for interpreting the world through phenomena. The more phenomena that a researcher studies, the better interpretation he or she can make. An empirical study does not degrade the value of theoretical research, but rather pursues theoretical explanations for the empirical phenomena.

The focus of this comparative research was to identify significant predictors of students' science achievement based on the common variables investigated in SES and Phase II SISS. Since no theoretical solution to this problem is in sight, an empirical approach was adopted to construct exploratory models. The underlying philosophy, however, is not empiricism. Prior knowledge has played an important role in the construction of latent factors and interpretation of results. Also, these exploratory results need further empirical reconfirmation and theoretical interpretation.

Most scientific theories are not based on a single empirical study. More empirical studies need to be conducted in the area of international comparisons. Results which are consistently reconfirmed by empirical studies form the foundation for theoretical interpretations. Hence, empirical studies are indispensable in the construction of theory.

REFERENCES

Anderson, L. W., Ryan, D. W. and Shapiro, B. J. (Eds.) (1989). The IEA classroom environmental study. Oxford, U.K.: Pergamon Press.

Anderson, O. R. (1990). The teaching and learning of biology in the United States. NY: Teachers College, Columbia University.

Ayres, Frank Jr. (1964). Theory and problems of differential and integral calculus. NY: Schaum Publishing Co.

Baker, R. (1989). Science achievement of non-science students. Doctoral Dissertation,. NY: Teachers College, Columbia University.

Bell, J. A. (1989, May). Where have all the young men (and women) gone? Paper presented at the Harvard University Symposium, "The Coming Revolution in Science Education", Cambridge, MA.

Berger, J. and Chen, S. Y. (1987). Minimaxity of empirical Bayes estimators derived from subjective hyperpriors. In *Advances in Multivariate Statistical Analysis*. A. K. Gupta (Ed.), 1-12. D. Reidel Publishing Company.

Beyer, S. L. (1990). Factors in the school environment associated with student achievement in science. Doctoral dissertation. NY: Teachers College, Columbia University.

Bilodeau, M. and Srivastava (1988). Estimation of the MSE matrix of the Stein estimator. The Canadian Journal of Statistics, 16, 153-159.

Bottenberg, R. & Ward, J. (1963). Applied multiple linear regression. (PRL-TDE-63-6). Lackland Air Force Base, TX.

Bradley, R. A. & Srivastava, S. S. (1979). Correlation in polynomial regression. The American Statistician, 1979, 33, 11-14.

Brophy, J. E. & Evertson, C. M. (1974). Process-product correlations in the Texas Teacher Effectiveness Study: Final report. Austin, Texas: University of Texas.

Budescu, D. V. (1980). A note on polynomial regression. Multivariate Behavioral Research, 15, 497-508.

Casella, George (1985). An introduction to empirical Bayes data analysis. The American Statistician, 39 (2), 83-87.

Casella, George and Berger, Roger L. (1990). Statistical inference. Pacific Grove, California: Wadsworth & Brooks/Cole Advanced Books & Software.

Chandavarkar, M. S., Doran, R. L., & Jacobson, W. J. (1991). Achievement of U.S. high-school physics students. Results of the Second International Science Study. The Physics Teacher, 23(6), 387-393.

Chandavarkar, M. S. (1988). The teaching and learning of physics in the United States. Doctoral dissertation. NY: Teachers College, Columbia University.

Chang, E. Y. (1990, April). Hierarchical levels of difficulty for test items with visual illustrations. Paper presented at the 63rd annual meeting of the National Association for Research in Science Teaching, Atlanta, Georgia.

Chang, E. Y. (1989, March). Science education in the 1970s and 1980s: What changes have taken place? Paper presented at the 62nd annual meeting of the National Association for Research in Science Teaching, San Francisco, CA.

Chang, E. Y. (1988). Science education in the 1970s and 1980s: What changes have taken place? Doctoral dissertation. NY: Teachers College, Columbia University.

Clive, T. (1983). The trial testing of items and instruments for the Second International Science Study: An analysis of results to verify the cumulative hierarchical nature of Bloom's taxonomy of educational objectives (cognitive domain). Doctoral dissertation. NY: Teachers College, Columbia University.

Cohen, J. (1968). Multiple regression as a general data-analytic system. Psychological Bulletin, 70, 426-443.

Cohen, J. & Cohen, P. (1983). Applied multiple regression/correlation analysis for the behavioral sciences. Hillsdale, NJ: Lawrence Erlbaum Associate, Publishers.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. A., Weinfeld, F. S., & York, R. L. (1966). Equality of educational opportunity (Report from the Office of Education). Washington, DC: U.S. Government Printing Office.

Comber, L. & Keeves, J. (1973). Science education in nineteen countries. NY: John Wiley & Sons.

Copas, J. (1983). Regression, prediction and shrinkage. Journal of Royal Statistical Society, series B, 45, 311-354.

Court, A. (1930). Measuring joint causations. Journal of the American Statistical Association. 25, 245-254.

Cramer, E. M. & Appelbaum, M. I. (1978). The validity of polynomial regression in the random regression model. Review of Educational Research, 48, 511-515.

Crew, K. A. (1988). Demographic illiteracy. Population Education Interchange, 17 (1). ERIC: ED292739.

Cronbach, L. J. (1957). The two disciplines of scientific psychology. American Psychologist, 12, 671-684.

Curry, E. W., Roberts, A. E., & Walling, D. D. (1986). Reinterpreting polynomial regression. Quality and Quantity, 20 (1), 19-26.

Dewit, D. J., Wister, A. V. & Burch, T. K. (1988). Physical distance and social contact between elders and their adult children. Research on Aging, 10 (1), 56-80.

Digman. (1966). Interaction and non-linearity in multivariate experiment. In R. B. Cattel (Ed.): Handbook of multivariate experimental psychology. Chicago: Rand McNally.

Doran, R. L. (1992, March). Alternative assessment of high school laboratory skills. Paper presented at the 65th annual meeting of the National Association for Research in Science Teaching. Cambridge, Massachusetts.

Doran, R. L., Boorman, J., Chan, F., & O'Rafferty, M. (1991, April). Assessment of science lab skills at the high school level. Paper presented at the 64th annual meeting of the National Association for Research in Science Teaching, Abbey at Lake Geneva, WI.

Doran, R. L. (1990). What research says: About assessment. Science and Children, 27(8), 26-27.

Doran, R. L. & Jacobson, W. J. (1984). What Research says: How are our fifth graders doing? Science and Children, 22(1), 41-42.

Dryden, M. A. (1987a, April). Modeling science classroom environments with large scale survey data: An analysis of the IEA 9th grade sample. Paper presented at the 60th annual meeting of the National Association for Research in Science Teaching, Washington, DC.

Dryden, M. (1987b). Modeling classroom environment: An analysis of achievement at the ninth grade level. Doctoral dissertation, State University of New York, at Buffalo.

Draper, Norman R. and Nostrand, Graig V. (1979). Ridge regression and James-Stein estimation: review and comments. Technometrics. 21, 451-466.

Efron, B. and Morris, C. (1972). Limiting the risk of Bayes and empirical Bayes estimators. -- Part II: The empirical Bayes case. J. Amer. Statist. Assoc., 67, 130-139.

Efron, Bradley and Morris, Carl N. (1973). Stein's estimation rule and its competitors--An empirical Bayes approach. Journal of the American Statistical Association 68, 117-130.

Efron, Bradley and Morris, Carl N. (1975). Data analysis using Stein's estimator and its generalizations. Journal of the American Statistical Association. 70, 311-319.

Efron, Bradley and Morris, Carl N. (1977). Stein's paradox in statistics. Scientific American. 236(5), 119-127.

Eichinger, J. (1990). Science education in the United States: Are things as bad as the recent IEA report suggests? School Science and Mathematics. 90 (1), 33-39.

Ekeocha, E. (1986). Correlates of science achievement: A U.S. study of fifth grade students. Doctoral dissertation, State University of New York, at Buffalo.

Ferko, A., Doran, R. L., & Jacobson, W. J. (1991). Advanced science student performance. NY: Teachers College, Columbia University.

Ferko, A. (1989). An analysis of achievement test scores and some associated findings of the Second IEA Science Study. Doctoral dissertation. NY: Teachers College, Columbia University.

Ferko, A., Doran, R. L., & Jacobson, W. J. (1990). SISS-USA: Advanced science student achievement. Studies in Educational Evaluation. 16, 457-476.

Flanders, N. A. (1970). Analyzing teacher behavior. Menlo Park, CA: Addison-Wesley.

Ghosh, M. and Lin, Pi-Erh (1986). Preface. Commun. Statist.-Theor. Meth., 15(7), vii-viii.

Ghosh, M., Saleh, A. K. Md. E. and Sen, P. K. (1989). Empirical Bayes subset estimation in regression models. Statistics and Decision, 7, 15-35.

Graybill, F. A. (1976). Theory and application of the linear model. Pacific Grove, California: Wadsworth & Brooks/Cole Advanced Books & Software.

Hebel, P. (1989). Utilisation d'estimateurs a retrecisseurs dans les modeles predictifs. Doctoral dissertation, Station de Biometrie et d'Intelligence Artificielle, INRA, Centre de Recherche de toulouse, 31326 Castanct Tolosan, France.

Hebel, P., et.,al. (1993). Shrinkage estimators applied to prediction of French winter wheat yield. Biometrics, (to be published).

Helgeson, S. L. (1988). The second IEA science study: Data related to precollege science in the U.S.A. ERIC, Clearinghouse for Science, Mathematics, and Environmental Education, Columbus, Ohio.

Hoerl, A. E. and Kennard, R. W. (1970). Ridge Regression: Biased estimation for nonorthogonal problems. Technometrics, 12, 55-67.

Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. Journal of Educational Psychology, 24, 417-441, 498-520.

Humrich, E. (1991). Sex differences in science attitude and achievement. NY: Teachers College, Columbia University.

Humrich, E., Jacobson, W. J. & Doran, R. L. (1990). Characteristics of female advanced science students in the United States. Studies in Educational Evaluation. 16(3), 477-488.

Humrich, E. (1989a, March). Characteristics of female advanced science students - Second IEA Science Study. Paper presented at the 62nd annual meeting of the National Association for Research in Science Teaching, San Francisco, CA.

Humrich, E (1989b, March). Sex difference in the Second IEA Science Study. Paper presented at the 61st annual meeting of the National Association for Research in Science Teaching, St. Louis, Missouri.

Humrich, E. (1988). Sex and science achievement. Doctoral dissertation. NY: Teachers College, Columbia University.

Husen, T. (1987).Policy impact of IEA research. Comparative Education Review, 31 (1), 29-46.

IEA (1988). Science achievement in seventeen countries: A preliminary report. NY: Pergamon Press.

Jacobson, W. & Doran, R. (1988; 1991). Science achievement in the United States and sixteen countries: A report to the public. NY: Teachers College, Columbia University.

Jacobson, W., Doran, R., & Schneider, A. E. (1991). The second IEA science study: United States. The International Journal of Educational Research. 16(1).

Jacobson, W. (1990). Guest Editoral: A dismal picture. Rays of hope. Journal of Research in Science Teaching. 27(4), 287-288.

Jacobson, W. (1988). Demographic illiteracy. Population Education Interchange. 17(1), 1-2.

Jacobson, W., Doran, R., Humrich, E., Kanis, I. (1988). International science report card. NY: Second IEA Science Study, Teachers College, Columbia University.

Jacobson, W. & et. al. (1987, September). The Second IEA Science Study--U.S. revised edition. Paper presented at a Meeting of the General Assembly of IEA, New York, NY.

Jacobson, W., Takemura, S., Doran, R., Kojima, S., Humrich, E., & Miyake, M. (1986). Analyses and comparisons of science curricula in Japan and the United States. Washington, D. C.: United States Studies of Education in Japan and the Center for Educational Statistics, Office of Educational Research and Improvement.

Jacobson, W. & Doran, R. (1986). What do our ninth graders think of science and school? The Science Teacher, 53(5), 59-61.

Jacobson, W. J., Keeves, J., Soydhurum, P., Hernandez, D., Fonacier, J., Tamir, P., Doran, R., Connelly, M., Crocker, R., Kass, H. & Finegold, M. (1986, March). Second International Science Study -- A multi-national report.Paper presented at the 59th annual meeting of the National Association for Research in Science Teaching, San Francisco, CA..

Jacobson, W. & Doran, R. (1985a). How are our ninth graders doing? The Science Teacher, 52(8), 24-27.

Jacobson, W., & Doran, R. (1985b). The second international science study: U.S. results. Phi Delta Kappan, 66, 414-417.

Jagodzinski, W., Weede, E. & Tiefenbach, H. (1981). Testing curvilinear propositions by polynomial regression with particular reference to the interpretation of standardized solutions. Quality and Quantity, 15 (5), 447-463.

James, W. and Stein, C. (1961). Estimation with Quadratic Loss. Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, 1, 361-380. Berkeley, CA: University of California Press.

Jennrich, Robert I. and Oman, Samuel D. (1986). How much does Stein estimation help in multiple linear regression? Technometrics, 28(2), 113-121.

Johnson, D. E. (1992). Multivariate Statistical Methods. Textbook of STAT 720 at Statistics Department, Kansas State University.

Joreskog, K. G. & Sorbom, D. (1984). Analysis of linear structural relationships by maximum likelihood, instrumental variables, and least squares methods. Mooresville, Indiana: Scientific Software, Inc.

Judge, G. G. and Bock, M. E. (1978). Statistical Implications of Pre-test and Stein-Rile Estimators in Econometrics. NY: North-Holland.

Kanis, I. (1991). Ninth grade lab skills. The Science Teacher. 58(1), 29-33.

Kanis, I. (1990). What is the current status of laboratory skills of our grade five students? The Science Teacher's Bulletin. 54(1), 20-29.

Kanis, I. (1988). An analysis of the science process practical examination administered to grade five and grade nine students in the United States of America. Doctoral dissertation. NY: Teachers College, Columbia University.

Kanis, I., Doran, R. & Jacobson, W. (1990). Assessing science process laboratory skills at the elementary and middle/junior high levels. NY: Teachers College, Columbia University.

Keeves, J. & Rosier, M. (1981). Guidelines for the Second IEA Science Study. (IEA/SISS/4), Australian Council for Educational Research. Hawthorn, Victoria 3122, Australia.

Keeves, J. (1992). The IEA Study of Science III: Changes in science education and achievement: 1970 to 1984. NY: Pergamon Press.

Kelly, T., Beggs, D., McNeil, K., Eichelberger, T. & Lyon, J. (1969). Research design in the behavioral science: Multiple regression approach. Carbondale, Ill.: Southern Illinois University Press.

Kerlinger, F. & Pedhazur, E. (1973). Multiple regression in behavioral research. NY: Holt, Reinhart & Wiston.

Kotz, S. and Johnson, N. L. (1988). Encyclopedia of statistical science. NY: John Wiley & Sons, Inc.

Krieger, J. H. (1984). Study indicates some improvement in U.S. science education. Chemical and Engineering News, 62 (24), 26-28.

Lemmer, Hermanus H. (1988). Shrinkage estimators. Encyclopedia of Statistical Science, 8, 452-456. NY: John Wiley & Sons, Inc.

Lindley, D. and Smith, A. F. M. (1972). Bayes estimate for the linear model. J. Roy. Statist. Soc. (Ser B), 34, 1-41.

Lindquist, E. F. (1951). Design and analysis of experiments in psychology and education. Boston: Houghton Mifflin Co..

Liski, Erkki P. (1982). A test of the mean square error criterion for shrinkage estimators. Commun. Statist.--Simula. Computa., 11(5), 543-562.

Liu, Hsiao-Ling (1981). Biased estimation in polynomial regression. Doctoral dissertation, The University of Alabama.

Loucks, S. F. (1975). A study of the relationship between teacher level use of the innovation of individualized instruction and student achievement. Doctoral dissertation. Austin: University of Texas.

Matloff, Norman S. (1982). James-Stein regression estimation in a prediction context. Commun. Statist.--Simula. Computa., 11(5), 589-601.

McNeil, K. (1970). Meeting the goals of research with multiple linear regression. Multivariate Behavioral Research, 5, 375-386.

McNeil, K. & Spaner, S. (1971). Highly correlated predictor variables in multiple regression. Multivariate Behavioral Research, 6, 117-125.

Micik, J. (1986). Science achievement in an American School: A case study. Doctoral dissertation. NY: Teachers College, Columbia University.

Miller, J. K. (1988). An analysis of science curricula in the United States. Doctoral dissertation. NY: Teachers College, Columbia University.

Miller, J. K. (1986). An analysis of science curricula in the United States. Second IEA Science Study (SISS), United States. ERIC: ED275534.

Miller, J. K. (1985). An analysis of science curricula in the United States. NY: Teachers College, Columbia University.

Nebebe, Fassil and Stroud, T. W. F. (1986). Bayes and empirical Bayes shrinkage estimation of regression coefficients. Canadian Journal of Statistics, 14, 267-280.

Nebebe, Fassil and Stroud, T. W. F. (1988). Bayes and empirical Bayes shrinkage estimation of regression coefficients: A cross-validation study. Journal of Educational Statistics 13(3), 199-213.

O'Rafferty, M. H. & Doran, R. L. (1991, April). Science teachers: A comparative perspective. Paper presented at the 64th annual meeting of the National Association for Research in Science Teaching, Abbey at Lake Geneva, WI.

O'Rafferty, M. H. (1991). A descriptive analysis of the performance of grade 9 pupils in the united states on practical science tasks. Doctoral dissertation, State University of New York at Buffalo.

Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. Philosophical Maganize, 6 (2), 559-572.

Postlethwaite, T. N. & Wiley, D. E. (1992). The IEA Study of Science II: Science achievement in twenty-three countries. NY: Pergamon Press.

Postlethwaite, T. N. (1974). Introduction. Comparative Education Review, 18 (2), 1974.

Quast, W. G., Cole, W. L., Sparks, T. M. & Haubner (1983). Mathematics. Boston: Houghton Mifflin Company.

Rim, E. D. & Coller, A. R. (1978). In search of nonlinear process-product functions in existing schooling effects data: I. A reanalysis of the first grade reading and mathematics data from the Stallings and Kaskowitz Follow Through Study. ERIC: ED179289.

Rolph, John E. (1976). Choosing shrinkage estimators for regression problems. Commun. Statist.--Theoryand Method, 5(9), 789-802.

Rosier, M. (1987). The second international science study. Comparative Education Review, 31(1), 106-128.

Rosier, M. & Keeves, J. (1991). The IEA study of science I: Science education and curricula in twenty-three countries. NY: Pergamon.

SAS (1982). SAS user's guide: statistics. Cary, NC: SAS Institute Inc.

Saunder, D. (1956). Moderator variables in prediction. Educational and Psychological Measurement, 16, 209-222.

Schneider, A., Muller, E., Doran, R. & Jacobson, W. (1991). The second international science study: Discussion, items, and results: Population 2 -- Grade 9. NY: Teachers College, Columbia University.

Schneider, A., Muller, E., Doran, R. & Jacobson, W. (1990). The second international science study: Items, and results: Population 1 -- Grade 5. NY: Teachers College, Columbia University.

Sennetti, John T. and Kakar, Rajesh K. (1980). On the inadmissibility of the James-Stein estimator. American Statistical Association Proceedings of the social Statistics Section, 499-501.

Singh, R. S. (1991). James-Stein rule estimators in linear regression models with multivariate-t distributed error. Austral. J. Statist., 33(2), 145-158.

Smith, V. K. (1990). Estimating recreation demand using the properties of the implied consumer syrplus. Land Economics, 66(2), 111-120.

Soar, R. S. (1966). An integrative approach to classroom learning. Philadelphis, Pa.: Temple University.

Soar, R. S. (1968). Optimum teacher-pupil interaction for pupil growth. Educational Leadership Research Supplement I, 275-280.

Soar, R. S. (1971). Follow Through classroom process measurements. Gainesville, Fla.: Institute for Development of Human Resources, University of Florida.

Soar, R. S. (1973). Follow-Through classroom process measurement and pupil growth (1970-71): Final report. Gainesville: College of Education, University of Florida.

Soar, R. S. & Soar, R. M. (1972). An empirical analysis of selected Follow Through programs: An example of a process approach to evaluation. In I. J. Gordon (Ed.), Early childhood education. Chicago: National Society for the Study of Education.

Soar, R. S. & Soar, R. M. (1976). An attempt to identify measures of teacher effectiveness from four studies. Journal of Teacher Education, 27, 261-267.

Sockloff, A. (1976). The analysis of nonlinearity via linear regression with polynomial and product variables: an examination. Review of Educational Research, 46 (2), 267-291.

Stanley, T. D. (1990). Stein-rule least squares. The American Statistician, 44(3), 263-264.

Stein, C. (1956). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability 1, 197-206. Berkeley, CA: University of California Press.

Stein, C. (1981). Estimation of the mean of a multivariate distribution. Ann. Statist., 43, 1481-1490.

Stimson, J. A., Carmines, e. G. & Zeller, R. A. (1978). Interpreting polynomial regression. Sociological Methods and Research. 6 (4), 515-524.

Tamir, P., Doran, R., Vari, P., Chye, Y. O., Mlyake, M. & Im, I. (1989, March). Results of the practical laboratory process tests in the Second International Science Study. Paper presented at the 62nd annual meeting of the National Association for Research in Science Teaching, San Francisco, CA.

Thompson, J. R. (1968). Some shrinkage techniques for estimating the mean. Journal of American Statistical Association, 63, 113-122.

Walberg, H. (1971). Generalized regression models in educational research. American Educational Research Journal, 8, 71-91.

Walberg, H. J. (1983). Scientific literacy and economic productivity in international perspective. Scientific Literacy. Issued as Volume 112, Number 2, of the proceedings of the American Academy of Arts and Science.

Walberg, H. J. (1981). A psychological theory of educational productivity. ED206042.

Walker, D. (1976). The IEA six subject survey: An empirical study of education in twenty-one countries. NY: John Wiley & Sons.

Wang, Jianjun (1993). Shrinkage Regression for Prediction of Kansas Wheat Yields. M.S. Report at Department of Statistics, Kansas State University.

Weigel, K. A. (1991). Improving estimates of fixed effects in a mixed linear model. Journal of Dairy Science, 74(9), 3137-3182.

Whittemore, Alice S. (1989). Errors-in-variables regression using Stein estimates. The American Statistician, 43(4), 226-228.

Wold, H. (1975). Another hour another day. In Hauser, R. M., Sewell, W. H., and Alwin, D. (eds.) Schooling and Achievement in American Society. Standford University. Standfort, CA.

Wold, H. (1979). Model construction and evaluation when theoretical knowledge is scarce. Cahier 79.06. University of Geneva, Department of Economics, Geneva.

Wold, H. (1982) The basic design and some extensions. In: Joreskog, K. G., Wold, H. (eds.) System under Indirect Observation, Part II. Amsterdam: North-Holland Press.

APPENDIX 1: THE CHINESE VERSION OF SISS INSTRUMENTS[1]

LEGEND

STUDENT BOOKLET[2]

2M: Core Test[3]

2A: Rotated Test[4]

2B: Rotated Test

2C: Rotated Test

2D: Rotated Test

2Q: Mathematics Test

ST: Student Questionnaire, Opinion Questionnaire and Description of Science Learning

TEACHER BOOKLET

2OTL: Opportunity to Learn (OTL) Questionnaire

TE: Teacher Questionnaire

PRINCIPAL BOOKLET

SC: School Questionnaire

NOTES

[1] The items which are employed in the Chinese and U.S. surveys are noted by "*" in this appendix.

[2] Word Knowledge Test in Table 1 is excluded from this appendix because the test is not valid for Chinese students.

[3] Core Test is a science achievement test which has been taken by every student who participated in the international study.

[4] Rotated Tests are science achievement tests which are grouped pairwisely into six combinations, 2A & 2B, 2A & 2C, 2A & 2D, 2B & 2C, 2B & 2D and 2C & 2D. Based on the SISS design, each student should be randomly assigned to take one of the six combined tests.

INTERNATIONAL STUDY
OF SCIENCE EDUCATION

理科测验

卷 号 **2M**

中国IEA国家中心
北京

这个测验包括与理科各科目有关的题目。每一个问题都已提供了A、B、C、D、E、五个答案，请选择你认为正确的一个答案，在此答案旁的方框里划一“✓”，在每页答完后，请你按指示将已经填好的答案抄写在答卷纸上。

如果你不知道某一问题的答案，不必浪费时间，先把它搁下，继续答下题。如有时间剩余，则可尝试先前漏答的。如果不能确知题目的答案也可尝试作答，但不可瞎猜。

练习例题

P1.地球绕太阳一周要多少时间？

- ☐A 一天
- ☐B 一星期
- ☐C 一个月
- ☑D 一年
- ☐E 以上都不对

地球需要一年时间才能围绕太阳转一周，所以D是正确答案，应在D旁的方框内划一“✓”，如上面所示。

请翻往下一页。

现在继续练习以下P2、P3和P4三个例子。在你认为正确的一个答案的方框里划一"✓"。

P2. 怎样将水变成冰？

- ☐A 加热
- ☐B 快速搅动
- ☐C 加进食盐
- ☐D 倒进一个浅碟子内
- ☐E 加以冷却

P3. 在南半球，一年中的哪一天太阳光照射时间最长？

- ☐A 一月二十一日
- ☐B 三月二十一日
- ☐C 十二月二十二日
- ☐D 九月二十三日
- ☐E 六月二十二日

你可能要在部分题目中找出一个错误的答案或是和其他不相配的，例如

P4. 下列哪一项和其他的不属同一类？

- ☐A 鹰
- ☐B 狮
- ☐C 鼠
- ☐D 象
- ☐E 鹿

现在请将练习题P2、P3和P4的答案抄写在答卷纸上。
未得指示前，请勿翻往后页。

现在开始作答，请选出你认为正确的一个答案，并在答案旁的方框里划一“✓”。

* 1. 在我们的太阳系中，太阳是唯一发出大量光和热的物体。为什么我们能看见月球？

☐A 月球反射太阳的光。

☐B 月球缺少大气层。

☐C 月球是一颗恒星。

☐D 月球是太阳系中最大的物体。

☐E 月球比太阳较接近地球。

* 2. 一艘火箭大约需要多少时间才能抵达月球？

☐A 两小时

☐B 数小时

☐C 数月

☐D 一光年

☐E 数年

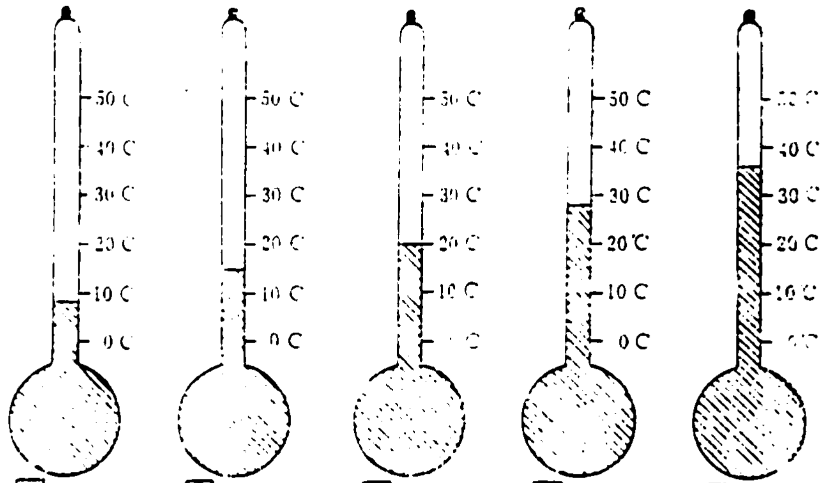
* 3. 一男孩在树下看到一只鸟从树皮的缝隙中捕捉昆虫。哪一图显示这只鸟的喙形？

现在请将第一、二和三题已填妥的答案抄写在答卷纸上。
抄写完毕后，请翻往下一页。

下表显示在三日内不同时间所记录的温度读数。以下二题皆与此表有关

	上午6时	上午9时	中午12时	下午3时	下午6时
星期一	15°C	17°C	20°C	21°C	19°C
星期二	15°C	15°C	15°C	10°C	9°C
星期三	8°C	10°C	14°C	14°C	13°C

* 4. 下列哪一个图显示星期三上午6时的温度？



☐ A ☐ B ☐ C ☐ D ☐ E

* 5. 其中一日有冷风吹过，发生在什么时候呢？

☐A 星期一早上

☐B 星期一下午

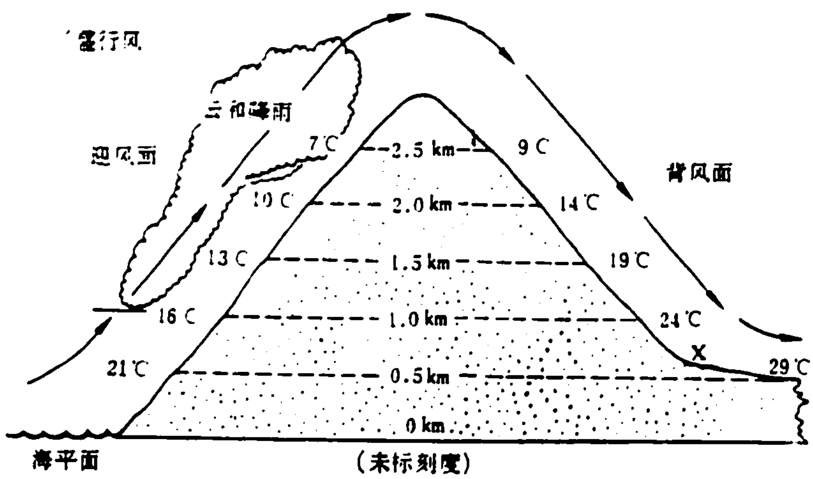
☐C 星期二早上

☐D 星期二下午

☐E 星期三下午

现在请将已填妥的答案抄写在答卷纸上。
抄写完毕后，请翻往下一页。

*6. 下图表示盛行风向及山两面不同高度的平均温度。



背风面的山脚下（位置X处）应属于哪一种地表特征？

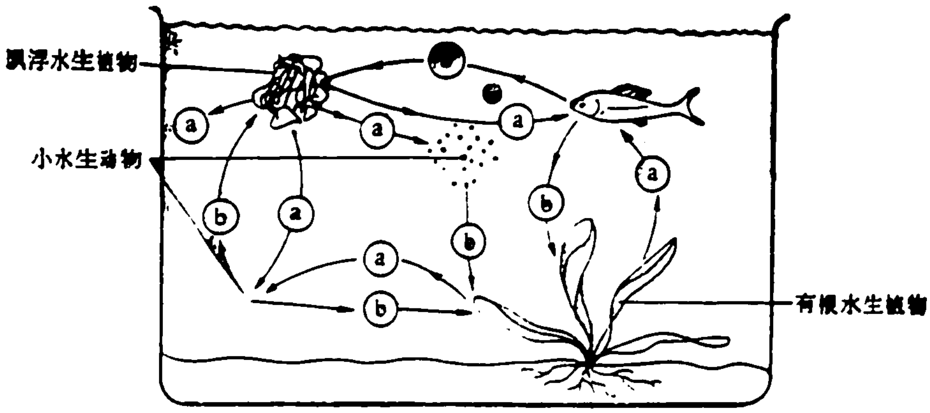
- ☐A 干旱地区
- ☐B 丛林
- ☐C 冰川
- ☐D 大湖
- ☐E 雨林

*7. 在高山的岩石中发掘出与现代的海洋贝壳类动物形状极相似的化石。最合理的解释是什么？

- ☐A 此海洋贝壳类动物可在海中或陆上生存。
- ☐B 海洋生物曾一度具备能够呼吸大气中空气的器官。
- ☐C 埋藏此化石的岩石是在海底形成的。
- ☐D 在某些情况下，海洋生物移居到了陆上。
- ☐E 海洋生物是由陆上生物进化而来的。

现在请将已填妥的答案抄写在答卷纸上。
抄写完毕后，请翻往下一页。

* 8. 下图是一个水生生物互相依赖的例子。白天时这些生物用去或放出ⓐ或ⓑ如箭头所示。



从各项中选择ⓐ及ⓑ的正确答案。

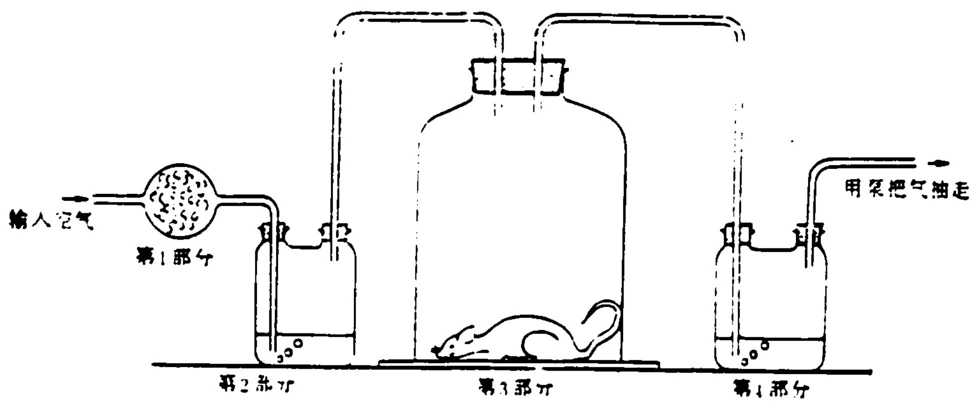
- ☐A ⓐ是氧及ⓑ是二氧化碳。
- ☐B ⓐ是氧及ⓑ是碳水化合物。
- ☐C ⓐ是氮及ⓑ是二氧化碳。
- ☐D ⓐ是二氧化碳及ⓑ是氧。
- ☐E ⓐ是二氧化碳及ⓑ是碳水化合物。

*9. 一女孩找到一动物的头骨，她不知道这属何种动物。她只能确定这动物是以捕食其它动物为生。
哪一线索引致此结论？

- ☐A 眼窝朝向两侧。
- ☐B 头骨的长度比宽度大得多。
- ☐C 沿头骨上部有一凸出的脊纹。
- ☐D 其中四只牙齿长而尖。
- ☐E 颚骨可以上下及左右移动。

现在请将已填妥的答案抄写在答卷纸上.
抄写完毕后，请翻往下一页.

* 10．下图装置显示动物呼吸时放出二氧化碳。



第1部分所盛的物质能除去所通过的空气中所含的二氧化碳。第2及第4部分所盛的液体，当通过二氧化碳时均起可见的变化。

使用下列各盛载动物的容器，哪一个能最快获得结果？

☐A　小的容器

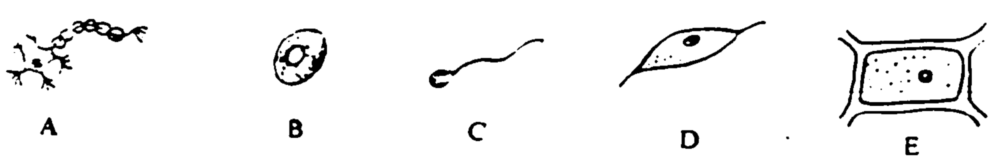
☐B　大的容器

☐C　置强光下的容器

☐D　用黑布包裹的容器

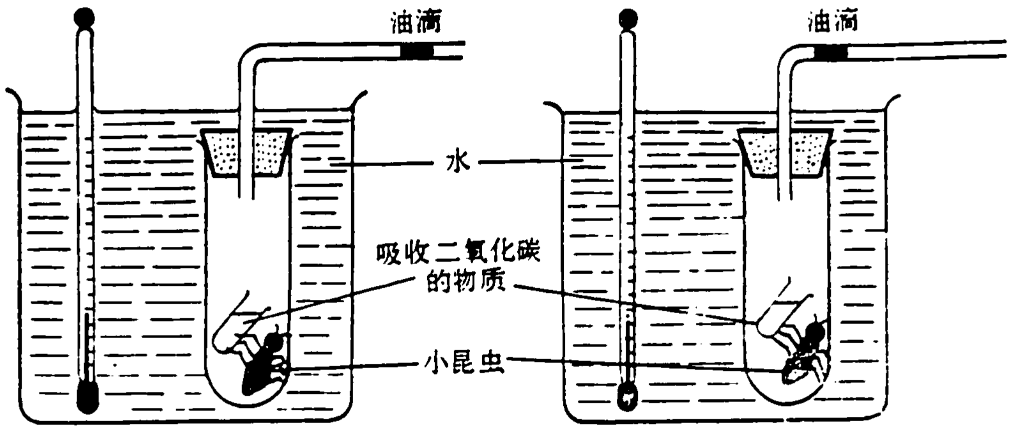
☐E　置有湿棉花的容器，用以保持空气湿润。

* 11．下面哪一种细胞常见于人类的神经系统？



现在请将已填妥的答案抄写在答卷纸上。
抄写完毕后，请翻往下一页。

* 12. 动物吸入氧气及放出二氧化碳。普通空气中含有很少二氧化碳。



刚装置好的设备　　　　5分钟以后的设备

下列哪一项可由上图装置量度？

☐A　动物的运动率。

☐B　动物产生的热量。

☐C　动物的呼吸率。

☐D　二氧化碳对动物的影响。

☐E　动物吸收二氧化碳的份量。

* 13. 以下哪一项有关种子的描述是正确的？

☐A　所有植物都会产生种子。

☐B　所有果实皆含有大量种子。

☐C　所有种子都可食用。

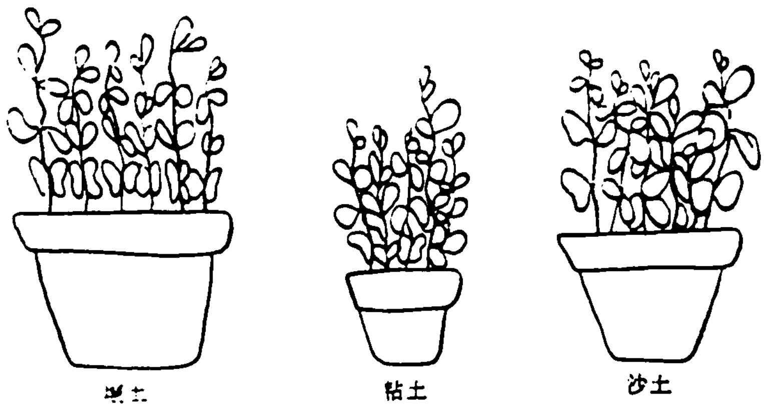
☐D　每一颗种子都含有幼苗，养份及种皮。

☐E　种子的养份经常贮存在子叶中。

现在请将已填妥的答案抄写在答卷纸上。
抄写完毕后，请翻往下一页。

* 14. 一女孩希望研究粘土、砂土及壤土中哪一种土壤最适宜种植豆类。她用三个花盆分别盛载三种不同的土壤，然后分别在每盆种植相同数目的豆类植物。她把花盆平排在窗槛上，并给每盆浇同样的水。

下图表示她所采用的花盆及数日后所得的结果。



为什么这实验<u>不能</u>达到予期的目的？

☐A 其中一盆植物较其他植物获得较多的阳光。

☐B 每盆所盛的泥土份量不同。

☐C 其中一盆应置于黑暗中。

☐D 应浇不同份量的水。

☐E 植物于窗槛上可能太热。

*15. 为什么牛奶放在冰箱内比置于室温下需要较长时间才会变酸？

☐A 低温使牛奶中的水份变成冰。

☐B 低温将乳脂分离。

☐C 低温减慢了细菌的作用。

☐D 低温使苍蝇不会飞近。

☐E 低温使牛奶的表面形成一层奶皮。

现在请将已填妥的答案抄写在答卷纸上。
抄写完毕后，请翻往下一页。

* 16．将一昆虫种群中的雄虫经过处理，使其失去产生精子的能力。这样将会不会使该种群的昆虫数目减少？

- □A 不会，因为雌虫仍会产卵。
- □B 不会，因为昆虫仍然会交配。
- □C 不会，因为这样不会改变后代的突变率。
- □D 会，因为这样会急剧地降低生殖率。
- □E 会，因为雄虫会死亡。

* 17．将2g（克）锌和1g硫混合加热产生硫化锌，差不多所有锌及硫均用去而没有剩余。如将2g硫及2g锌混合加热，将会发生什么？

- □A 产生的硫化锌将会含有约两倍份量的硫。
- □B 只剩下约1g的硫。
- □C 只剩下约1g的锌。
- □D 剩余硫及锌各约1g。
- □E 不会产生任何作用。

* 18．两种单质经化合后产生一种有毒的化合物。下列有关这两种单质性质的结论，哪一项可从上述资料获得？

- □A 这两种单质都一定是有毒的。
- □B 最少有一种单质一定是有毒的。
- □C 一种单质是有毒的而另一种则是无毒的。
- □D 两种单质都不是有毒的。
- □E 两种单质都不必是有毒的。

现在请将已填妥的答案抄写在答卷纸上．
抄写完毕后，请翻往下一页．

* 19. 在铁的表面涂上油漆是为防止铁生锈。下列哪一项是最主要的原因？

- A 油漆防止氮与铁接触。
- B 油漆与铁产生化学作用。
- C 油漆防止二氧化碳与铁接触。
- D 油漆使铁的表面较平滑。
- E 油漆防止氧及湿气与铁接触。

* 20. 在化学变化过程中，下列哪一类粒子可从其他原子获得，从本身原子失去或与其他原子共有？

- A 距离原子核最远的电子
- B 距离原子核最近的电子
- C 从原子核中来的电子
- D 从原子核中来的质子
- E 从原子核中来的中子

* 21. 图中所示的木块有多长？

- A 10厘米
- B 20厘米
- C 25厘米
- D 30厘米
- E 35厘米

现在请将已填妥的答案抄写在答卷纸上。
抄写完毕后，请翻往下一页。

* 22. 玛莉与珍妮皆购买了同类的橡胶球。玛莉说："我的球的弹力比你的好。"珍妮回答说："我希望你能证实你的说法。"玛莉应该怎样办？

☐A 让两球于离地等高处堕下，留意哪一个反弹得较高。

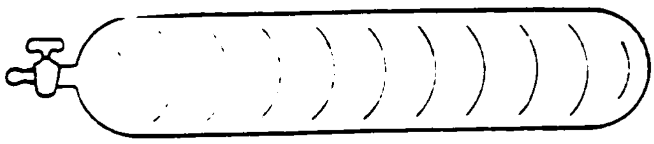
☐B 把两球向墙掷去，留意它们反弹时离墙的距离。

☐C 让两球于离地不同高度堕下，留意哪一个反弹得较高。

☐D 把两球向地面掷下，留意它们反弹的高度。

☐E 用手触摸两球找出哪一个较硬。

* 23. 将铁容器内的空气抽出（抽真空）后，称它的重量。将容器充满氢气后，再称它的重量。



充满氢气容器的重量与真空容器的重量相比是怎样的？

☐A 有氢气的轻些。

☐B 有氢气的重些。

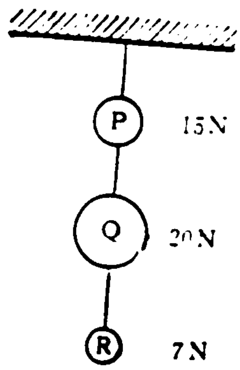
☐C 两个的重量一样。

☐D 要视容器内气体的体积而定。

☐E 要视容器内气体的温度而定。

现在请将已填妥的答案抄写在答卷纸上。
抄写完毕后，请翻往下一页。

* 24. 物体P、Q和R的重量分别为15N（牛顿）、20N和7N。如下图所示，用细线将它们悬挂起来。



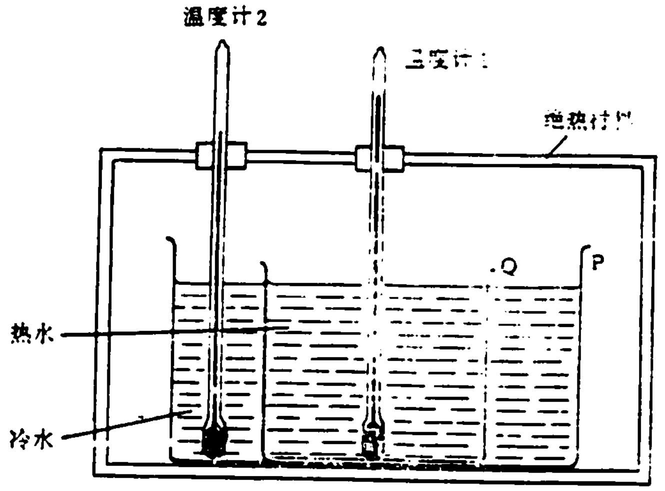
在P和Q之间的一段线的张力是多少？

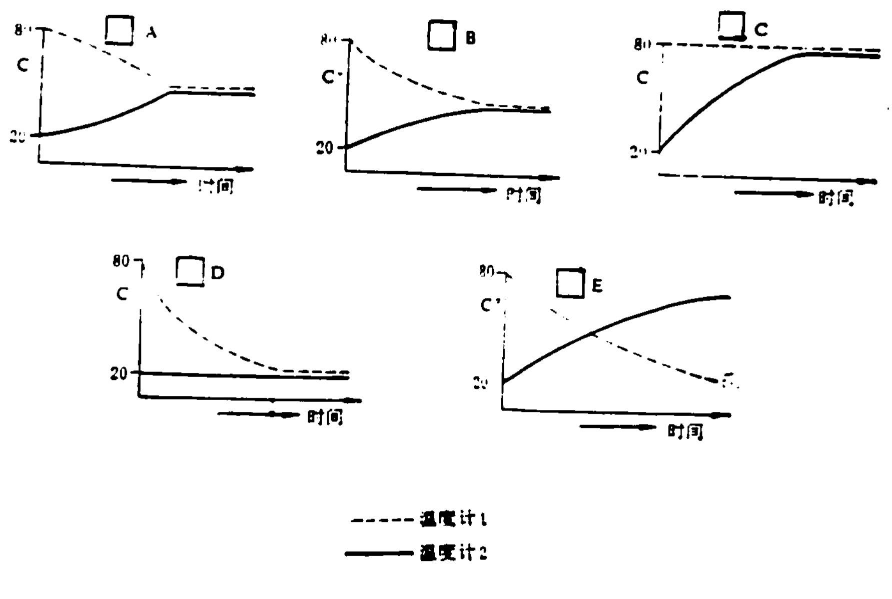
- ☐A 42N
- ☐B 35N
- ☐C 27N
- ☐D 15N
- ☐E 7N

现在请将已填妥的答案抄写在答卷纸上。
抄写完毕后，请翻往下一页。

★ 25. 用图中所示装置，将100g（克）20°C的水倒进外面的容器P内，然后每隔一段时间用温度计2量度其温度。同时，将100g80°C的水倒进里面的容器Q内，并每隔一段时间用温度计1量度其温度。

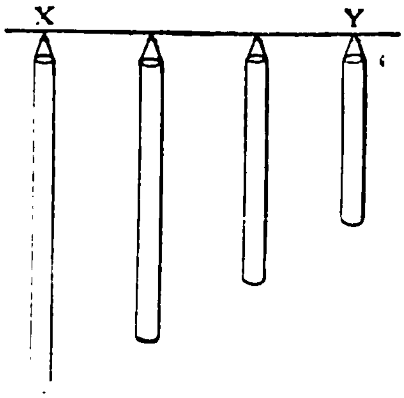
下面的图表，哪一幅最能表达两个容器内水温的变化？





现在请将已填妥的答案抄写在答卷纸上.
抄写完毕后，请翻往下一页。

* 26．把一条长金属管切成四段，长度各不相同。然后将四段金属管挂起来（如图），造成一组鸣钟（乐器）。用锤子敲击这管时，哪一只管发出的音调最低。



☐A 管X

☐B 管Y

☐C 所有管发出的音调高低一样。

☐D 要试验了才知道。

☐E 要看敲在哪一部分。

* 27．在一个炎热的晴天，靠窗的桌子上放着两只同样的杯子。一杯装满了水，另一杯装满了汽油。数小时后，发觉两个杯子里的液体都减少了，剩下的汽油比剩下的水较少。这实验说明什么？

☐A 所有液体都会蒸发。

☐B 汽油比水变得热些。

☐C 有些液体比其他液体蒸发得快些。

☐D 只有在太阳照射下，液体才会蒸发。

☐E 水比汽油变得热些。

现在请将已填妥的答案抄写在答卷纸上。
抄写完毕后，请翻往下一页。

* 28. 一只可装上两个干电池的手电筒。要它亮起来，两个干电池应怎样放置？

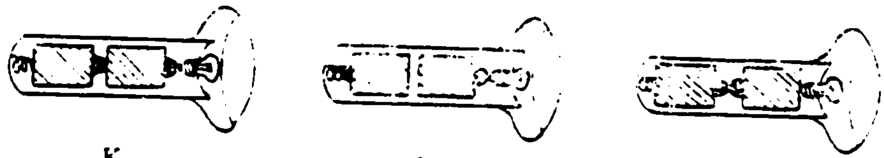
□A 如图K

□B 如图L

□C 如图M

□D 图L或图M其中一种

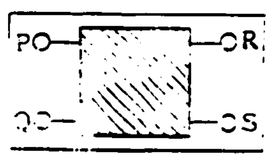
□E 以上各项所示均不正确



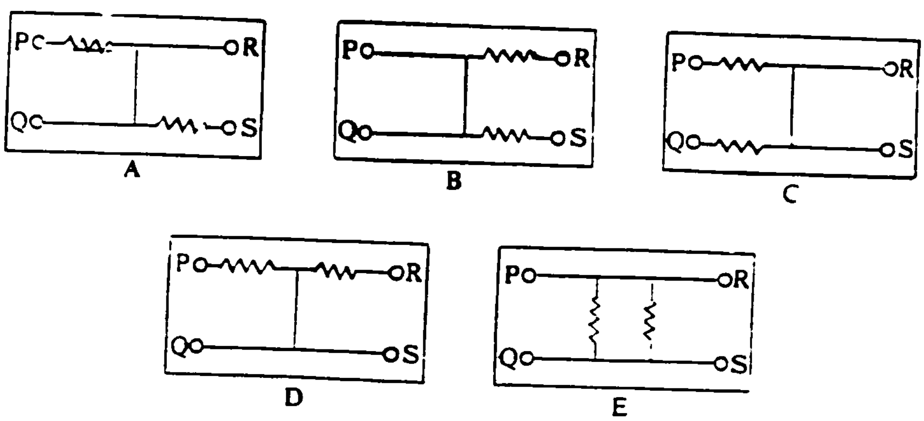
K　　L　　M

* 29. 下图表示一有四个端钮P、Q、R和S的盒子。
以下是观察所得。

1. P和Q之间有若干电阻。
2. P和R之间的电阻等于P和Q之间电阻的两倍。
3. Q和S之间差不多没有电阻。



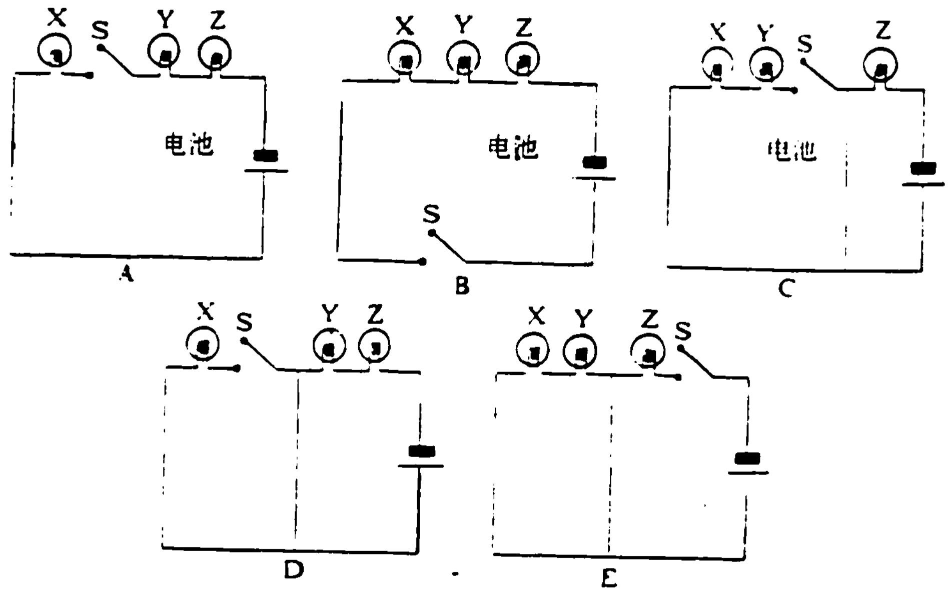
假如图中电阻符号都代表相同的电阻。下列哪一图最可能表示上述盒子内的电路？



现在请将已填妥的答案抄写在答卷纸上。
抄写完毕后，请翻往下一页。

* 30．X、Y和Z代表在一电路中的三个灯炮，该电路也包括电池和电键S，当电键打开时，X不亮而Y和Z亮了。

下列哪一图表示上述的电路？



现在请将已填妥的答案抄写在答卷纸上。

INTERNATIONAL STUDY
OF SCIENCE EDUCATION

理科测验

卷号 **2A**

中国IEA国家中心
北京

1. 有些种子于黑暗中萌发得最好，有些则需在光照下才萌发得最盛，更有些则无论于黑暗中或光照下都萌发得同样地好。一女孩进行一项实验，以分辨种子的类别，她应该将种子放在湿新闻纸上，并

☐A 置于温暖而黑暗处.

☐B 将一部分放于有光处，将另一部分放于黑暗处.

☐C 置于温暖而有光处.

☐D 将一些种子放在干新闻纸上，一同置于有光处.

☐E 将另一些种子放在干新闻纸上，一同置于黑暗处.

2. 花通常不能产生种子除非

☐A 曾与昆虫接触.

☐B 于夏天时开放.

☐C 开花的那株植物生长于良好的泥土中.

☐D 能制造花蜜.

☐E 适当的花粉落在柱头上.

3. 饭馆出售再经过加热的肉类通常不受鼓励，有时甚至为法律所禁. 下列哪一项是主要的原因?

☐A 大多数人不喜欢.

☐B 再加热使有价值的矿物盐失去.

☐C 将食物加热两次是不经济的.

☐D 微菌会在加热了的肉类迅速繁殖.

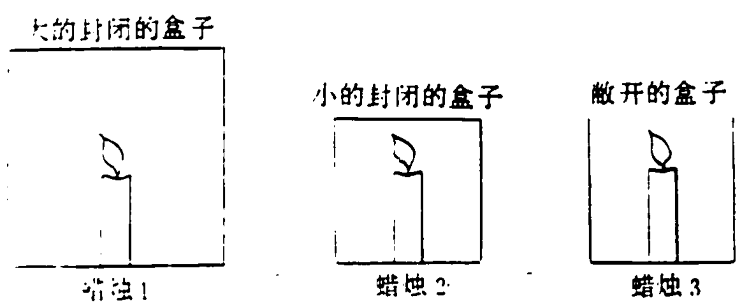
☐E 再加热使蛋白质含量降低.

现在请将已填妥的答案抄写在答卷纸上。
抄写完毕后，请翻往下一页。

4. 一男孩使用气筒给自行车轮胎打气，不久他觉得使用气筒非常吃力。为什么？

- ☐A 轮胎中的空气向气泵推压。
- ☐B 空气由气泵中漏出。
- ☐C 气泵发热以致难以握持。
- ☐D 气泵变得胶着以致难以推动。
- ☐E 轮胎比气泵较大。

5. 图中有三只不同的盒子，里面各放着一枝相同的蜡烛。把蜡烛同时燃点。



蜡烛熄灭的次序应是怎样的？

- ☐A 1，2，3
- ☐B 2，1，3
- ☐C 2，3，1
- ☐D 1，3，2
- ☐E 3，2，1

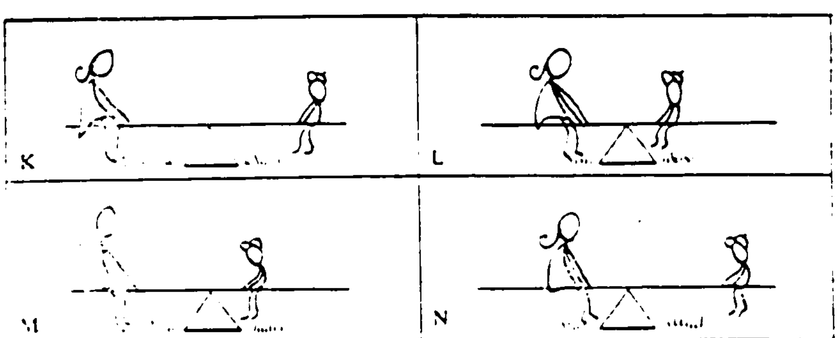
现在请将已填妥的答案抄写在答卷纸上。
抄写完毕后，请翻往下一页。

6. 凝固点是液体凝固时的温度；沸点是液体沸腾时的温度。

下列各项是盐水和清水的凝固点和沸点的比较，哪一项是正确的？

	盐水的凝固点	盐水的沸点
□A	比清水低	比清水低
□B	比清水低	比清水高
□C	比清水高	比清水低
□D	比清水高	比清水高
□E	和清水一样	和清水一样

7. 一女孩想和她的弟弟玩跷跷板。女孩的体重是50千克力，弟弟是25千克力。下面哪一图所示的做法最能使他们保持平衡？



□A 图K

□B 图L

□C 图M

□D 图N

□E 以上都不对

现在请将已填妥的答案抄写在答卷纸上。
抄写完毕后，请翻往下一页。

8. 从实验所得，以下是各种哺乳类动物从初生下来到体重增加一倍所需的时间。

哺乳类动物	初生动物体重增加一倍所需日数	母乳内蛋白质含量的百分率
人	180	1.6
马	60	2.0
牛	17	3.5
猪	18	5.9
羊	10	6.5
狗	8	7.1
兔	6	10.4

这实验结果提示了什么？

☐A 较大的哺乳类动物乳汁的蛋白质含量较高。

☐B 较小的哺乳类动物乳汁的蛋白质含量较高。

☐C 哺乳类动物乳汁的蛋白质含量愈高，则初生哺乳类动物体重增加一倍所需的时间愈短。

☐D 哺乳类动物乳汁的蛋白质含量愈高，则初生哺乳类动物体重增加一倍所需的时间愈长。

☐E 哺乳类动物乳汁的蛋白质含量，与初生哺乳类动物体重增加一倍所需的时间看来无关。

9. 在海上，两艘船上的水手可以藉高呼来互相通讯，为何在太空里同样距离的两艘太空船的船员不能利用此方法？

☐A 温度太低，

☐B 声音会反射，

☐C 太空船内的压力太大.

☐D 已超过声速，

☐E 太空船之间没有空气.

现在请将已填妥的答案抄写在答卷纸上。
抄写完毕后，请翻往下一页。

10． 一支粗细均匀的棒，以中点为支点，在同一平面上，承受着两个各等于10N（牛顿）的力。在哪一种情形下该棒会转动？

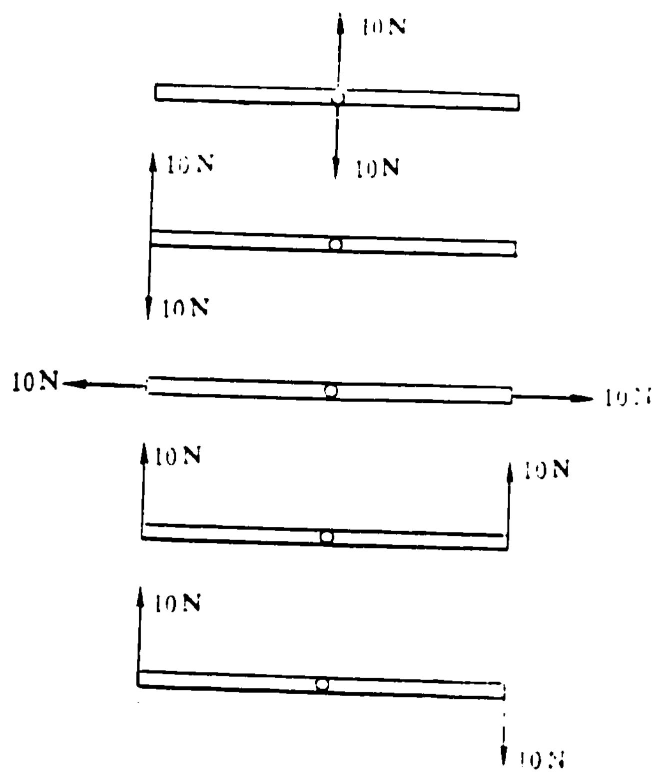
☐ A

☐ B

☐ C

☐ D

☐ E



11． 人体患急性炎症验血时，往往发现哪种细胞有增多的现象？

☐A 白细胞

☐B 血小板

☐C 表皮细胞

☐D 肌肉细胞

☐E 红细胞

现在请将已填妥的答案抄写在答卷纸上。
抄写完毕后，请翻往下一页。

12. 下列各种动物，哪个不属于脊索动物门？

□A 扬子鳄

□B 乌贼

□C 喜鹊

□D 蝙蝠

□E 文昌鱼.

13. 6克碳在20克氧气里充分燃烧．对生成物及剩余物的下列叙述哪个是正确的？

□A 生成26克一氧化碳，反应物没有剩余.

□B 生成26克二氧化碳，反应物没有剩余.

□C 生成12克二氧化碳，剩余14克氧气.

□D 生成22克二氧化碳，剩余4克氧气.

□E 生成12克一氧化碳，剩余14克氧气.

14. 在一张比例尺为一百万分之一的地图上，量得甲、乙两地的距离为5厘米，这两地的实际距离是：

□A 5公里

□B 25公里

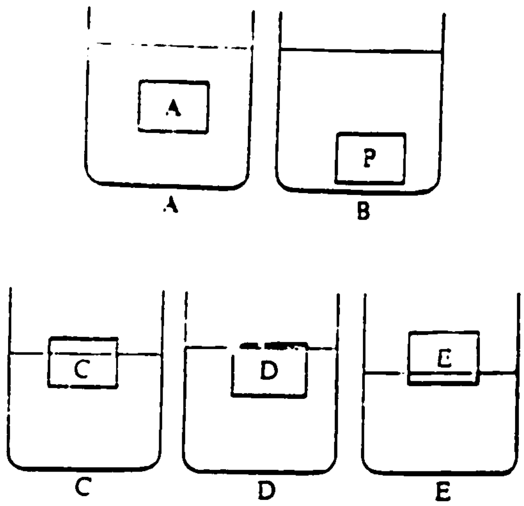
□C 50公里

□D 250公里

□E 500公里.

现在请将已填妥的答案抄写在答卷纸上．
抄写完毕后，请翻往下一页．

15. 有五块体积、形状相同的不同材料做成的物体，分别放在水中，出现了图中的沉浮情况，哪个物体受到的浮力最小？



□A B块受到的浮力最小。

□B E块受到的浮力最小。

□C D块受到的浮力最小。

□D 无法判断。

□E C块受到的浮力最小。

现在请将已填妥的答案抄写在答卷纸上。

INTERNATIONAL STUDY
OF SCIENCE EDUCATION

理科测验

卷号 **2B**

中国IEA国家中心
北京

1. 进行光合作用所用的能量通常从哪里取得？

- A 叶绿素
- B 叶绿体
- C 阳光
- D 碳水化合物
- E 二氧化碳

2. 以下哪一个器官的位置<u>不在</u>腹腔内

- A 肝脏
- B 肾脏
- C 胃
- D 膀胱
- E 心脏

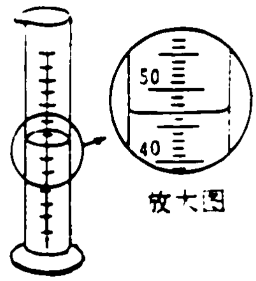
3. 乙酸（醋成分之一）的分子式是CH_3COOH。

   一个乙酸分子含有多少个原子？

- A 1
- B 2
- C 3
- D 6
- E 8

现在请将已填妥的答案抄写在答卷纸上。
抄写完毕后，请翻往下一页。

4. 一量筒盛有一定体积的水，图中放大部分显示出其旁边现有水面的情况。水的体积是多少？



放大图

A 50cm

B 45cm

C 43cm

D 47cm

E 46cm

5. 配戴同样的装备，一个人在月球上能跳得比在地球上高，下列哪一项是最好的解释？

A 在月球上，他的质量较小。

B 月球上的重力较地球上的小。

C 在月球上，他距离地球较远。

D 月球上没有空气，因而没有阻力

E 牛顿运动定律在月球上不适用。

现在请将已填妥的答案抄写在答卷纸上。
抄写完毕后，请翻往下一页。

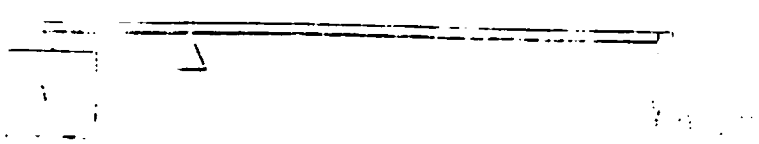
6. 将粉状的铁和硫混和加热，会产生下列哪一项结果？

- A 一种单质
- B 另两种单质
- C 一种溶液
- D 一种合金
- E 一种化合物

7. 在下[illegible]一种情况，水分蒸发得最快？

- A [illegible]热而[illegible]的日子
- B [illegible]热而[illegible]的日子
- C [illegible]冷而[illegible]的日子
- D [illegible]的日子
- E [illegible]的日子

8. [illegible]所示，[illegible]杆提起重物W，比直接提起它[illegible]？



- A [illegible]
- B [illegible]些
- C [illegible]
- D [illegible]
- E [illegible]少

现在请将已填妥的答案抄写在答卷纸上。
抄写完毕后，请翻往下一页。

[illegible]. 触摸一个金属盘子会觉得它比塑料的要冷些，为什么？

- A 金属的温度总是较塑料低。
- B 金属辐射出来的热，比塑料要多得多，所以冷却得较快。
- C 金属更容易把热从手中传导出来。
- D 塑料比金属是较佳的热传导体
- E 一个平滑的表面较一个粗糙的表面容许更密切的接触。

[illegible]. 请参阅下图。

[illegible]进入[illegible]，为什么？

- A 空气的压力比水大。
- B 空气比水重。
- C 空气占有空间，所以一定要先把空气放出，水才能进入。
- D 玻璃管太细。
- E 水的质量比空气的质量大。

现在请将已填妥的答案抄写在答卷纸上。
抄写完毕后，请翻往下一页。

11． 目前根据不同地层的植物化石表明，地球上最古老、最原始的陆生植物是哪一类？

__A 蓝藻

__B 裸蕨类植物

__C 裸子植物

__D 被子植物

__E 苔藓类植物

12． 28°c时，100克饱和硫酸铜溶液中加入1克无水硫酸铜粉末后，该饱和溶液的质量符合下列的哪一项？

__A 减少

__B 增加

__C 不变

__D 先增加，后减少

__E 无法知道

13． 人在岸边看到水里的鱼，其实是

__A 变浅了的鱼的实象．

__B 变浅了的鱼的虚象．

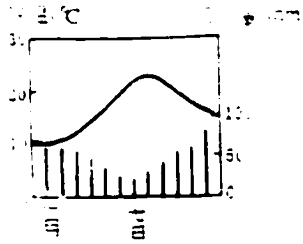
__C 变深了的鱼的实象．

__D 变深了的鱼的虚象．

__E 既未变浅也不变深的鱼本身．

现在请将已填妥的答案抄写在答卷纸上．
抄写完毕后，请翻往下一页．

14. 根据下图判断该地是属于下列气候类型中的哪一种：



□ A　地中海式气候

□ B　热带季风气候

□ C　温带季风气候

□ D　热带草原气候

□ E　亚热带季风性湿润气候。

[illegible]. 氯元素的原子结构示意图是：＋17 287 下列说法错误的是：

□ A　氯原子核中有17个质子，

□ B　氯原子很容易失去最外层的7个电子.

□ C　氯元素是非金属元素，

□ D　氯原子容易得到一个电子

□ E　氯原子K层、L层电子已排满

现在请将已填妥的答案抄写在答卷纸上。

INTERNATIONAL STUDY
OF SCIENCE EDUCATION

理科测验

卷号 **2C**

中国IEA国家中心
北京

1. 水壶和锅通常都用铜制成。下列是一些可能的原因，哪一个是错误的？

□A 铜是不良导热体。

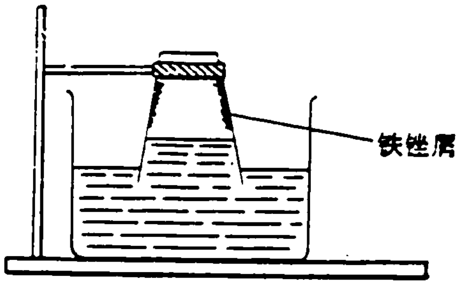
□B 铜是一种坚韧的金属。

□C 铜经打磨后较悦目。

□D 铜容易被制成不同形状。

□E 铜不溶于热水。

2. 在一只玻璃杯的内壁粘上铁粉，然后将杯拑紧，垂直置于水中。如图所示，杯内的水位逐渐上升一个短距离。



下列哪一项是最好的解释？

□A 水会在杯内凝结。

□B 铁放出一种能溶于水的气体。

□C 铁锈代替了原来的铁，因而所占的空间比原来为少。

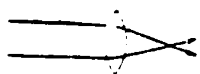
□D 铁和杯内空气中的氧产生作用。

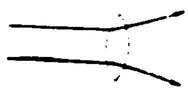
□E 杯内的氧溶于水中。

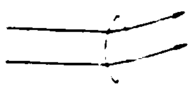
现在请将已填妥的答案抄写在答卷纸上。
抄写完毕后，请翻往下一页。

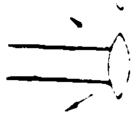
5. 下列各图中，哪一幅最能表示光通过放大镜后的情形？

□A 

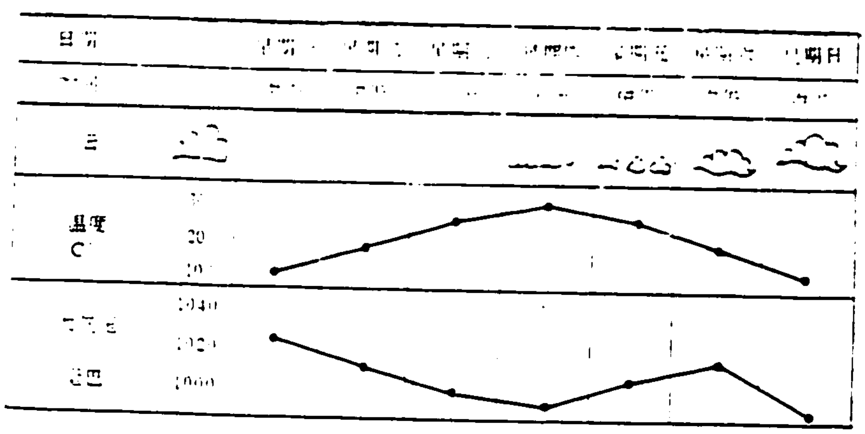
□B 

□C 

□D 

□E 

6.



从以上图表所得的资料，下列哪一项可作为预[illegible]会下[illegible]

□A 气压在下降及温度在上升。

□B 气压在上升及温度在下降。

□C 气压在下降及温度在下降。

□D 气压在上升及温度在上升。

□E 气压及温度均稳定。

现在请将已填妥的答案抄写在答卷纸上。
抄写完毕后，请翻往下一页。

7. 许多年前农民已发现如果将腐烂的鱼埋在接近种植玉蜀黍的地方，可以使玉蜀黍生长得更好。腐烂的鱼可能提供了一些什么来刺激植物的生长？

☐A 能量

☐B 矿物质

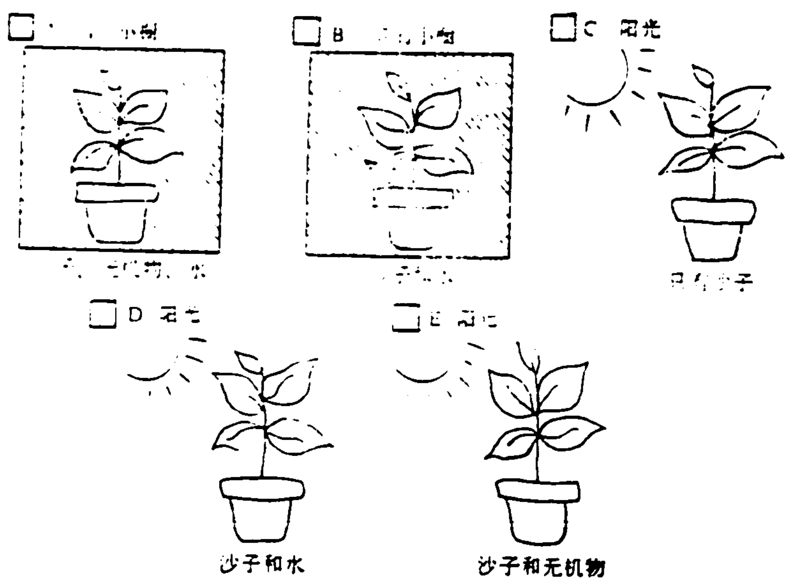
☐C 蛋白质

☐D 氧

☐E 水

8. 一女孩有一个想法，认为植物需要泥土中的矿物质才能健康成长，她将一株植物放在阳光下，如下图所示。

为了验证她的想法，她还需要用另一株植物来做实验。她应该采用下列哪一个做法？



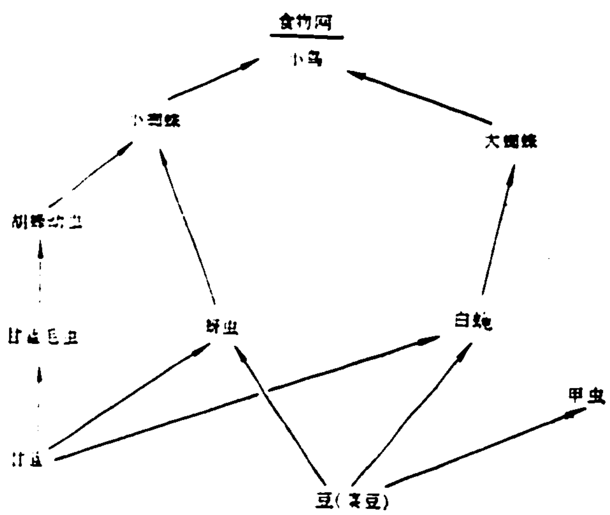
现在请将已填妥的答案抄写在答卷纸上。
抄写完毕后，请翻往下一页。

9. 肾脏的主要功能是什么？

- ☐A 制造抗体抵抗疾病
- ☐B 消化食物
- ☐C 使血液循环
- ☐D 制造红血球
- ☐E 从血液中移去废物

10. 下图为一食物网。食物网表示出各种动物所吃的食物。有些动物吃植物，而其本身可为别种动物所吃，吃动物的也可能被别的动物所吃。箭号由食物指向其食者，

例如：椰菜→蚜虫（蚜虫吃椰菜）



如果将豆类全部摇去毁灭，哪一种动物会消失？

- ☐A 大蜘蛛
- ☐B 甲虫
- ☐C 蚜虫
- ☐D 白蝇
- ☐E 小鸟

现在请将已填妥的答案抄写在答卷纸上。
抄写完毕后，请翻往下一页。

11. 在幼苗的根部施用过量化肥，会引起植株萎蔫是因为

□A 化肥使土壤变干燥了.

□B 化肥对根有刺激作用.

□C 没有及时浇水，影响根的吸水.

□D 化肥使土壤溶液的浓度大于根毛细胞液的浓度.

□E 幼苗太小了。

12. 保温瓶的瓶胆是用双层玻璃做的，夹层里的空气抽得很稀薄，夹层的玻璃壁上镀了银，光亮得象镜子一样，瓶口上还有软木塞. 这样保温瓶的保温是由于

□A 主要避免了空气的对流和热的传导.

□B 主要避免了辐射热的反射和吸收辐射热.

□C 主要避免了热的传导和热的辐射.

□D 主要避免了空气的对流和热的辐射.

□E 三种热的传递方式都尽可能避免了.

13. 取少量下列物质分别放在较多的水里，振荡后不形成溶液的物质是

□A 硫酸铜

□B 浓硫酸

□C 食盐

□D 细砂土

□E 生石灰

现在请将已填妥的答案抄写在答卷纸上.
抄写完毕后，请翻往下一页.

14．乙醇（酒精的化学名称）的分子式为CH_3CH_2OH，一个乙醇分子中

(1)有四个氢原子和一个氢分子。

(2)有六个氢原子。

(3)共有九个原子。

(4)共有八个原子。

两种说法都是错误的为：

□A (1)，(2)

□B (3)，(4)

□C (1)，(4)

□D (2)，(4)

□E (2)，(3)

15．A地位于东经100°，北纬60°，B地位于东经70°，北纬25°，B地位于A地的

□A 东北

□B 西北

□C 东南

□D 西南

□E 南面

现在请将已填妥的答案抄写在答卷纸上.

INTERNATIONAL STUDY
OF SCIENCE EDUCATION

理科测验

卷号 **2D**

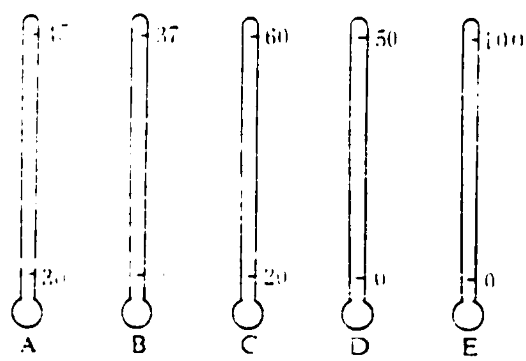
中国IEA国家中心
北京

1. 假如早上你面向北方，太阳应该在哪里？

- A 你的左方
- B 你的右方
- C 你的后方
- D 你的前方
- E 你的上方

2. 图中是五只不同的摄氏温度计。人体的正常温度是37°C（摄氏），病人的体温大致在36°C至42°C之间。图中哪一只温度计最适宜用来准确地量度体温？

- A 温度计A
- B 温度计B
- C 温度计C
- D 温度计D
- E 温度计E



3. 一个碟子内盛有2g（克）盐和8g水的溶液。将碟子置于阳光下，碟子内的东西有5g蒸发到空气中。剩下来的5g溶液里有些什么？

- A 2g盐和3g水
- B 多于1.5g盐和2.5g水
- C 1g盐和4g水
- D 只有5g水
- E 少于0.5g盐和4.5g水

现在请将已填妥的答案抄写在答卷纸上。
抄写完毕之后，请翻往下一页。

下表表示五种单质的熔点。

单质	熔点
铝	660°C
镁	610°C
铁	1535°C
铅	327°C
铜	1083°C

把这些单质的小样置于炉中加热至1555°C，如此时把炉的温度降低，下列哪一个样最先凝固？

A 铝

B 镁

C 铁

D 铅

E 铜

现在请将已填妥的答案抄写在答卷纸上。
抄写完毕之后，请翻往下一页。

5. 下表列出些一烷烃系化学药品的名称、分子式及沸点。

名称	分子式	沸点
甲烷	CH_4	-161°C
乙烷	C_2H_6	-88°C
丙烷	C_3H_8	-42°C
戊烷	C_5H_{12}	36°C
已烷	C_6H_{14}	69°C
庚烷	C_7H_{16}	99°C

丁烷为一种烷烃，它的沸点是0°C，分子式最可能是什么?

☐A C_8H_{18}

☐B C_3H_6

☐C C_3H_8

☐D C_4H_{10}

☐E C_3H_{12}

6. 地球的表面虽经过数百万年来风雨的侵蚀，仍然是不平坦的。下列哪一项最能解释这个现象?

☐A 海面的水位不断转变。

☐B 地壳运动不断发生。

☐C 时间还未足够。

☐D 地球表面的温度差距不够大。

☐E 侵蚀不够剧烈。

现在请将已填妥的答案抄写在答卷纸上。
抄写完毕后，请翻往下一页。

7. 如果进食同等份量的下列各种食物，哪一种会供给身体最多的蛋白质？

☐A 马铃薯
☐B 苹果
☐C 饭
☐D 面包
☐E 鸡

8. 流汗对身体最主要的帮助是什么？

☐A 保持身体凉快。
☐B 保持皮肤湿润。
☐C 防止感冒。
☐D 排除体内的盐份。
☐E 排除体内过量的水份。

9. 血液在人体内有很多功能，以下哪一项不是血液的功能？

☐A 消化食物
☐B 抵抗疾病
☐C 把养料输送到细胞
☐D 把废物带离细胞
☐E 把氧气输送到身体各部分

现在请将已填妥的答案抄写在答卷纸上。
抄写完毕后，请翻往下一页。

10. 为什么绿色植物对于动物重要？

- ☐A 绿色植物消耗食物及氧气.
- ☐B 绿色植物消耗食物及放出氧气.
- ☐C 绿色植物消耗食物及放出二氧化碳.
- ☐D 绿色植物制造食物及放出氧气.
- ☐E 绿色植物制造食物及放出二氧化碳.

11. 以下各种物质具有电解质导电性的是：

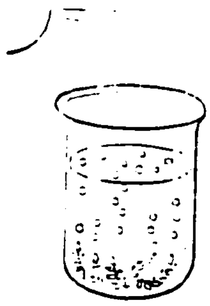
- ☐A 酒精
- ☐B 糖水
- ☐C 铁
- ☐D 石墨
- ☐E 食盐水

12. 血液流经肾小球，经过滤作用形成原尿，原尿流经肾小管，肾小管重吸收了

- ☐A 全部葡萄糖、大部分水分和部分无机盐.
- ☐B 葡萄糖.
- ☐C 水分和无机盐.
- ☐D 部分葡萄糖、水分和无机盐.
- ☐E 全部葡萄糖、水分和无机盐.

现在请将已填妥的答案抄写在答卷纸上，
抄写完毕后，请翻往下一页.

13．根据下图指出气泡中主要含有什么气体？

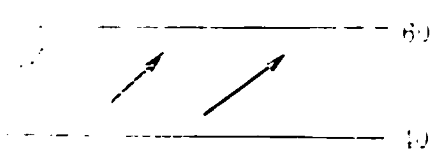


- ＿A 二氧化碳
- ＿B 水蒸汽
- ＿C 氧气
- ＿D 一氧化碳
- ＿E 氮气

14．关于力的概念，下面哪句话是正确的？

- ＿A 力是产生运动的原因。
- ＿B 力是物体间的相互作用，是改变物体运动状态的原因。
- ＿C 运动的物体必然受力的作用。
- ＿D 有力作用于物体上，不一定有施力物体。
- ＿E 力是维持物体运动的原因。

15．看图找出它是下列各风带中的哪个风带？



- ＿A 北半球的西风带
- ＿B 北半球的东风带
- ＿C 北半球的东北信风带
- ＿D 南半球的东风带
- ＿E 南半球的西风带

现在请将已填妥的答案抄写在答卷纸上。

INTERNATIONAL STUDY
OF SCIENCE EDUCATION

数学测验

卷号 **2Q**

中国IEA国家中心
北京

1.
```
  1.27
  2.91
  3.06
+ 8.22
```

和的答案是

□A 14.36

□B 14.46

□C 15.46

□D 27.16

□E 54.06

2.
```
  8265
− 5437
```

差的答案是

□A 2828

□B 2832

□C 3228

□D 3232

□E 3838

3. $2^3 \times 3^2$ 的值是

□A 30

□B 36

□C 64

□D 72

□E 以上都不对

4. 下面哪个数是29×32的最接近的估计值?

□A 600

□B 700

□C 900

□D 1100

5. 下面哪个数是12×75的平方根？

- □A 6.25
- □B 30
- □C 87
- □D 625
- □E 900

6. 多少个七人球队能够组成7个九人球队？

- □A 7
- □B 8
- □C 9
- □D 16
- □E 63

7. 小敏有48块糖果。她自己留了一半，把其余的平均分给她的三个同学，小英，小磊和露露。露露得了多少块糖果？

- □A 16
- □B 12
- □C 8
- □D 6
- □E 1

8. $52.01\div7$等于

- □A 7.04
- □B 7.33
- □C 7.40
- □D 7.43
- □E 7.50

9. 下面哪一个算式是正确的？

☐A 64÷9=8

☐B 0÷7=7

☐C 12÷6=6

☐D 81+9=8

☐E 48+6=8

10. $0.004\overline{)24.56}$

上面除式的正确答案是

☐A 0.614

☐B 6.14

☐C 61.4

☐D 614

☐E 6140

11. 一只蜗牛如果3分钟爬了90厘米，那么一分钟它能爬多少厘米？

☐A 3

☐B 30

☐C 87

☐D 93

☐E 270

12. 一所学校有227个学生。每个学生都要参加一个校内课外小组，音乐小组或体育小组，有的学生可参加两个小组。音乐小组有120人，其中36人也参加了体育小组。参加体育小组的共多少人？

☐A 84

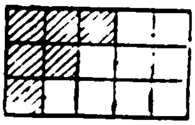
☐B 107

☐C 120

☐D 143

☐E 191

13.



图中所有小方格的大小都是相同的，整个长方形的面积等于1。
阴影部分的面积等于

- □A $\frac{2}{15}$
- □B $\frac{1}{3}$
- □C $\frac{2}{5}$
- □D $\frac{3}{8}$
- □E $\frac{1}{2}$

14. 下面哪个算式是正确的？

- □A $\frac{3}{7}\times\frac{7}{9}=\frac{63}{21}=3$
- □B $\frac{3}{7}\times\frac{7}{9}=\frac{21}{16}=1\frac{5}{16}$
- □C $\frac{3}{7}\times\frac{7}{9}=\frac{10}{16}=\frac{5}{8}$
- □D $\frac{3}{7}\times\frac{7}{9}=\frac{21}{63}=\frac{1}{3}$
- □E $\frac{3}{7}\times\frac{7}{9}=\frac{27}{49}$

15. $\frac{a}{15}-\frac{b}{5}$ 等于

- □A $\frac{a-3b}{15}$
- □B $\frac{5a-15b}{15}$
- □C $\frac{a-b}{10}$
- □D $\frac{a-b}{75}$
- □E $\frac{3a-5b}{75}$

16. 小明要绘制一张条图，来表示四天的最高温度。他制了这张表，来帮助他绘图。

日　　期	星期一	星期二	星期三	星期四
最高温度	16°C	18°C	21°C	24°C
条的高度	8cm	9cm		12cm

星期三条的高度应该是多少？

☐A　9.5cm

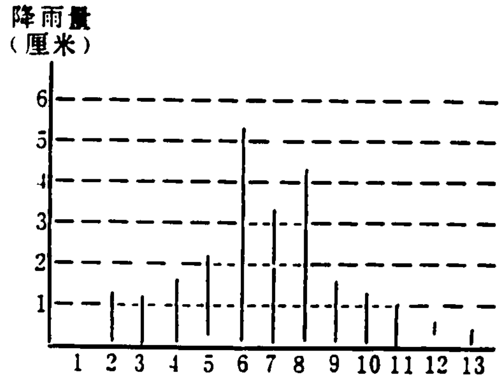
☐B　10cm

☐C　10.5cm

☐D　21cm

☐E　42cm

17. 下图描绘出13周的降雨量（以厘米为单位）。



在这期间每周的平均降雨量接近于

☐A　1cm

☐B　2cm

☐C　3cm

☐D　4cm

☐E　5cm

18． 在下面数列的空白处应该填上什么数，才能使它完整？

2， 4， …， 48， 240

☐A 8

☐B 12

☐C 16

☐D 24

☐E 32

19． 四乘以什么数的积，比这个数大24？

☐A 5

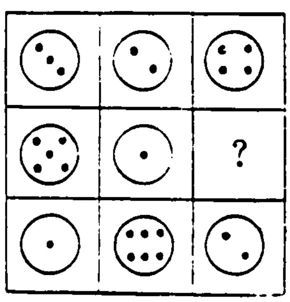
☐B 6

☐C 8

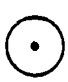
☐D 12

☐E 18

20．



上图中的问号部分，应该填入哪个图，才能使它完整？

☐A 

☐B 

☐C 

☐D 

☐E 

普通中学初中科学
教育评价研究

学生问卷

卷号 **ST**

中国IEA国家中心
北京

这份学生问卷的内容，与你和你的科学课程的学习有关。这不是考试，也不是档案材料。请你不要有顾虑，选择与你的真实情况相符合的答案。

请你把你所选择的答案旁括号内的数字填写在答卷纸上相应题号的方框内。

有些题目中有用(a)、(b)、(c)、(d)等符号标记的多项问题组，要逐一作答。

1. 你属于哪个民族？

(1)汉 (2)蒙 (3)壮 (4)回
(5)满 (6)藏 (7)其它

* 2. 你的性别是什么？

(1)男 (2)女

3. 你进过幼儿园或学前班吗？

(1)没进过 (2)进过1年 (3)进过2年
(4)进过3年 (5)进过4年或以上

4. 你是共青团员吗？

(1)是 (2)不是

5. 你是近视吗？

(1)是 (2)不是

6. 你在中学期间请病假吗？

(1)从不 (2)有时 (3)经常

7. 请指明你上学期期末考试下列各科的成绩属于哪一类。

		0—40分	41—60分	61—75分	76—90分	91—100分
(a)	数 学	(1)	(2)	(3)	(4)	(5)
(b)	物 理	(1)	(2)	(3)	(4)	(5)
(c)	化 学	(1)	(2)	(3)	(4)	(5)
(d)	生 物	(1)	(2)	(3)	(4)	(5)
(e)	语 文	(1)	(2)	(3)	(4)	(5)

8. 你家共有几口人?

请在答卷纸上相应题号旁方框内填上你的答案。

9. 有祖父、祖母、外祖父或外祖母等老人与你们同居吗?

(1)有　　(2)没有

10. 你有多少兄弟姐妹呢?(不计你自己在内)

(1)没有　　(2)一个　　(3)两个　　(4)三个及以上

11. 如有兄弟姐妹，你排行第几?

(1)第一　　(2)第二　　(3)第三　　(4)第四或其他

★ 12. 请指明你父亲和母亲的文化程度。

(a) 父亲的文化程度是

(1)文盲　(2)小学　(3)初中　(4)高中或中专
(5)大学　(6)研究生

(b) 母亲的文化程度是

(1)文盲　(2)小学　(3)初中　(4)高中或中专
(5)大学　(6)研究生

13. 请指明你父亲现在的职业？

(1)教师　(2)科学研究人员　(3)工程技术人员　(4)医务工作者
(5)文化体育工作者　(6)国家机关干部　(7)商业和服务行业工作人员
(8)军人　(9)工人　(10)农民　(11)个体劳动者　(12)其他

14. 请指明你母亲现在的职业？

(1)教师　(2)科学研究人员　(3)工程技术人员　(4)医务工作者
(5)文化体育工作者　(6)国家机关干部　(7)商业和服务行业工作人员
(8)军人　(9)工人　(10)农民　(11)个体劳动者　(12)其他

15. 你家住在下列哪种地区？

(1)市区　(2)县镇　(3)农村　(4)其它

16. 你家的住房有几间。

请在答卷纸上相应题号旁方框内填上你的答案。

17. 你家有没有下列各物品？

		有	无
(a)	电灯	(1)	(2)
(b)	收音机	(1)	(2)
(c)	录音机	(1)	(2)
(d)	电视机	(1)	(2)
(e)	电冰箱	(1)	(2)
(f)	缝纫机	(1)	(2)
(g)	自行车	(1)	(2)
(h)	照像机	(1)	(2)
(i)	计算器	(1)	(2)
(j)	自来水	(1)	(2)
(k)	煤气灶	(1)	(2)
(l)	暖气设备	(1)	(2)
(m)	可供你学习的桌子	(1)	(2)
(n)	洗衣机	(1)	(2)
(o)	电扇	(1)	(2)

18. 你家是否种（养）下列各类植物和动物，用途何在？

		自己家消费、观赏	卖	我家不种（养）
(a)	蔬　菜	(1)	(2)	(3)
(b)	花　木	(1)	(2)	(3)
(c)	家　禽	(1)	(2)	(3)
(d)	家　畜	(1)	(2)	(3)

* 19. 你家共计约有多少本书（不包括杂志、报纸等）？

(1)没有或很少　(2)（20—50本）
(3)(51—100本)　(4)（100—500本）　(5)500本以上

20. 你家的书籍主要是什么种类？

(1)科技类　(2)文学类　(3)艺术类　(4)政治类

21. 你家订阅报纸或杂志吗？

		按期订阅	有时购买	很少买或从不买
(a)	报纸	(1)	(2)	(3)
(b)	杂志	(1)	(2)	(3)

22. 你家订阅的报纸或杂志中有专门为你订的吗？

(1) 有　(2) 无

23. 你家里有一个安静的学习环境吗？

(1)有　(2)无　(3)一般

24. 你父母对你上学是什么态度？

(1)反对　(2)不太支持　(3)一般　(4)支持

25. 请注明你父亲希望你将来从事什么工作？

(1)科学家　(2)国家干部　(3)医生　(4 农民　(5)工人　(6)教师
(7)商业服务员　(8)个体劳动者　(9)文化体育工作者　(10)军人

26. 请注明你母亲希望你将来从事什么工作？

(1)科学家　(2)国家干部　(3)医生　(4)农民　(5)工人　(6)教师
(7)商业服务员　(8)个体劳动者　(9)文化体育工作者　(10)军人

27. 你是否常去下列各种地方？

		经常去	有时去	很少去或从不去
(a)	图书馆	(1)	(2)	(3)
(b)	少年宫	(1)	(2)	(3)
(c)	博物馆	(1)	(2)	(3)
(d)	影剧院	(1)	(2)	(3)
(e)	体育场	(1)	(2)	(3)
(f)	公园	(1)	(2)	(3)
(g)	同学或亲属家	(1)	(2)	(3)
(h)	书店	(1)	(2)	(3)
(i)	郊游、野餐	(1)	(2)	(3)

* 28. 请指出你每周大约用多少课外时间做所有学科的作业。

(1)不到1小时　(2)1—4小时　(3)5—9小时
(4)10—14小时　(5)14小时以上

* 29. 请指出你每周大约用多少课外时间做所有科学课程的作业。

(1)不到1小时　(2)1—4小时　(3)5—9小时
(4)10—14小时　(5)14小时以上

30. 请指出你每周大约用多少时间干家务活或参加生产劳动。

(1)不到1小时　(2)1—4小时　(3)5—9小时
(4)10小时或以上　(5)我不干家务或不参加生产劳动

31. 请指出你每周大约用多少时间参加下列各项课外活动。

		不到1小时	1—4小时	5—9小时	10小时或以上	不参加
(a)	体育锻炼	(1)	(2)	(3)	(4)	(5)
(b)	自然学科兴趣小组	(1)	(2)	(3)	(4)	(5)
(c)	社会学科兴趣小组	(1)	(2)	(3)	(4)	(5)
(d)	无线电小组	(1)	(2)	(3)	(4)	(5)
(e)	航空(海)模型小组	(1)	(2)	(3)	(4)	(5)
(f)	文艺社团	(1)	(2)	(3)	(4)	(5)
(g)	集邮	(1)	(2)	(3)	(4)	(5)
(h)	听广播	(1)	(2)	(3)	(4)	(5)
(i)	看电视	(1)	(2)	(3)	(4)	(5)
(j)	校办工厂、农场	(1)	(2)	(3)	(4)	(5)

32. 你的老师在课外时间同你一起参加活动吗?

		经常	有时	很少	从不
(a)	物理	(1)	(2)	(3)	(4)
(b)	化学	(1)	(2)	(3)	(4)
(c)	生物	(1)	(2)	(3)	(4)
(d)	地理	(1)	(2)	(3)	(4)

33. 你的家长辅导你学习吗?

(1)从不辅导 (2)每天约1小时 (3)每天2小时 (4)每天2小时以上

34. 你每天睡眠时间平均大约多少小时?

(1) 9小时 (2) 8小时 (3) 7小时 (4) 6小时 (5) 5小时

35. 你学习中最喜欢的是哪一学科？

(1)物理　(2)化学　(3)生物　(4)地理　(5)没有最喜欢的

36. 你学习中感到最困难的是哪一学科？

(1)物理　(2)化学　(3)生物　(4)地理　(5)没有最困难的

37. 你对考试成绩采取什么态度？

(1)得多少分无所谓　(2)比较想得高分　(3)非常想得高分

38. 你住校吗？

(1)住校　(2)不住校

39. 请指出从你家到学校的距离

(1)不到1公里　(2)1—2公里　(3)3—4公里　(4)4公里以上

40. 你平时怎样去学校？

(1)步行　(2)骑自行车　(3)乘公共汽车　(4)其他

41. 你到学校单程需要多少分钟？

请在答卷纸上相应题号旁方框内填上你的答案

42. 你们班里有图书角吗？

(1)有　(2)没有

43. 你希望将来最终接受何种教育？

(1)初中毕业　(2)高中、中专或技校　(3)大学
(4)研究生　(5)自学成才

44. 你自己喜欢将来从事何种职业？（只选一种）

(1)科学家、工程师　(2)教师　(3)工人　(4)农民　(5)军人
(6)医务人员　(7)国家干部　(8)商业服务员　(9)个体劳动者
(10)文艺、体育工作者

答题指示

这份问卷的目的，是让学生讲出自己对科学、科学课程及就读学校的看法。这不是一项测验，答案无所谓对和错。

在下列各题中，请在答卷纸上相应题号旁的方框内填上你所选择的答案下括号内的数字，以代表你的意见。

(1) 表示你同意。

(2) 表示你不同意。

(3) 表示你不肯定或不知道。

		同意	不同意	不肯定
1.	科学实验室中的工作并不是一种有趣的工作。	(1)	(2)	(3)
2.	科学对解决日常生活中的问题有用处。	(1)	(2)	(3)
3.	上学督促人奋进。	(1)	(2)	(3)
4.	科学破坏了自然环境。	(1)	(2)	(3)
* 5.	如果教得适当学生都可以学好科学课。	(1)	(2)	(3)
* 6.	将来大多数的工作都需要科学知识。	(1)	(2)	(3)
7.	科学对一个国家的发展非常重要。	(1)	(2)	(3)
8.	上学是很愉快的。	(1)	(2)	(3)
9.	把钱花在科学上是很值得的。	(1)	(2)	(3)
10.	科学课程在学校里是饶有趣味的。	(1)	(2)	(3)
11.	要获得一份好工作，懂得科学并不是重要的。	(1)	(2)	(3)
12.	现代社会中许多问题都是由科学引起的。	(1)	(2)	(3)
13.	我喜欢我们的学校。	(1)	(2)	(3)
14.	在我们的社会中懂得科学的人生活得较为充实。	(1)	(2)	(3)
15.	学校中所教授的科学是有趣的。	(1)	(2)	(3)
16.	富于创造性的人更适合从事科学研究。	(1)	(2)	(3)
17.	过去几年花在科学上的公款是用得不明智的。	(1)	(2)	(3)
18.	在学校里的大部分时间我都觉得沉闷。	(1)	(2)	(3)
19.	科学发明改善我们的生活。	(1)	(2)	(3)
20.	科学课是深奥的。	(1)	(2)	(3)

	同意	不同意	不肯定
21. 我希望在将来的工作中能用到在学校所学的科学知识.	(1)	(2)	(3)
22. 政府应该在科学研究方面少花些钱.	(1)	(2)	(3)
23. 学校中许多科目我都不喜欢.	(1)	(2)	(3)
24. 科学发明使世界变得太复杂了.	(1)	(2)	(3)
25. 科学课在涉及计算时就显得深奥.	(1)	(2)	(3)
26. 我希望尽量多接受些教育.	(1)	(2)	(3)
27. 科学发明使人与人之间的关系更加紧张.	(1)	(2)	(3)
28. 我生活中最愉快的光阴是在学校中渡过的.	(1)	(2)	(3)
29. 科学使未来世界变得更美好.	(1)	(2)	(3)
30. 科学课在涉及仪器操作时就显得深奥.	(1)	(2)	(3)
31. 我不打算在离开学校后成为--名担任科学课程的教师.	(1)	(2)	(3)
32. 科学发现弊多利少.	(1)	(2)	(3)
33. 我通常不喜欢学校的功课.	(1)	(2)	(3)
* 34. 利用现代发明（如电脑）的工作是更富趣味的工作.	(1)	(2)	(3)
* 35. 科学课中可以学到很多东西.	(1)	(2)	(3)
* 36. 科学家对怎样应用他们的发现是有责任的.	(1)	(2)	(3)
37. 科学与技术是世界上许多问题的起因.	(1)	(2)	(3)
38. 在未来五年间，我们国家的形势可能会变得更好.	(1)	(2)	(3)
39. 我希望多了解些我们生活的世界.	(1)	(2)	(3)
40. 科学与日常生活有密切的关系.	(1)	(2)	(3)

答题指示

这一部分的各题均与学生学习科学时的活动有关。

请在答卷纸上相应题号旁的方框里填上你所选择的答案下括号内的数字，以表示学习科学时所出现的情况。

（1）表示经常出现的

（2）表示有时会出现的

（3）表示从未出现的

		经常	有时	从不
* 1.	上科学课时，我们用课本。	(1)	(2)	(3)
* 2.	学习科学课时，我们用课本以外的书籍。	(1)	(2)	(3)
3.	每一科学课开始时，教师与我们重温以上几课所学习过的东西。	(1)	(2)	(3)
4.	每一科学课开始时，教师说明这一课要学习的东西。	(1)	(2)	(3)
5.	每一科学课结束时，教师总结该课教过的东西。	(1)	(2)	(3)
6.	教师准许我们自己选择科学课题来学习。	(1)	(2)	(3)
7.	教师在计划科学课时，采纳我们的意见和建议。	(1)	(2)	(3)
* 8.	教师运用示范来帮助解释科学概念。	(1)	(2)	(3)
9.	教师把科学课教得很有趣。	(1)	(2)	(3)
10.	上科学课时，我们把教师的笔记从黑板抄到我们自己的本子上。	(1)	(2)	(3)
11.	我们把实验或其他活动的报告写出，作为科学课作业。	(1)	(2)	(3)
12.	教师解释我们所学的科学知识如何与生活有关。	(1)	(2)	(3)
13.	教师与我们讨论将来在科学领域内可能从事的职业。	(1)	(2)	(3)
* 14.	科学课经常都有测验。	(1)	(2)	(3)
15.	科学课教师帮助学习科学有困难的学生。	(1)	(2)	(3)

	经常	有时	从不
16. 上科学课时我们使用计算器.	(1)	(2)	(3)
17. 我们到教室以外进行实地学习作为科学课的一部分.	(1)	(2)	(3)
* 18. 实验是我们科学课的一部分.	(1)	(2)	(3)
19. 科学课作实验的时候，学生分成小组.	(1)	(2)	(3)
20. 做实验的时候，教师指示我们怎样做.	(1)	(2)	(3)
21. 做实验的时候．我们依照书本或教师的指示进行	(1)	(2)	(3)
22. 做实验时，我们自己提出疑问，然后教师帮助我们设计实验去解答.	(1)	(2)	(3)
23. 做实验时，教师提出问题，让我们自己去找出解决的方法和答案.	(1)	(2)	(3)
24. 做实验时，我们提出问题，自己找出方法去研究问题.	(1)	(2)	(3)

答题指示：第二部分

这一部分的问题与学习科学有关。每一题有五个可供选择的答案，分别用(1)、(2)、(3)、(4)及(5)来表示。

请把你所选择的答案括号内的数字填写在答案纸上相应题号的方框内，以表示学习科学时所出现的情况。

25. 上课时教师向你发问，你会怎样做？

(1)教师从不向我发问　(2)我总是尽我所能作答
(3)我通常会尝试作答　(4)我通常不会尝试作答　(5)我从不尝试作答

26. 当教师在班上向另一个同学发问时，你会怎样做？

(1)上课时教师从不发问　(2)我总是尽力暗自作答　(3)我通常会尝试暗自作答
(4)我通常不会尝试暗自作答　(5)我从不尝试暗自作答

27. 当教师介绍一些新事物的时候，你注意吗？

(1)我总是尽量注意　(2)我通常会注意　(3)我有时会注意，有时不注意
(4)我通常不会注意　(5)我从不注意

28. 你课上有问题时，会怎样做？

(1)我从来没有问题　(2)我总是向老师提问　(3)我通常向老师提问
(4)我通常不向老师提问　(5)我从不向老师提问

29. 当你做作业而感到需要帮助时，你会找家里人帮忙吗？

(1)我从没有作业　(2)我从不需求助　(3)我通常找家里人帮忙
(4)我通常不找家里人帮忙　(5)我从不做作业

30．当你做作业而感到需要帮助时，你会找同学帮忙吗？

(1)我从没有作业　(2)我从不needs求助　(3)我通常找同学帮忙
(4)我通常不找同学帮忙　(5)我从不做作业

31．你帮助同学做作业吗？

(1)从不帮助　(2)没有人要我帮助　(3)有时帮助
(4)经常帮助　(5)我们几乎没有作业

32．当你想认识一些事物时，你查阅课外书籍吗？

(1)我差不多每天都查阅课外书籍　(2)我常常查阅课外书籍
(3)我有时会查阅课外书籍　(3)我很少查阅课外书籍
(5)我从不查阅课外书籍

33．你曾否阅读课本中教师没有指定的阅读内容？

(1)我没有课本　(2)我经常阅读教师没有指定要看的部份
(3)我有时会阅读教师没有指定要看的部份
(4)我很少阅读教师没有指定要看的部份
(5)我从不阅读教师没有指定要看的部份

34．你做完作业后，会核对一下吗？

(1)我从没有作业　(2)我通常会核对的　(3)我有时会核对的
(4)我通常不会核对的　(5)我从不做作业

35．做作业时，你是否尽力而为？

(1)我从没有作业　(2)我总是尽力而为　(3)我通常会尽力而为
(4)我通常不会太努力　(5)我从不做作业

36. 做家庭作业时，你的周围经常有噪音和其他干扰吗？

(1)从没有　(2)很少有　(3)有时有

(4)经常有　(5)我从不做家庭作业

37. 当有家庭作业时，你会把它全部完成吗？

(1)我从没有家庭作业　(2)我会全部完成　(3)我会完成大部份，留下小部份

(4)我会完成小部份，留下大部份　(5)我从不做家庭作业

38. 当你要交家庭作业的时候，你会准时交吗？

(1)我从没有家庭作业　(2)我通常会准时交　(3)我有时会准时交

(4)我通常不会准时交　(5)我从不做家庭作业

39. 你缺课后，能把所缺课程补上吗？

(1)我从不缺课　(2)我缺课后总能补上　(3)我缺课后通常能补上

(4)我缺课后通常不补上　(5)我缺课后从来不补

40. 你每周在校外接受科学课程补习的时间有多少？

(1)我没有接受科学课程的补习　(2)每周少于半小时科学课程的补习

(3)每周半小时至一小时科学课程的补习　(4)每周超过一小时科学课程的补习

INTERNATIONAL STUDY
OF SCIENCE EDUCATION

学生学习机会评估

卷 号

2OTL

中国IEA国家中心
北京

说　明

1. “学生学习机会评估”目的在于评估你校的入样学生在何种程度上学习过这次测验题目所涉及的科学概念。要进行两种评估，即评估1：学习机会百分比，即占本学年教学时间的百分比和评估2：年级水平，即学生在哪一年级学过这一内容。

2. 评估时请注意以下事项：

2.1 评估表要与实际测验题目对照使用

即：测验2M（题号1—30）

测验2A（题号1—15）

测验2B（题号1—15）

测验2C（题号1—15）

测验2D（题号1—15）

2.2 先仔细研读测验题目，然后再确定评估1和评估2的等级。

2.3 在每一题（题目编号2A12即指测验2A中的第12题）后的四个字母中选择一个用铅笔圈起来。

2.4 学校联络员可组织有关教师完成学生学习机会评估。

3. 请注意对每一测验题都要进行评估。

题目编号	评估1：学习机会百分比（占本学年数学时间的百分比）				评估2：年级水平			
	76-100%	26-75%	1-35%	0%	初二或更早	初三	高一或更晚	不学
2M01	A	B	C	D	A	B	C	D
2M02	A	B	C	D	A	B	C	D
2M03	A	B	C	D	A	B	C	D
2M04	A	B	C	D	A	B	C	D
2M05	A	B	C	D	A	B	C	D
2M06	A	B	C	D	A	B	C	D
2M07	A	B	C	D	A	B	C	D
2M08	A	B	C	D	A	B	C	D
2M09	A	B	C	D	A	B	C	D
2M10	A	B	C	D	A	B	C	D
2M11	A	B	C	D	A	B	C	D
2M12	A	B	C	D	A	B	C	D
2M13	A	B	C	D	A	B	C	D
2M14	A	B	C	D	A	B	C	D
2M15	A	B	C	D	A	B	C	D
2M16	A	B	C	D	A	B	C	D
2M17	A	B	C	D	A	B	C	D
2M18	A	B	C	D	A	B	C	D
2M19	A	B	C	D	A	B	C	D
2M20	A	B	C	D	A	B	C	D
2M21	A	B	C	D	A	B	C	D
2M22	A	B	C	D	A	B	C	D
2M23	A	B	C	D	A	B	C	D
2M24	A	B	C	D	A	B	C	D
2M25	A	B	C	D	A	B	C	D
2M26	A	B	C	D	A	B	C	D
2M27	A	B	C	D	A	B	C	D
2M28	A	B	C	D	A	B	C	D
2M29	A	B	C	D	A	B	C	D
2M30	A	B	C	D	A	B	C	D

题目编号	评估1：学习机会百分比（占本学年数学时间的百分比）76-100%	26-75%	1-25%	0%	评估2：年级水平 初二或更早	初三	高一或更晚	不学
2A01	A	B	C	D	A	B	C	D
2A02	A	B	C	D	A	B	C	D
2A03	A	B	C	D	A	B	C	D
2A04	A	B	C	D	A	B	C	D
2A05	A	B	C	D	A	B	C	D
2A06	A	B	C	D	A	B	C	D
2A07	A	B	C	D	A	B	C	D
2A08	A	B	C	D	A	B	C	D
2A09	A	B	C	D	A	B	C	D
2A10	A	B	C	D	A	B	C	D
2A11	A	B	C	D	A	B	C	D
2A12	A	B	C	D	A	B	C	D
2A13	A	B	C	D	A	B	C	D
2A14	A	B	C	D	A	B	C	D
2A15	A	B	C	D	A	B	C	D
2B01	A	B	C	D	A	B	C	D
2B02	A	B	C	D	A	B	C	D
2B03	A	B	C	D	A	B	C	D
2B04	A	B	C	D	A	B	C	D
2B05	A	B	C	D	A	B	C	D
2B06	A	B	C	D	A	B	C	D
2B07	A	B	C	D	A	B	C	D
2B08	A	B	C	D	A	B	C	D
2B09	A	B	C	D	A	B	C	D
2B10	A	B	C	D	A	B	C	D
2B11	A	B	C	D	A	B	C	D
2B12	A	B	C	D	A	B	C	D
2B13	A	B	C	D	A	B	C	D
2B14	A	B	C	D	A	B	C	D
2B15	A	B	C	D	A	B	C	D

题目编号	评估1：学习机会百分比（占本学年教学时间的百分比）76-100%	26-75%	1-25%	0%	评估2：年级水平 初二或更晚	初三	高一或更早	不学
2C01	A	B	C	D	A	B	C	D
2C02	A	B	C	D	A	B	C	D
2C03	A	B	C	D	A	B	C	D
2C04	A	B	C	D	A	B	C	D
2C05	A	B	C	D	A	B	C	D
2C06	A	B	C	D	A	B	C	D
2C07	A	B	C	D	A	B	C	D
2C08	A	B	C	D	A	B	C	D
2C09	A	B	C	D	A	B	C	D
2C10	A	B	C	D	A	B	C	D
2C11	A	B	C	D	A	B	C	D
2C12	A	B	C	D	A	B	C	D
2C13	A	B	C	D	A	B	C	D
2C14	A	B	C	D	A	B	C	D
2C15	A	B	C	D	A	B	C	D
2D01	A	B	C	D	A	B	C	D
2D02	A	B	C	D	A	B	C	D
2D03	A	B	C	D	A	B	C	D
2D04	A	B	C	D	A	B	C	D
2D05	A	B	C	D	A	B	C	D
2D06	A	B	C	D	A	B	C	D
2D07	A	B	C	D	A	B	C	D
2D08	A	B	C	D	A	B	C	D
2D09	A	B	C	D	A	B	C	D
2D10	A	B	C	D	A	B	C	D
2D11	A	B	C	D	A	B	C	D
2D12	A	B	C	D	A	B	C	D
2D13	A	B	C	D	A	B	C	D
2D14	A	B	C	D	A	B	C	D
2D15	A	B	C	D	A	B	C	D

普通中学初中科学
教育评价研究

教师问卷

TE

中国IEA国家中心
北京

普通中学初中科学教育评价研究是一项在内蒙古、辽宁、江苏、湖北、广西、四川、甘肃七省、自治区之间的合作研究，其目的是探讨普通中学初中课程、学生、教师及学校特点等因素与学生科学课程（包括物理、化学、生物、地理）成绩的关系。

你校部分学生将被选出参加测验及回答问卷，作为研究的一部分。我们请你帮助回答的这份问卷是以"第二次国际科学研究"总体2所用的教师问卷为蓝本，结合我国当前的教育实际修订而成的。请在卷内每题题号旁的方框内填上你所选择的答案旁括号内的数字，有些填空题目，请在该题内的方框里填上你的答案。

这项研究必须依当教师所提供的正确资料，因此请你尽可能准确地填写这份问卷，所填写的一切资料绝对保密。

对你的合作及帮助，谨此致谢。

学校名称：______________________

教师姓名：______________

教师识别编码：__________

□1. 你的性别是.

(1)男　(2)女

2. 你出生于＿＿＿年＿＿＿月

□3. 你的学历是什么?

(1)初中及以下　(2)高中或中专　(3)高等专科学校　(4)大学及以上

□4. 你受过何种师范教育?

(1)没受过师范教育　(2)师范性的在职培训　(3)中等师范教育　(4)高等师范教育

□5. 你在高等学校所学的专业是什么?

(1)物理　(2)化学　(3)生物　(4)地理　(5)其他

6. 你的教龄＿＿＿年

7. 你教科学课程＿＿＿年

□8. 如可能，你还希望受多少年师范教育?

(1)1年　(2)2—3年　(3)4年　(4)没有这种必要

□9. 你的健康状况怎样?

(1)健康　(2)有不明显影响教学的疾病　(3)有明显影响教学的疾病

□10. 你有病时将怎样?

(1)及时到医院治疗　(2)无时间到医院治疗　(3)无力到医院治疗

□11. 你家人均住房面积是多少?

(1)5平方米及以下　(2)6—10平方米　(3)11—15平方米　(4)16平方米及以上

12. 1987年你家全年人均收入＿＿＿元

□13. 你每天用于家务的时间是多少?

(1)2小时以内　(2)2—3小时　(3)4—5小时　(4)5小时以上

□14. 你主要教哪一学科?

(1)物理　(2)化学　(3)生物　(4)地理　(5)其他

□15. 你是否兼教下列课程？

(1)物理　(2)化学　(3)生物　(4)地理　(5)其他

16. 你每周授课____课时

17. 你每周用于备课____小时

18. 你每周用于批改作业____小时

19. 你每周用于辅导学生学习____小时

□20. 你是否兼管实验室工作？

(1)是　(2)否

□21. 在你的教学中，教科书上规定的演示实验能做多少？

(1)基本上不能做　(2)20%及以下　(3)21—50%　(4)51—80%　(5)80%以上

□22. 你在教学中使用教学挂图吗？

(1)从不　(2)很少　(3)有时　(4)经常

□23. 你在教学中使用模型标本吗？

(1)从不　(2)很少　(3)有时　(4)经常

□24. 你在教学中使用幻灯机、投影仪吗？

(1)从不　(2)很少　(3)有时　(4)经常

□25. 你在教学中使用教学电影吗？

(1)从不　(2)很少　(3)有时　(4)经常

□26. 你怎样要求学生完成你布置的作业？

(1)尽量在课堂内完成　(2)大部分在课堂内完成　(3)基本上在课外完成

□27. 你怎样批改学生的作业？

(1)抽批学生互批的作业　(2)全批(全部划出对或错)不改　(3)全批粗改　(4)全批细改

□28. 你用作业评定来评定学生的学习吗？

(1)从不　(2)　很少　(3)有时　(4)经常

☐29. 你用平时测验来评定学生的学习吗？

(1)从不　(2)很少　(3)有时　(4)经常

☐30. 你用课堂提问来评定学生的学习吗？

(1)从不　(2)很少　(3)有时　(4)经常

☐31. 你用定期考试来评定学生的学习吗？

(1)从不　(2)很少　(3)有时　(4)经常

☐32. 你组织学生参观学习吗？

(1)从不　(2)很少　(3)有时　(4)经常

☐33. 你组织学生现场考察吗？

(1)从不　(2)很少　(3)有时　(4)经常

☐34. 你组织学生看演出吗？

(1)从不　(2)很少　(3)有时　(4)经常

☐35. 你带领学生一起自制仪器、教具吗？

(1)从不　(2)很少　(3)有时　(4)经常

☐36. 你是否参加教学研究活动，如听课、专题讨论等？

(1)从不参加　(2)每学期参加5次以内　(3)每学期参加6—10次

(4)每学期参加10次以上

☐37. 你是否阅读与数学有关的学报或期刊？

(1)从不阅读　(2)很少阅读　(3)有时阅读　(4)经常阅读

☐38. 你在什么报刊上发表过与教学有关的文章？

(1)校级　(2)县级　(3)市级　(4)省级以上

☐39. 你对你所教的现行教材有何看法？

(1)内容太浅　(2)内容适当　(3)内容太深

☐40. 请估计一下你所教的学生中爱学习的学生比例。

(1)20%及以下　(2)21—40%　(3)41—60%　(4)61—80%　(5)80% 以上

普通中学初中科学
教育评价研究

学校问卷

SC

中国IEA国家中心
北京

普通中学初中科学教育评价研究是一项在内蒙古、辽宁、江苏、湖北、广西、四川、甘肃七省、自治区之间的合作研究，其目的是探讨普通中学初中课程、学生、教师及学校特点等因素与学生科学课程（包括物理、化学、生物、地理）成绩的关系。

你校部分学生将被选出参加测验及回答问卷，作为研究的一部分。我们请你帮助回答的这份问题是以“第二次国际科学研究”总体2的学校问卷为蓝本，结合我国当前的教育实际修订而成的，请在卷内每题题号旁的方框内填上你所选择的答案旁括号内的数字。有些填空题目，请在该题内的方框里填上你的答案。

这项研究必须依赖校长所提供的正确资料，因此请你尽可能准确地填写这份问卷。所填写的一切资料绝对保密。

对你的合作及帮助，谨此致谢。

学校编码：________________

学校名称：______________________

校长姓名：__________

□1. 你的性别是

(1)男　(2)女

2. 你出生于[　　]年[　]月

□3. 你的学历是

(1)初中及以下　(2)高中或中专　(3)高等专科学校　(4)大学及以上

□4. 你受过何种师范教育？

(1)没受过师范教育　(2)师范性的在职培训　(3)中等师范教育　(4)高等师范教育

5. 你担任校长[　　]年

6. 你的教龄[　　]年

7. 你兼教[　　]课，每周[　　]课时

8. 你校自开办至今有[　　]年历史

9. 请填写你校1987—1988学年度各年级的班数和男、女学生数

年级	班数	学生数	
		男	女
初一			
初二			
初三			
高一			
高二			
高三			

10. 本学年度初三学生从小学升入你校初中一年级时，录取的最低分数是[　　]分，最高是[　　]分，平均是[　　]分

11. 你校共有男教师[ ]名，其中民办教师[ ]名

12. 你校共有女教师[ ]名，其中民办教师[ ]名

13. 你校初中科学课程教师的学历

学科 / 性别 / 人数 / 学历	物理		化学		生物		地理	
	男	女	男	女	男	女	男	女
初中及以下								
高中或中专								
高等专科学校								
大学及以上								

14. 你校初中科学课程教师的级别

级别 / 人数 / 学科	中教6级及以下	中教3—5级	中教2级及以下
物理			
化学			
生物			
地理			

15. 你校初中科学课程教师如已评出职务职称，请再填写下表

人数＼职称 / 学科	三级教师	二级教师	一级教师	高级教师
物理				
化学				
生物				
地理				

16. 你校图书馆（室）藏书[ ]册

☐17. 你校专辟有教师阅览室吗？

(1)有 (2)无

☐18. 你校专辟有学生阅览室吗？

(1)有 (2)无

19. 你校下列学科各有实验室（即学生分组实验的教室）几个？其面积多大？

物理[ ]个，[ ]平方米

化学[ ]个，[ ]平方米

生物[ ]个，[ ]平方米

地理[ ]个，[ ]平方米

20. 你校各科教学仪器贮藏室共约[ ]平方米

21. 你校实验室的专职工作人员共[ ]人

22. 请估计你校下列学科的实验室平均每周使用多少课时

物理[ ]课时

化学[　　　]课时

生物[　　　]课时

地理[　　　]课时

23　请估计你校初中学生参加课外科技活动的人数百分比

(1)0%　(2)1—25%　(3)26—50%　(4)51—75%　(5)75%以上.

24.　请估计你校初中学生参加课外文娱活动的人数百分比

(1)0%　(2)1—25%　(3)26—50%　(4)51—75%　(5)75%

25.　请估计你校初中学生参加课外体育活动的人数百分比

(1)0%　(2)1—25%　(3)26—50%　(4)51—75%　(5)75%以上

26.　你校校办工厂的情况怎样?（无校办工厂的免答）

固定资产约合[　　　]元

年产值[　　　]元

年利润[　　　]元

27.　你校本学年度的经费（不包括教职工的工资）来源:

政府拨款[　　　]元

集　　资[　　　]元

校办工厂提成[　　　]元

其他来源[　　　]元

28.　你校本学年度的经费开支:

基本建设[　　　]元

图书资料费[　　　]元

添置仪器和购买实验消费品等的费用[　　　]元

学校福利费[　　　]元

ABSTRACT

Appropriate assessment of students' science achievement is a fundamental question in science education. One statistical approach to assessment suggests the establishment of a prediction model. Yet, no prediction model is uniformly supported by theories. The research presented in this dissertation explores a possible empirical model for prediction of students' science achievement in China and the United States. Construction of the model is based on the ninth grade data sets from the Phase .: of the Second IEA Science Study (SISS) in the United States and the SISS Extension Study (SES) in Hubei province of China.

Previous research divides prediction models into linear vs. non-linear categories. However, as an empirical exploration, neither linear nor non-linear relations should be imposed as a pre-condition of the model construction. In this research, both linear and non-linear functions are treated as special cases of a Taylor polynomial series. The shrinkage method favored by Copas (1983) and Hebel, et. al. (1993) is employed to construct the polynomial coefficients in the truncated Taylor model. The common variables observed in the SES and Phase II SISS projects are classified into five categories, students' gender, attitudes, home background, classroom experience, and personal effort, based on the distinction of visible and latent characteristics and the scree plots from principal component analyses. The latent categories, students' attitudes, home background, classroom experience, and personal effort, are represented by their first principal components. The factors of prediction are constructed by polynomials of the visible variable (gender), the latent principal components, and their interactions. Significant factors are selected through the backward elimination procedure in SAS.

Factor structures are expressed by factor loadings in each category. The differences in the factor structure and the model complexity between the United States and China are interpreted in terms of the differing educational, political, social and cultural contexts in each country. The empirical results are: 1) Gender has a significant linear effect on students' science achievement; 2) The effects of attitude, home background, classroom experience, and personal effort, are curvilinear. Curvature functions are derived for each factor to elaborate the

curvilinearities. 3) In both countries, most significant interactions are at the third polynomial level.