A Comparison of Automated Scoring Engines and Human Raters on the Assessment of English Essay Writing

Thesis submitted by

CHAN Kin Yee

M.A.

in February 2012

for the degree of Doctor of Philosophy

in the School of Education

James Cook University

STATEMENT OF ACCESS

I, the undersigned, author of this work, understand that James Cook University will make this thesis available for use within the university library and, via the Australian Digital These network, for use elsewhere.

I understand that, as an unpublished work, a thesis has significant protection under the Copyright Act; and I do not wish to place any further restriction on access to this work.

19 February 2012

Signature ______________________ Date ______________________

DECLARATION ON SOURCES

I declare that this thesis is my own work and has not been submitted in any form for another degree or diploma at any university or other institution of tertiary education. Information derived from the published or unpublished work of others has been acknowledged in the text and a list of references is given.

19 February 2012

Signature   Date

STATEMENT OF SOURCES

ELECTRONIC COPY

I, the undersigned, the author of this work, declare that the electronic copy of this thesis provided to the James Cook University Library is an accurate copy of the print thesis submitted, within the limits of the technology available.

19 February 2012

Signature   Date

CONTRIBUTIONS OF OTHERS

Financial Support:

School of Education, JCU: Postgraduate Research Student Support Funds
Faculty of Arts, Education and Social Sciences, JCU:
Graduate Research Scheme Funds, Internal Research Awards

Editorial Support

Thanks to Professor Trevor G. Bond for providing intelligent feedback and advice on the research design, data analysis and thesis writing. He was consistently available for consultation by email, video-conferencing or face-to-face supervision meeting. It made my off-campus doctoral study much easier.

Dr Helen Boon provided timely responses and insightful advice on the thesis skeleton and the more completed chapters. Thanks to her for providing me with fully support throughout my doctoral journey.

DECLARATION ON ETHICS

The research presented and reported in this thesis was conducted within the guidelines for research ethics outlined in the National Statement on Ethics Conduct in Research Involving Human (1999), the Joint NHMRC/AVCC Statement and Guidelines on Research Practice (1997), the James Cook University Policy on Experimentation Ethics, Standard Practices and Guidelines (2001) and the James Cook University Statement and Guidelines on Research Practice (2001). The proposed research methodology received clearance from the James Cook University Experimentation Ethics Review Committee (approval number: H3477).

19 February 2012

Signature Date

ACKNOWLEDGEMENTS

In the winter of 2008, I had an opportunity to start my doctoral study.

Prof Trevor G. Bond and Dr Helen Boon played significant roles in it. They were thoughtful and remarkably helpful to me during this research. I would like to express my sincere gratitude to Prof Bond for his superb guidance through the entire research process. He provided me with a package of rigorous training and invaluable advice on problems I encountered in this three-year study. This research was finished under the auspices of Prof Bond. Dr Boon always provided prompt feedback and advice, and her kind support had been of great value in this research.

Special thanks to the MetaMetrics Inc. for access to the data containing the US student essay writing. Thank you, also, to Hal Burdick, Dr Carl Swartz and the technology team at the MetaMetrics Inc. for their ongoing support and help. In particular, I would like to give my heartfelt thanks to the Chairman and CEO of the MetaMetrics Inc., Dr A. Jackson Stenner, for his timely, precise and concise words that were always greatly encouraging to me.

Without the help of Dr Mike Linacre, I could not have finished this thesis. He was patient in providing advice and suggestions for the MFRM analyses in this research. His prompt replies erased the time difference between Hong Kong and the United States.

I also sincerely thank the principals in my Hong Kong partner school, Principals Mimi Lai Yung Ng and Lam Yuet Har, for the permission to access data and their enthusiastic and professional English teachers, Anna Poon, Angela Lai, Angie Lo and Jamie Yau, for scoring more than 300 student essays voluntarily for this research.

Thank you for the unwavering support from Prof Jerry Chan and his wife Kennis Leung, Prof Ling-po Shiu, Josephine Chan, Whitney Kwong and Esther Shaw at every stage of my study. Without their encouragement, I am sure I would have gotten lost during these years.

Finally, I would like to thank my family. Thank you to my mother and father, and my three elder brothers - in particular Kin Wai, my little brother - for supporting me in achieving this goal. Their love and support provides me with the bliss of doing whatever I want in my life.

I dedicate this research to my beloved father
- I am truly proud to be your daughter.

ABSTRACT

Essay scoring operates both in the classroom and in high-stakes testing and the results of essay scoring in high-stakes assessment impact on the students' academic development. Thus, teachers, students and parents are under considerable pressure in the educational system in Hong Kong.

This research investigates how effective a new Automated Essay Scoring (AES) system, the Lexile Analyzer, is in measuring essay writing ability as scored by trained, professional human raters. To do this, an estimation of rater, genre, student and rubric effects must be undertaken. Then, the estimates from the recently developed the Lexile Analyzer is modeled against those of Intelligent Essay Assessor (IEA), an AES widely used in high-stakes testing in the United States. Finally, a concurrent co-calibration of the scores of human raters, the Lexile Analyzer and IEA needs to be undertaken.

Automated Essay Scoring (AES) employs computer technology to evaluate and score written prose in the place of the usual human grading. It has become more widely accepted as an educational technology for both assessment and classroom instruction, but little research has been done on validating the performance of the AES models in an applied educational context, and only some foundational work has been performed in their calibration against human ratings (Chung & Baker, 2003). This research attempts to apply the Rasch model to co-calibrate the scales of AES systems and human raters for grading essay writing and to implement the Many-facets Rasch Measurement (MFRM) to examine the fairness and consistency of raters and adjust students' final essay scores for the measured differences in prompt difficulty and rater severity in the essay writing. Rasch measurement models can provide a framework for obtaining objective and fair measurements of writing ability which are statistically invariant over raters, writing tasks and other aspects of the writing assessment process.

In this research, there are two sets of essay writing data; one from the United States and a second from Hong Kong. All the students were administered the writing prompts from the United States National Assessment of Educational Progress (NAEP)

and essays were human raters scored by using the NAEP holistic essay marking rubrics covering the narrative, informative and persuasive genres. For the data set of students in the United States, 3453 essays were collected from 589 Grades 4 to 12 students in one school district in north-central Mississippi of the United States. All the students were administered four to six writing prompts. Every essay was rated by four of the nineteen independent paid trained raters from an established US testing company. Moreover, a sub-sample of essay writing of the students in the United States - selected to be the best matched to the levels of essay writing of the samples of Hong Kong students - were scored by two of the four experienced Hong Kong teachers of English from one local primary school. For the set of Hong Kong student essays, 408 essays were written by 137 Grade 6 Hong Kong students who were from the same local primary school as the Hong Kong raters. The students were administered two to three writing prompts, each essay was scored against the same rubric by two of the four Hong Kong English teachers. The US essays were scored using the Lexile Analyzer and IEA scoring engines; the HK essays by the Lexile Analyzer alone.

The MFRM analyses revealed that all human raters including 19 raters from the United States and 4 Hong Kong teachers of English scored essays consistently, but these two groups of raters scored essays very differently from each other according to the infit and outfit statistics. Moreover, the Lexile Analyzer and IEA scored consistently with the trained and professional human raters in this research. For the essay writing performance of students, most of the misfitting cases were students who performed erratically, produced off-topic essays or wrote essays that were scored by the most severe or lenient rater(s) in this research. For example, the students misinterpreted the instructions of the prompt; as a result, the ratings for that prompt were far worse than the students' average performances on the other prompts.

This research was conducted to calibrate the reliability of a new AES system, so teachers, students and parents are provided insight into the nature of the essay scoring process. Most importantly, the research probes issues about the fairness and consistency of essay scoring when the Lexile Analyzer is used as an assessment tool and highlights the time- and stress-saving potential of this tool for Hong Kong language teachers by overcoming problems of validity and reliability and by encouraging students' independence during their English language learning.

Undoubtedly, the results might be significant for teachers, students and parents elsewhere. The limitations of this research are related to the need for a larger set of data on student essays, scorings of human raters and AES systems. If a standard scale to measure the relative validity and reliability of the various AES systems and human raters is established, future research might provide a reference guide for the stakeholders who are involved in English essay writing assessment.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter One

Introduction

Background

English is a highly valued component of the core curriculum in Hong Kong schools, even after the handover of the Hong Kong Special Administrative Region (HKSAR) by the United Kingdom to the People's Republic of China in 1997.

The significance of English means that a high score in high-stake English assessment is essential for university entrance, and for general academic success. The importance attached to such assessment, highlights the necessity of an assessment system that is reliable and fair. Given that English curriculum and syllabus documents are organized around the four macro skills of reading, writing, listening and speaking, appropriate assessment of any and all the skills is crucial. For Hong Kong teachers of English, the most time consuming of these is writing; the marking and provision of feedback on writing for English constitutes a major element of their workload.

Essay writing operates as a form of performance assessment both in the classroom and in high-stakes testing. If consistency in scoring student performances contributes to reliable assessment, teachers and other stakeholders can then have confidence in the results and those data can be aggregated or combined across classrooms, grade levels, departments, schools, districts or states (Gottlieb, 2006). In other words, the judgment of teachers or paid essay raters usually affects the ranking of students' academic performance, and this subsequently impacts on the opportunity for students to earn a place in high school, college or university.

Judging the quality of student writing is an extremely complex task. Although there has been considerable discussion about fairness and consistency of essay scoring in educational assessment, routine questions remain: How is any essay given the mark that it is worth? How do all essays of an equivalent worth get the same mark?

In the 1980s, language teachers in some Hong Kong school systems seldom used scoring rubrics to score essays; they were required merely to score students' essay writings impressionistically and to allocate marks in different areas, such as content, language and organization. Impressionistic marking is the precursor of holistic scoring, and is still thought of as synonymous with holistic scoring, although holistic scoring has developed considerably from the original concept of impressionistic marking. Each teacher reads each sample and assigns a grade without any scoring guidelines or discussion of the shared meanings of scores (Hamp-Lyons, 1991b). Sakyi (2000) found that in most cases there were combinations of 'good' and 'bad' qualities in impressionistic marking, and raters could use their own discretion to decide which ones were important. Their decisions to focus on different qualities or characteristics were found to be influenced by different factors including their reading style and their own expectations as well certain distinctive characteristics of the essays. Thus, teacher A and teacher B might have discrepancies on scoring the same piece of essay writing. Similarly, they might be unable to explain the difference between the scores of 75 and 80 in the own ratings.

In the past decade, many Hong Kong English teachers have tried to achieve a common understanding of the criteria in standardized holistic or analytic marking to ensure reliability with colleagues teaching in the same year level by having sample-marking meetings or sharing sessions. The process of moderation in the school has become a regular part of essay scoring for Hong Kong teachers. In fact, the general quest for what is seen as objectivity and reliability in grading drives many language teachers to try a range of different grading approaches. This is a recognition of the importance attached to good academic results in the public examinations in Hong Kong; good results being the first crucial key for students to enjoy success in this society. Inevitably however, there are some teachers who trust their own marking experiences intuitively and discrepancies in scoring still exist. It seems that no completely reliable way may ensure the consistency of the given marks of student written work when teachers in the same year level mark students' work only in their own classes. So clearly there are problems to be addressed in essay rating processes in Hong Kong schools.

As an English language teacher, this researcher understands thoroughly how difficult it is to ensure fairness and consistency of essay scoring both in classroom assessment and high-stakes assessment. Rather, this researcher's experience would suggest that it is difficult for colleagues to achieve compromise and agree on every rating: discrepancies are almost inevitable.

An additional pressure on marking, aside from achieving intrinsic fairness and transparency, is the importance attributed to marks by students and parents. This researcher observed an illustrative episode in a school. A colleague received a complaint from a Grade 5 student's parent about her son's receiving a grade of 84/100 rather than 85 in the school English writing assessment. It seems a trivial complaint. Why did the parent take the one mark difference seriously? In this highly competitive city, education is seen as the essential key to any of life's chances (Hamp-Lyons, 2007). The results of that Grade 5 examination would be reported to the Hong Kong Education Bureau and would affect that child's opportunity to earn a place in a prestigious secondary school in the near future. It is difficult to imagine the strange scenarios in Hong Kong families in this past decade. For example, some parents schedule their children to learn languages, musical instruments, sports and to have tutorials for academic subjects every single day of the week. Some students receive training on interviewing skills to prepare for entering a target primary or secondary school. Some parents queue up for the limited application forms of their favourite school at four or five o'clock in the morning; and some of them work overnight to prepare a book-size portfolio to record all the prizes and learning experiences of their children. When a child is accepted in a prestigious school through the centralized school admission system, the mother might cry in front of the camera in the television news report and claim, ''I'm much happier now than the day I got married".

This kind of constant pressure on teachers, parents and even students makes it particularly vital that essay scoring in Hong Kong is consistent, fair and transparent. Again, this reinforces the importance of ensuring fairness and consistency of essay scoring in both classroom assessment and high-stake assessment in this small corner of the world.

To achieve such rigour in the marking of large quantities of essay writing, Hong Kong schools have tried to implement a range of essay-marking strategies in recent years. For years this researcher has worked closely with English teachers in Hong Kong secondary and primary schools to provide professional support for the implementation of assessment strategies in English language teaching and learning. During this time she observed that teachers started to adopt the use of specific holistic or analytic rubrics for particular writing tasks in marking student essays. In the past decade, teachers have had opportunities to explore the use of rubrics in essay scoring and have adopted scoring rubrics from high-stake assessment to apply them to developing practical scoring systems in their schools. The use of rubrics at least avoids complaints from parents about why their child could not have one more mark to achieve the level of Grade A in the school English writing assessment. Teachers might explain the achieved grading to students and parents according to the descriptors of the scoring rubrics.

In fact, the variability from teacher to teacher that affects scores is a significant issue in the assessment of writing. The variation in grades is most likely the result of teacher subjectivity, which has the potential to result in unfair essay scoring. It is not common for teachers to achieve perfect agreement on scores even though they judge the essays against the same scoring rubrics. Interpretation of such rubrics might be affected by the teachers' particular language backgrounds, the various genres of the prompts and assumptions about the criteria underlying good writing. To attempt to solve this problem, it is a routine practice in some Hong Kong schools for teachers to take turns to mark all student English essays in the same year level in the examinations within a short period of time. This is done to help ensure objectivity and reliability of rating, important issues given that academic results will be one of the major factors impacting on entry to secondary school or university places in the future. Conversely, it has been found (Popham, 1990) that if a rating procedure requires raters to work for several days, raters might be prone to fatigue or boredom. For example, some raters might find themselves more attentive to the rating task in the morning, but less attentive in the afternoon.

The problem of variability remains evident during the complex rating process. The definition of the criterion and standards of achievement, rater training or control of the consistency of rater standards might be among the numerous factors of variability in the writing setting that could influence the possible outcome for any student.

In light of the difficulties outlined above in human rating, it cannot be denied that judging the quality of student writing is also a time-consuming task. Teachers might learn more about the achievement of student learning by assessing essay writing continually since day-to-day teacher observations of students' writing are the best way to monitor student progress in written language learning. Thus, essay scoring is a key routine practice of every language teacher. Scoring student essays not only raises question about how fair, consistent and objective professional judgments can be in the light of the obvious subjectivity of essay assessment, but also, as indicated above, occupies most of language teachers' working time.

In the 1960s, Ellis Page developed a system to set the stage for automated writing evaluation. Automatic Essay Scoring (AES) systems are engines that employ computer technology to evaluate and score written prose in the place of the usual human rater. At least any AES system is consistent in the way it scores essays, and saves time and cost for scoring the usual written language tasks. The AES method for assessing English language written essays has become quite widely accepted, especially in the United States, as an educational technology for both low-stakes classroom assessment and large scale high-stakes standardized tests (Dikli, 2006).

Over the last forty years, a number of studies have been conducted to assess the appropriateness of the AES systems for writing assessment. The research demonstrates that AES scores correlate well with those of human raters behavior that they reveal a high level of construct validity (Shermis & Burstein, 2003). Burstein and Chodorow (2002) described the advances of computer scoring of essays with a focus on the reliability of the software, a key issue for faculty members who consider appropriate ways of incorporating such technology into their courses and programmes. They concluded:

> A goal of current research in automated essay analysis and scoring is to develop applications to ensure that systems maintain a relevant link to what writing experts and test developers believe are critical to the teaching and learning of writing. Future research should continue to enhance the validity of automated scoring so that computer-based methods of essay analysis will be in line with the educational goals in writing instruction and so that systems can adequately represent the underlying aspects of writing assessment. (p. 497)

Rudner and Gagne (2001) conducted a study demonstrating that developed software has been used to score essays with considerable success. They concluded that the reliability of three software systems, Project Essay Grade (PEG), Intelligent Essay Assessor (IEA) and the e-rater, is high and suggest that educational systems in states and schools in the US consider using automated scoring services. One study compared the correlations of the automated essay ratings with those of the ratings of six human raters on an overall holistic assessment of writing, the computer ratings (r =.83) achieved statistically significant higher average reliability than did the human raters with themselves (r =.71) (Shermis, Koch, Page, Keith, & Harrington, 2002). Moe (1980) concluded that the analysis of text by computers has been demonstrated to be both feasible and practical. Computers are now more readily available and the programmes are rapidly improving. The versatility and the efficiency of computers for text analysis should convince educators that computers are enormously useful tools for studying language. But the question that is regularly raised remains: To what extent can the AES systems replace expert human raters for essay scoring in classroom assessment or even in high-stakes assessment?

In spite of the growing body of data surrounding the use of AES systems, little research has been done on validating the performance of the AES models in an applied educational context, and only some foundational work has been performed on their calibration against human rating responses (Chung & Baker, 2003).

Purpose of this Research

The maintenance of fairness and consistency is a crucial element in essay scoring in both classroom assessment and standardized high-stakes assessment as the results of essay scoring will impact directly on the academic development of students. This research aims to explore how a new AES system, the Lexile Analyzer, ensures the fairness and consistency of essay scoring and what level of similarity there is between the scores given by this AES system, those provided by trained human raters and those given by volunteer teachers of English from one Hong Kong primary school. It might be that a handy and reliable tool could reduce the assessment workload for teachers in Hong Kong and to release them from some of the pressure of scoring student essay writing by providing a solution which avoids potentially unfair situations in classroom and high-stakes assessment.

For part of analyses, this research attempts to apply the Rasch model to co-calibrate the scales of AES systems and human raters for grading essay writing. Rasch measurement models can provide a framework for obtaining objective and fair measurements of writing ability which are statistically invariant over raters, writing tasks and other aspects of the writing assessment process. Both the difficulty of the writing task components and the ability (as well as prompt difficulty and rater severity) are expressed on a single linear interval measurement scale (Bond & Fox, 2007).

The other crucial part of this research is to implement Many-facets Rasch Measurement (MFRM) to adjust for writing prompt (essay topic) difficulty as well as to detect and manage variations rater severity in the scoring of essay writing. The rater-scale interaction resembles the subject-instrument interaction in that the rater-scale interaction can be considered as a 'test' of the raters and the scale in the way that the subject-instrument interaction is a test of the students and the instrument (McNamara, 1996). The MFRM is the extended version of the Rasch model to situations in which more than the usual two test facets (person ability and item difficulty) interact to produce an analysis. It enables the construction of a frame of reference in which quantitative comparisons no longer depend on which students happened to be scored by which raters on which essay (Linacre, 1989). The long-term aim of any genuine measurement system in the human sciences should be accessed to a series of co-calibrated testing situations, such that the raw score or locally calculated person measure could be expressed on the single scale used world-wide (Bond, 2004).

The focus of this research is to investigate how effective the Lexile Analyzer is in measuring essay writing ability as judged by trained professional raters. This new AES engine is particular in that it applies the same principle to measure writing ability on the same scale as does the Lexile scale of reading ability and that the essay scoring is independent of using a sample of human-rated essays for training.

To do this, an estimation of rater, genre, student and rubric effects must be undertaken. Then, the estimates from the recently developed the Lexile Analyzer will be modeled against those of Intelligent Essay Assessor (IEA), an AES already widely used in high-stakes testing in the United States. Finally, a concurrent co-calibration of the scores of human raters, the Lexile Analyzer and IEA needs to be undertaken.

In this research, there are two sets of essay writing data; one from the United States and a second from Hong Kong. All the students were administered the writing prompts from the United States National Assessment of Educational Progress (NAEP) and essays were human rater scored by using the NAEP holistic essay marking rubrics. For the data set of students in the United States, every essay was rated by four independent paid trained raters from an established US testing company. Moreover, a sub-sample of essay writing of the students in the United States - selected to be the best matched to the levels of essay writing of the samples of Hong Kong students - were scored by four experienced Hong Kong teachers of English from one local primary school. For the set of Hong Kong student essays, each essay was scored against the same NAEP rubric by two of the four Hong Kong English teachers. The US essays were scored using the Lexile Analyzer and IEA scoring engines; the HK essays by the Lexile Analyzer alone.

Linacre and Wright (2002) found that dramatic misfit of the data in the MFRM model can reveal that the allocated raw scores might have doubtful meaning. The MFRM analysis conceptualizes the essay writing setting in terms of facets and shows the interactions of the facets determine the likelihood of particular raw scores, distinguish between raw scores and what they may indicate of underlying ability in students, and introduces the possibility of a technology that can compensate for aspects of the test situation which vary from student to student (e.g., essay topic and essay rater) (McNamara, 1996). Thorndike (1904) concluded that the development of ideal valid scales has steps of equal difficulty between the calibrated objects on the scale. Many-facets Rasch Measurement (MFRM) is part of the solution to the problem of rater and prompt variability for high-stakes testing. It is a tool to investigate rater consistency and adjust rater severity and students' scores for the measured differences in prompt difficulty in the writing assessment. It is common for students to choose one of a selection of topics/genres to answer in a high-stakes writing assessment, and all of the prompts are assumed to be equally difficult. Thus, the MFRM analysis may avoid the problem of raw scores being an unreliable indicator of student writing ability.

One aim of the writing process should be that the measures of performance given to each student, as derived from the ratings given by raters, be as fair, accurate and useful as possible. Consequently, although the measure given to an examinee is derived from the particular rater(s) who rated that essay, it must, in meaning and implication, be independent of them. The measure must be 'judge-free' (Linacre, 1989). Thus, Many-facets Rasch Measurement (MFRM) was implemented to detect and account for rater effects and the effect of a range of other variables in reporting student writing performance in the United States and Hong Kong in this research. This Rasch measurement model possesses desirable statistical and psychometric properties related to the separability of parameters with sufficient statistics available for estimating these parameters (Engelhard, 1992).

If this research proves to be successful in calibrating the new AES system, the Lexile Analyzer, against the ratings usually provided by trained raters in the United States and Hong Kong teachers, it should then provide insights into the nature of the essay scoring process and provide more options for educators, policymakers, government bodies, school administrators, teachers as well as students, for scoring essay writing, and subsequent related learning and teaching strategies. Most importantly, it can help teachers and students in Hong Kong because the results of essay writing in standardized external high-stakes assessment will have a serious influence on the future of students. Further, if the Lexile Analyzer can help ensure the maintenance of fairness and consistency of essay scoring, it could lighten Hong Kong language teachers' grading load and relieve their considerable stress in the process of scoring, and encourage students to have daily self-learning activities in essay writing for improving their English writing skills. The results, if confirmed independently might be significant for teachers, students and parents elsewhere.

Summary

In countries where English is an official, or even highly valued second language, the learning and teaching practices for reading and writing will significantly affect the development of students, as well as impact on their academic and professional prospects. The significance of English results for the students makes it salient that fairness and consistency on essay scoring in classroom settings and high-stakes assessment is maintained, regardless of the difficulties this entails.

The AES systems appear to offer promise of such reliability, at least. Research over the last forty years has provided strong evidence that these systems can be effective for scoring essay writing. This current research will investigate the ratings generated by a new AES system, the Lexile Analyzer, comparing it with human raters in the United States and Hong Kong on essay scoring by using the Rasch model for measurement. The results of this research should provide insights to the stakeholders who are involved in classroom and high-stakes essay assessment in the educational systems. Moreover, the new AES system could provide an assessment tool as a solution to the heavy load of essay rating for the Hong Kong teachers, as well as addressing issues related to validity and reliability.

Chapter Two

Literature Review

Essay writing is usually a core requirement in all English language assessment, both where English is the student's first language and in places such as Hong Kong where Cantonese language users are required to demonstrate their English language proficiency. However, basic problems persist: How might raters mark independently of their own perspectives (objectivity)? How might raters be able to give the same score to the same essay repeatedly (reliability)? How might objectivity and reliability of essay scoring be implemented and sustained, especially in high-stakes assessment? A number of studies have been conducted to assess various practices of English essay writing assessment.

Essay Writing

Within the language teaching field no consensus has emerged on what is the right way to teach or learn, nor has any single dominant model for language teaching been established (Hamp-Lyons, 2000). However, it does seem to have been agreed by language teachers and researchers that essay writing provides a significant performance-based assessment in language teaching and learning.

Writing is a literacy process by which students use their prior experiences and knowledge of the world to apply a variety of strategies to, ultimately, make meaning (Peregoy & Boyle, 2005). It is a process routinely used to determine student competence in integrating new language and comprehension information with previous knowledge and experience (Applebee, 1984). Essay writing is a form of performance assessment in which students activate their linguistic knowledge in the given contexts and make use of their creativity, express personal ideas and feelings, and develop and demonstrate their critical thinking skills. It is one of the tasks crucial for representing what students have learnt in their own words.

Further, essay writing is an essential element of English language learning. Teachers use essay writing to enhance and consolidate the language structures and vocabulary items that students acquire in other aspects of language learning. In other words, student writing skills, such as choosing the correctly spelled word from a list, selecting the appropriate punctuation for an unpunctuated sentence or picking out the best organised paragraph from among several offered, can all be revealed from their essay writing (Godshalk, Swineford, & Coffman, 1966).

Assessment of Essays

In any academic year, Hong Kong students have to experience a large number of on-going formative writing assessment and summative assessment at the term end in English language learning. The assessment results usually provide evidence of what students have learnt and will be used as part of the criteria for promoting students to the next year level. Of course, teachers play a crucial role in assessing student essays and providing sustained constructive feedback to students to enhance their access and make meaning in their essay writing. Primarily, teacher’s feedback is offered to help students to identify their strengths and weaknesses for improving both their essay writing skills and other aspects of their English language learning, as well as helping teachers review their teaching strategies to improve the quality of instruction (Hong Kong Education and Manpower Bureau [HKEMB], 2004b).

In the United States, students are generally required to take a writing assessment test for placement at universities and colleges (Carlisle & McKenna, 1991). Student responses to the writing assessment in state departments of education in colleges and universities are often associated with high-stakes decisions (Johnson, Penny, & Gordon, 2000). For Hong Kong students, achieving good academic results in high-stakes assessment is crucial for entry to university where participation rates are limited by the provision of comparatively smaller numbers of university places. According to the report on Hong Kong student enrolment by level of education, only the top 18.2% of all Grade 13 (Secondary 7) students entered university in 2010 (Hong Kong Census and Statistics Department, 2010).

Common Genres in Classroom and High-stakes Assessment

Hong Kong students are exposed to a number of genres of essay writing including the four most common genres in essay writing: narrative, informative/expository, descriptive and persuasive (Graesser, McNamara, & Louwerse, 2003) during their school learning experiences. The HK curriculum outlines illustrate that writing stories that have a clear sequence of events and descriptions of characters, presenting and elaborating the main ideas and supporting details is as important as using persuasive devices effectively in essay writing at primary and secondary levels (Hong Kong Education Department [HKED], 2002; HKEDB, 2007a; HKEMB, 2004a).

Each type of genre has its own linguistic and writing demands. Narrative writing requires students to provide chronological development and description of events or experiences while informative/expository writing expects students to support the ideas presented with detailed explanation and logical development of thought (Quellmalz & Burry, 1983). Descriptive writing asks students to describe a person, object, emotion or situation and develop an expression of a particular experience. Persuasive writing requires the use of complex language to analyse and resolve controversies and make diverse points of view in a clear, convincing and considerate way (Nippold, Ward-Lonergan, & Fanning, 2005).

A number of studies have focussed on comparing student performance on different genres. Engelhard, Walker, Gordon, and Gabrielson (1994) conducted a 3-year study on ratings of the essays of Grade 8 student in high-stakes examinations in the United States. Narrative, descriptive and expository writing tasks were randomly assigned to students and each student responded to one writing task. The results revealed that narrative writing tasks received the highest ratings and that expository writing tasks received the lowest ratings. On the other hand, Quellmalz, Capell, and Chou (1982) found that Grade 11 and 12 students had higher scores when they responded to an informative prompt rather than a narrative prompt. They explained that students might have more practice in informative writing in the high school English curriculum in the United States.

Persuasive essays have been found to be syntactically more complex than either informative or narrative essays (Crowhurst & Piche, 1979). In persuasive writing, students attempt to distinguish fact from opinion, offer multiple reasons for their views and anticipate and reply to counter-arguments (Nippold, Ward-Lonergan, & Fanning, 2005). Students have acquired a sufficient amount of background knowledge, so they are able to appreciate the complexity of controversial topics and address different points of view (Chall, 1983). Elementary school curricula in the United States usually leave out persuasive writing because it involves the skills of formulating, analysing, and synthesising reasons (Burkhalter, 1995). Thus, training in persuasive writing is usually planned in English language curricula in high school or tertiary level. When Hamp-Lyons and Prochnow (1990) conducted a study of undergraduates, who learn English as a Second Language (ESL), wrote informative and persuasive essays in Michigan English Language Assessment Battery (MELAB), they found that the students were higher on average for persuasive prompts than for informative prompts.

Given that the link between student results and essay prompts is inconclusive, it might be more reasonable to ask students to write a range of text types as part of their writing assessment.

The Hong Kong English Language Curriculum Guide Primary 1-6 claimed that the introduction of a variety of text types to learners is crucial. Students at primary level should learn to produce essays in major genres including narrative, informative and persuasive texts (HKEMB, 2004a). According to the Writing Outlines of the Hong Kong English Language Curriculum Guide Secondary 4-6 (Hong Kong Education Bureau [HKEDB], 2007a), students are expected to present different views and arguments clearly and logically, and use persuasive devices effectively at secondary level. Thus, in Hong Kong high-stakes writing assessment, students are assessed by writing a variety of genres to reveal their English writing proficiency levels.

The Territory-wide System Assessment (TSA) of Hong Kong is a high-stakes assessment for teachers and schools because the assessment results will be used to promote better learning and teaching. In the TSA writing paper in 2008, the two Grade 6 English essay writing prompts required students to write a narrative either about meeting a dog on the street or planning how to spend the spare time (Hong Kong Examination and Assessment Authority [HKEAA], 2008b). Students would either have to describe a dog or provide possible ways to spend the spare time.

In another high-stakes assessment, the Hong Kong Certificate of Education Examination (HKCEE) for Grade 11 (Secondary 5) students, examinees were offered three genres in the Writing paper 2005: narrative, informative and persuasive (HKEAA, 2007a). Examinees had to finish a compulsory written task and selected one of three questions to answer.

In the United States, the National Assessment of Educational Progress (NAEP) is the only nationally representative and continuing assessment for Grades 4 to 12 students. It requires students to respond to the prompts that covered the narrative, informative and persuasive genres in the writing assessment section (US Department of Education [USDOE], 2009).

Essay Scoring

In the early 1940s, the College Entrance Examination Board (CEEB) in the United States had successfully piloted the Scholastic Aptitude Test (SAT) for the college admission of special populations of students, and sought ways to expand the use of the SAT. The CEEB immediately suspended essay testing because the new SAT could produce admission data for students more quickly, allowing them to start college sooner, so the deferment of military service was shortened for college students (Huot & O'Neill, 2009).

Currently, direct tests of writing remain useful in deciding student placement into writing programmes, in test batteries used for admissions decisions for schools and colleges, and in evaluating writing programmes in terms of instructional effectiveness (Hamp-Lyons, 1991a).

Essay Scoring Rubrics

The use of rubrics to score essays is one of the ways that educators tried to address issues of validity and reliability in essay marking.

> [A] couple of decades ago, rubric began to take on a new meaning among educators. Measurement specialists who scored students' written compositions began to use the term to describe the rules that guided their scoring. They could have easily employed a more readily comprehensible descriptor, such as scoring guide, but scoring guide lacked adequate opacity. Rubric was a decisively more opaque, hence technically attractive, descriptor. (Popham, 1997, p.72)

Brookhart (1999) describes an analytic rubric as a checklist allowing for the separate evaluation of each of the criteria, which are scored on a different descriptive scale. On the other hand, the criteria in a holistic rubric are considered in combination on a single descriptive scale which requires a broader judgment concerning the quality of essay writing. Rubrics have become common evaluation tools for student work from elementary level to college level (Moskal, 2000).

Using Analytic Rubrics on Essay Scoring

Quellmalz and Burry (1983) concluded that more resources may be required in analytical rating procedures, but the greater time commitment required by analytic rating is offset by instructional advantages. Analytic rubrics play a crucial role in instructional improvement by identifying particular strengths and weaknesses for individual students and indicating areas where school curricula may need improvement. Bacha (2001) claimed that advanced students benefit by the use of analytical evaluation instruments because they provide more information about their essay writing. Criteria in rubrics used to assess students' essays should reflect agreed-upon standards of good writing (Quellmalz & Burry, 1983).

The use of criteria as the basis of an analytic assessment of essay writing has its foundation in the work of Diederich, French, and Carlton (1961) who used factor analysis to determine the main influences on raters' scores in essays written as homework by 300 college freshmen in the United States in 1958. The five factors, Ideas, Form, Flavor (Style), Mechanics and Wording, were the basis of what would become known as analytic scoring by which raters could score student writing. Scores for each of the factors could then be weighted depending on the purpose of assessment. In contrast, Vaughan (1991) and Freedman (1979) claimed that 'content' was the most crucial factor in the final judgement of students' essays. Stewart and Grobe (1979) concluded from a study of teacher-markers of Grades 5, 8 and 11 student essays that the raters were primarily influenced by the length of the compositions and simple mechanical errors. The findings of Carlisle and McKenna (1991) and Santos (1988) revealed that content was more important than language errors in the ratings of ESL students' essays at universities and in large-scale placement tests. Hamp-Lyons (1991a) suggested that quality of content should be the most prominent criterion for making judgements, in contrast to Grobe (1981) who earlier found that vocabulary diversity was the most important feature of what teachers perceived to be a piece of good writing. Schaefer (2004) used Rasch analysis in his doctoral research about ratings of Japanese EFL essays to conclude that the order of how much each rubric category influenced native English speaking rater judgments was: Language Use (the most difficult category), Fluency, Mechanics, Organisation, Content, and Style and Quality of Expression (the easiest categories).

Although it is not easy to have agreement on the key criteria for any marking scheme, scoring rubrics with clearly described performance standards are helpful and significant. Bacha (2001) stated that the choice of criteria categories in rubrics should meet the purposes of the writing assessment.

Using Holistic Rubrics on Essay Scoring

Hillegas (1912) developed the scale which is accurate enough to be reference to researchers in the same area in measuring the quality of English compositions written in the upper grades of elementary school and high school. He stated that the purpose of the scale was to provide a ruler for a 'holistic' quality judgement. The actual characteristics of composition merit were not teased out; that is the composite elements of quality were not identified or made explicit. Diederich (1974) played a key role in developing and applying holistic scoring procedures in the school setting. He stated that it is not a blur to score essays with holistic scoring procedures and that it saves time if teachers have a common understanding on identifying the level (high, medium or low) of a paper. Charney (1984) described holistic scoring as a quick, impressionistic qualitative procedure for sorting or ranking samples of writing. It is not designed to correct or edit a piece, or to diagnose its weaknesses. Instead, it is a set of procedures for assigning a value to a writing sample which is based on the previously established criteria or general impression from a rater. Holistic scoring is a feasible method for testing and for research involving large numbers of writers.

With the support of the National Assessment of Educational Progress (NAEP), Lloyd-Jones (1977) and his colleagues from the National Council of Teachers of English (NCTE) revised analytic scoring using categories relevant to the writing task to develop a special kind of holistic scoring, Primary Trait Scoring, which focuses on identifying the presence of traits required by the rhetorical situation in the essay writing (Myers, 1980). The key steps in using the Primary Trait Scoring system were to define the universe of discourse, to define topics which sample that universe precisely, to ensure cooperation of the writers, to devise workable scoring guides and to use the guides. Primary Trait Scoring was used in early writing assessment for NAEP in the United States (Huot & O'Neill, 2009). In the current research, the writing prompts were selected from NAEP and holistic rubrics were used in essay scoring.

Vaughan (1991) claimed that the use of holistic assessment of essays to place and to grade students had become widespread in college programmes across the United States. Testing agencies, such as ETS, and numerous US schools now routinely assess writing samples with holistic scoring. When educational policy makers consider the issues of testing, one of the main concerns is the cost of testing. Thus, it was an important step to demonstrate that the holistic scoring of writing samples could take place quickly enough to be practical (White, 1984). Myers (1980) claimed that one of the most productive ways to score writing is to get a numerical scale on the holistic scoring rubrics by an overall impression of writing samples.

Sakyi (2000) summed up some significant observations of holistic raters from verbal protocol analysis as follows:

- experienced holistic raters are able to interact with text and express their personal impression in addition to the criteria outlined in the scoring guidelines; and
- holistic raters could be influenced by both content and language-related factors as well as their own expectations and personal reactions. (p.133)

Students' writing proficiency levels can be identified by holistic scoring rubrics; on the other hand, detailed criteria in analytic scales provide more specific feedback on students' progress and to evaluate students' proficiency levels for promotional purposes (Bacha, 2001). Moreover, it is a common practice for language teachers or paid professional raters to learn how to apply criteria of rubrics accurately through some marking training before proceeding to actual essay rating independently, especially scoring for high-stakes assessment. The subjectivity involved in the process of evaluating essays with a scoring rubric becomes more objective (Moskal, 2000). It helps develop common understanding of criteria and maintain high levels of agreement among raters to overcome the problems of being objective and reliable in essay scoring. Broad feature information is found in a well-designed scoring rubric which can be used for explanation of scores, diagnostics and instructional information (Burstein et al., 1998). Thus, not only can students be assisted in understanding their weaknesses, but their writing skills can also be strengthened. Further refinement or moderation can easily reduce the occurrence of score discrepancies among raters to ensure the maintenance of reliability of essay scoring.

Essay Scoring in High-stakes Assessment in Hong Kong

Hong Kong has a very rigid curriculum and a tight structure of assessment throughout the educational system. The dominance of an exam culture geared teachers towards preparing students for public examinations (Hamp-Lyons, 2007). Students start to face various high-stakes writing assessment when they are in Primary Grade 3 level. How do stakeholders in the educational system try to minimize the discrepancies and resolve discrepancies fairly?

The Hong Kong Examinations and Assessment Authority (HKEAA)

The Hong Kong Examinations and Assessment Authority (HKEAA), formerly known as the Hong Kong Examinations Authority, is a statutory body established in 1977 and is responsible for all high-stakes public examinations in Hong Kong. The HKEAA is currently investigating how certain aspects of the marking process can be computerized to modernize the marking systems for high-stakes examinations in line with developments elsewhere in the world (Legislative Council Panel on Education, 2005). The following part is the HKEAA's ongoing efforts to ensure that high-stakes examinations in Hong Kong are valid, reliable and fair. All the evidence shows that it is crucial to maintain fairness and objectivity in high-stakes examinations.

Onscreen Marking (OSM)

HKEAA (2010) reported that Basic Competency Assessment (BCA) is to provide information to the government for the purposes of school improvement; schools should make use of the assessment data to improve the learning and teaching effectiveness in the schools. In other words, BCA is a low-stakes assessment for students but a high-stakes assessment for teachers and schools. Territory-wide System Assessment (TSA) is one of the components of the Basic Competency Assessment which was commissioned in 2001 by the Education and Manpower Bureau (EMB), the Hong Kong government department responsible for education. The HKEAA is responsible for its development and implementation. It is similar to NAEP in the United States or the National Assessment Program Literacy and Numeracy (NAPLAN) in Australia. English is one of the key learning areas to assess students' English language proficiency. In May 2004, the TSA was first administered in Hong Kong. Since then, all the Grade 3, Grade 6 and Grade 9 (Secondary 3) students are required annually to attempt one writing task for the sub-paper of English language subject. The full paper is composed of a number of sub-papers and each sub-paper contains a number of common items which act as the anchor component in the assessment (Qian, 2008).

In order to improve the speed of essay scoring and overcome the problems of paper-based marking (Coniam, 2009a), the HKEAA has adopted an Onscreen Marking (OSM) system in 2008 for the scoring of TSA papers. Before that, the scoring of Chinese Language and English Language papers in the Hong Kong Certificate of Education Examination (HKCEE), which is a high-stakes assessment for students to take at the end of Grade 11 (Secondary 5), has already adopted OSM in 2007 (HKEAA, 2007b). It is anticipated that OSM will be fully implemented for all public examinations in Hong Kong from 2012 onwards (Coniam, 2009a).

All HK OSM raters are qualified serving teachers, they score the scanned student scripts on the computers at the assessment centres within a particular period. White (1984) claimed that the controlled essay marking eliminates irrelevant variables in the scoring process and establishes a positive social situation with the right condition for the success of the scoring and determines and enforces the standards of measurement. Dowling (2008) reviewed the marking reliability of the HKEAA of the Hong Kong public examinations in the light of international best practice during a secondment in the HKEAA. He revealed that grades are assigned by specially trained teachers while the OSM system presents the essays, recording the marks and continually monitors raters' performances. Most of the high-stakes examinations in Hong Kong are recognized for admission by all local and many overseas universities. The HKEAA uses this new OSM system not only to permit more monitoring of the reliability of marking throughout the process, but also to implement a system of double marking for essay papers in which every script is rated by two raters and might have a third marking by the assistant examiners if the discrepancies between the first two raters exceed the allowed range reference scores (HKEAA, 2010).

This is a crucial development as academic results of high-stakes assessment will be one of the major factors impacting on entry to university places, both in HK and overseas, making the quest for fair and consistent grading, including that of English language essays, more important than ever in school assessment practices.

The Rasch Analysis

The results of high-stakes assessment in Hong Kong also impact on people other than students; for example, those who plan to be teachers of English in local primary or secondary schools. According to the Circular Memorandum of Language Proficiency Requirement for English Teachers (HKED, 2000), the government placed heavy emphasis on teachers' English language proficiency with all serving English teachers in local primary and secondary schools being required to pass the five tests of the Language Proficiency Assessment for Teachers (English Language) (LPATE) including a writing test before the end of the 2005/06 school year. It meant that serving English teachers would lose their English teaching positions if they failed to meet the minimum language proficiency requirement. The proficiency levels of this high-stakes assessment are awarded objectively by expert judgment and employing the Rasch measurement model. Rasch model analysis provides the scale to describe both the level of difficulty of test items and the performance of candidates. Anchoring the scales across cohorts enables consistent measurement independent of variations in the ability of candidates and the difficulty level of different test items used in different settings of assessment. Expert judgment involves experts in the field reviewing the test content in combination with the information about item and test difficulty provided by the Rasch analysis in order to relate marks on the test to levels of performance and ability (HKEAA, 2008a).

Potential Problems of Human Raters on Essay Scoring

In spite of the above developments, writing assessment has had reliability and implementation problems since the 1910s, when Starch and Elliott (1912) reported that teachers have large differences in standards of essay marking.

In fact, essay scoring is not a simple task. There are potentially millions of possible combinations of words and ideas that appear in students' essay writing in response to the same prompt. Yet every Hong Kong English language teacher is expected to mark at least five essays as formative assessment for each student in a semester routinely as part of their core work. Their feedback on student essays is usually confined to the errors of grammar usage and vocabulary items with less reference to the specific assessment criteria related to the content and organization of writing (Lee, 2004). Calfee (2000) noted that teachers in the United States don't always have sufficient time or resources to effectively grade student compositions or provide feedback to students. Most teachers would be aware of the potential impact that the essay rater might have on the students' essay grades. Even the most conscientious teacher waxes and wanes in the capacity to be equally fair to all students, across all types of essays, from the start to the end of any essay marking session. This is highlighted for teachers when in-school examinations are graded; some sort of moderation or cross-marking between English language teachers is another routine practice. Moskal and Leydens (2000) claimed that reliability is of more concern on high-stakes testing than classroom assessment. In classrooms, student's knowledge is continuously assessed and this provides information to teachers to adjust in pursuit of effective pedagogy in learning and teaching. In high-stakes testing, the results impact on students' academic development. Thus, the problems associated with being fair and consistent (objective and reliable) are heightened when essay marking is used as an assessment component for large, high-stakes, public examinations of language competence.

The ideal scoring situation for examinations is one in which all raters agree on every rating, i.e., every essay for every child, as though the entire panel of raters could be replaced by any rater with no loss of information (Linacre, 1989). The raters must come from similar academic backgrounds to create a good chance for agreement and allow the raters to arrive at their own set of criteria, so that they will draw as much as possible from common experience and values (Follman & Anderson, 1967; McColly, 1970). Cooper (1977) concluded that a homogeneous group of raters can achieve high levels of reliability, as estimated by statistical tests. Charney (1984) argued that raters must be trained to use rubrics to score essays. She claimed that training procedures are designed to sensitize the raters to the agreed upon criteria and guide them to employ those standards, rather than their own.

In practice, this is not so easy to achieve, and in high-stakes examinations, this aimed-for uniformity is often achieved through some sort of moderating, averaging or other compromise in the allocation of final grades. Although a number of different programmes for training essay raters have been devised to try to improve the quality of essay ratings with the goal of reducing rater errors (Hedge & Kavanagh, 1988; McIntyre, Smith, & Hassett, 1984), ratings vary considerably from one rater to another as well as across essay types. Even the most diligent rater training fails to produce uniformity among raters, but any difference among raters threatens fairness because the student raw score depends on which rater awards the rating (Borman, 1978). Rater training is necessary for creating the conditions for a measurement process based on ratings by making raters more self-consistent (intra-reliability), but there is a limit to how far this process can be successful (McNamara, 1996). Some studies on rater scoring agree that experienced raters appear to bring well formulated strategies not only about the criteria with which to judge essay writings but also about how to conduct themselves during actual rating sessions (Cumming, 1990; Hamp-Lyons, 1991b; Huot, 1993; Janopoulos, 1993; Vaughan, 1991; Weigle, 1994). Quellmalz and Burry (1983) claimed less agreement among raters (inter-reliability) might be the result of fatigue, the range of quality in the essays being scored or the re-emergence of any rater’s idiosyncratic values.

Diederich et al. (1961) conducted a study in which 300 essays written as homework by US freshmen in three colleges were scored by 53 raters. The raters included 10 college English teachers and 43 other raters including social science teachers, natural science teachers, writers, educators, lawyers and business executives. The results showed that 94% of the essays received either seven, eight or nine different grades of the nine possible grades, with no paper receiving less than five separate grades; and the median correlation among raters was .31. Raters in each field agreed slightly better with the 10 college English teachers than with one another. Vaughan's (1991) study showed that trained raters agreed on the rating criteria outlined in the scoring guide in most cases but fell back on their own styles in situations where an essay did not fit well into the pre-defined standards outlined in the scoring guide. White (1984) reported on a two-year study at the California State University in which 1398 essays were rated again one year later by the same raters using a 6-point scale. The rating a year later produced scores that were identical to the first in only 20.7% of the cases. It revealed that most of the essay scores changed upon rescoring, and some scores changed a great deal.

A practical method to correct for those differences and their undesirable effects on the grading process has been unavailable. Carlson (1991) claimed that it is virtually impossible for human raters always to agree on the exact score assigned to a writing paper and raters cannot be trained to make absolute judgments, but rather to make informed, rather than, arbitrary evaluations.

Although individual raters were not entirely reliable and had personal biases, the use of more and more raters permitted a better approximation to the 'true' average rating of an essay. In fact, the requirements of the true quality of essays must be defined by human raters (Page, 2003).

Automated Essay Scoring (AES) Systems

Automated Essay Scoring (AES) systems adopt computer technology to evaluate and score written prose in the place of the usual human grading. This method has become more widely accepted as an educational technology for both low-stakes classroom assessment and large scale high-stakes standardized tests (Dikli, 2006). The system currently is most developed to score written work in English. All essay writing in school settings can be evaluated, except some particular forms of writing, for example, poetry (Shermis & Burstein, 2003). AES systems are time-saving, cost reducing and, especially, reliable in essay scoring (Bereiter, 2003; Burstein, 2003; Hamp-Lyons, 2001; Myers, 2003; Page, 2003). But the systems have been criticized for vulnerability to cheating (Chung & O'Neil, 1997; Kukich, 2000; Powers, Burstein, Chodorow, Fowles, & Kukich, 2002; Rudner & Gagne, 2001) and the need for a large corpus of sample texts to train the system (Chung & O'Neil, 1997).

Page (1966) was a pioneer researcher in AES. His system was able to generate scores that agreed strongly with those awarded by human raters. Foltz, Laham, and Landauer (1999) reported a study of essay writing by American university students which revealed a correlation of .8 between IEA and human raters. Page, Poggio, and Keith (1997) revealed superior correlations between results of marking by the AES Project Essay Grader (PEG) and human raters relative to correlations between human raters.

AES systems have developed considerably in recent years, with the following systems being widely adopted particularly in the United States.

Project Essay Grader (PEG)

A national network of the US universities, the College Board, supported Ellis Page to develop the first AES, Project Essay Grader (PEG), in the early 1960s to make large scale essay scoring practical and effective (Page, 2003; Rudner & Gagne, 2001). Shermis, Mzumara, Olson, and Harrington (2001) conducted a study to examine PEG's scoring of web-based essays as placement tests at a university. 617 essays were scored holistically by six human raters and PEG. The scores of PEG showed a correlation with human raters' scores of .71, but human raters scores correlated with each other's at .62.

Page and his team used the terms 'trins' and 'proxes' to explain how PEG generates a score. Trins refer to the intrinsic variables such as fluency, diction, grammar, punctuation, and proxes denote the approximation (correlation) of the intrinsic variables to refer to actual counts in an essay. One very useful prox has been found to be the length of the essay (Dikli, 2006). The PEG system needs from 100 to 400 manually marked sample essays for training purposes, and to determine values for up to 30 proxes (Rudner & Gagne, 2001). Proxes are determined for each essay and the variables, such as average sentence length, number of paragraphs and subject-verb openings, are entered into the PEG prediction equation. Then a score is assigned by computing beta weights (coefficients) from the training stage results (Chung & O'Neil, 1997).

Page and Petersen (1995) revealed that PEG programmes explore complex and rich variables, such as searching each sentence for soundness of structure and weighting these ratings across the essay. They used 300 essays to compare the performance of PEG and human raters. The essays were scored by six human raters and the ratings were compared with the predicted score generated by PEG. The results revealed that the PEG scores correlated with individual raters from .72 to .78 with an average correlation of .74. The human raters correlated with each other from .55 to .71 and the average correlation of six human raters was .65. Page and Petersen concluded that the PEG demonstrated better scoring performance than did human raters.

In other words, PEG can identify sentence length, number of paragraphs and elements of punctuation to determine how well the different variables correlate with the training sample scores of human raters. But the system focuses on the surface structures rather than on the semantic aspects of essays; nor does it provide instructional feedback to students as it does not detect the content-related features of an essay (Dikli, 2006). The fluency, diction, grammar and punctuation in an essay cannot be assessed directly but can be evaluated only though possible correlates (proxes) (after, Valenti, Neri, & Cucchiarelli 2003). Thus, the early version of PEG found only limited acceptance in the writing and education community (Kukich, 2000).

E-rater

In the 1990s, a team of US Educational Testing Service (ETS) researchers, led by Jill Burstein, hypothesized a set of linguistic features that might more directly assess general essay writing qualities - features they could automatically extract from essays using Natural Language Processing (NLP) and Information Retrieval (IR) techniques. They conducted a series of preliminary automated essay scoring studies, with data sets from the Analytical Writing Assessment (AWA) of the Graduate Management Admissions Test (GMAT). Each student in the AWA is required to write two essays, one to analyze an argument presented in a short text and another to express an opinion on a specific issue presented in a brief statement. 400 essays from each essay type in the assessment were scored by two human raters using the GMAT guidelines on a six-point holistic scale. Then, the team implemented computer algorithms to extract values for every feature out of more 100 automatically essay features from each essay, the optimal scoring models consisted of mainly linguistic features worked as well as those containing only proxy features. ETS patented the resulting automated essay scoring system, Computer Analysis of Essay Content (CAEC) and renamed it e-rater (Kukich, 2000).

The e-rater system uses NLP to model the performance of human raters. It is the application of computational methods to analyze characteristics of electronic files of text or speech. It also includes three NLP-based modules for scoring: structure (the syntax of sentences), organization (various rhetorical features) and content (prompt specific vocabulary). Moreover, the system uses NLP techniques to identify specific lexical and syntactical cues of the faculty-scored essays in its sample collection and stores them with their associated weightings in a database. It compares these features in any target essay to those in the database to assign a score to that new essay (Burstein, 2003). To measure syntactic variety, e-rater counts the number of complement, subordinate, infinitive, and relative clause and occurrences of modal verbs to calculate ratios of these syntactic features per sentence per essay (Rudner & Gagne, 2001). Coniam (2009b) suggests that NLP in language assessment will be widely used by educational researchers in the future.

In February 1999, e-rater became fully operational within ETS's Online Scoring Network (OSN) for scoring the essays of GMAT, which is the first large-scale assessment to incorporate automated essay scoring. The AES system is a tool to compliment, but cannot replace expert human raters (Rudner, Garcia, & Welch, 2006). For each rating application, e-rater is trained on a collection of 270 essays that have been previously scored by trained human raters (after, Valenti et al. 2003). The OSN sends examinees' essays to two or more ETS human raters to be scored in the usual manner on the new essay topic. Once a sufficiently large sample of human scored essays accumulates, a scoring model for that essay topic is created in the automated model builder of e-rater. With the introduction of e-rater, the overall ETS essay grading system for GMAT is changed so that each essay is scored by the engine and one human rater who can score creative and unexpected essays in this kind of high-stakes assessment. All essays receive at least two readings and require - one human, one machine - and will require a third human rater to resolve any discrepancy that differs by more than one point (Kukich, 2000).

IntelliMetric

In 1998, Vantage Learning developed the AES IntelliMetric engine (Rudner et al., 2006). The IntelliMetric is based on a blend of Artificial Intelligence (AI), Natural Language Processing (NLP) and statistical technologies. It is essentially a learning engine that internalizes the characteristics of the score scale through an iterative learning process. The IntelliMetric technology parallels processes of holistic scoring and human raters. For example, human raters are trained to be prompt-specific, and this AES system is able to create a solution for each stimulus prompt. It is capable of analyzing English into seventy-seven semantic, syntactic, and discourse level features in five different categories: focus and unity, development and elaboration, organization and structure, sentence structure, mechanics and conventions (Elliot, 2003).

The IntelliMetric engine uses a multi-stage procedure to score essay-type responses after training on about 300 human scored essays on any particular prompt

(Warschauer & Ware, 2006). In the first step, the IntelliMetric internalizes the known score points of a set of responses. Then, this AES system is tested against a smaller set of responses with known scores that aid in validation and generalizability of the system. Once these are confirmed, the system is used to score new essays whose scores are unknown. Essays are targeted if they are evaluated to be atypical with regards to the standards previously set by the essay scoring (Shermis & Barrera, 2002).

From January 2006, ACT Inc. became responsible for GMAT test development and scoring, and using the IntelliMetric system. Then, a study was conducted by Rudner, Garcia and Welch in 2006 to evaluate the results of the IntelliMetric on scoring the essays of the GMAT Analytic Writing Assessment. The results of the study showed that the Pearson *r* correlations of agreement between human raters and the IntelliMetric averaged .83, with a range of .67 to .94. With regard to the 78 fabricated responses deliberately planted into the calibration sets, the IntelliMetric correctly identified every instance of fabricated essays involving copying, i.e., those in the copied prompt, repeated paragraphs, and repeated prompt half genuine categories, but the issue of off-topic responses was not fully assessed in the study. The Graduate Management Admission Council (GMAC) uses the IntelliMetric system against a human rater, not as a primary scoring system in the assessment, but to validate scores provided by human raters; thus, the issue of off-topic essays is not a problem because the human rater will flag the problem responses (Rudner et al., 2006).

Bayesian Essay Test Scoring sYstem (BETSY)

The Bayesian Essay Test Scoring sYstem (BETSY) was funded by the United States Department of Education and was developed by Lawrence Rudner at the University of

Maryland located in the city of College Park in 2002 (Coniam, 2009b). BETSY is a non-commercial Windows-based programme and is computationally intensive. Moreover, it is freely downloadable and useable for researchers (after, Valenti et al. 2003). It requires a minimum of 200 scripts for training the system on a particular topic (Dikli, 2006).

It operates on Bayesian principles, determining the probability of certain features in an analysed essay being either complete, partially complete or incomplete and denote the essay score as Appropriate, Partial and Inappropriate respectively (Rudner & Liang, 2002). The underlying models for text classification adopted are Bernoulli models. With the multivariate Bernoulli model, each essay is viewed as a special case of all the calibrated features, and the probability of each score for a given essay is computed as the product of the probabilities of the features contained in the essay. With the Bernoulli model, the conditional probability of presence of each feature is estimated by the proportion of essays within each category that contain the feature. This model can require a long time to compute since every term in the vocabulary needs to be examined. Both the Bernoulli model and the multivariate Bernoulli model are considered naive Bayes models because they assume conditional independence (after, Valenti et al. 2003). After a set of parameters is specified, the system evaluates an essay for style and content by calculating into the number of words, average sentence length, number of verbs, specific words and phrases and other characteristics including the order in which concepts appear and the occurrence of certain noun-verb pairs. Each essay is considered as a sample of calibrated features such as stemming, stop words and feature selection. Stemming involves analyzing words for key content stems and stop words refer to the most frequent articles, pronouns and prepositions in English. Then, the feature selection involves identification of items with ‘maximum potential information’ (Coniam, 2009b).

Coniam (2009b) conducted the study of the validation of the automated scoring engine BETSY on ESL student writing scripts, the results showed good correlations (averaged .84, with a range of .83 to .86) between a calibrated score based on two good human raters’ scores and BETSY on the English language writing performance of Grade 11 (Secondary 5) ESL students in a high-stakes assessment, HKCEE.

Automated Essay Scoring (AES) Systems in this Research

In the current research, two AES systems are to be used to rate English essay writing: the chief research focus is on the performance of a new AES device, the Lexile Analyzer; the IEA is used to provide assessment based on a currently widely adopted AES technology.

Intelligent Essay Assessor (IEA)

The Intelligent Essay Assessor (IEA) was originally developed at the University of Colorado in the late 1990s and purchased by Pearson Education (Warschauer & Ware, 2006). It is an Internet-based AES system for scoring the quality of the conceptual content of essays based on the Pearson’s Knowledge Analysis Technologies (KAT) engine. IEA includes the implementation of Latent Semantic Analysis (LSA), a statistical language learning theory and computer model that measures the semantic similarity of words and documents by analyzing large bodies of relevant text. After autonomous learning from a large body of representative text, it is used to rate the adequacy of content of expository essays (Pearson, 2009a). The LSA of IEA is one of the engines in the field of essay assessment that incorporates a number of NLP methods to provide an overall approach to scoring essays and providing tutorial feedback. Also, the scoring of the LSA engine focuses primarily on the conceptual content, the knowledge conveyed in an essay, rather than grammar or mechanics. The LSA of IEA can provide scoring as the system stores various sample essays including pre-scored essays of other students, expert model essays and knowledge source materials, and internal comparison of an unscored set of essays (Landauer, Laham, & Foltz, 2003). LSA makes no use of word order since this is not the most important factor for getting the sense of a passage. In other words, IEA can recognize and score different words accurately describe a concept equally, even though the sentences share no words in common (after, Valenti et al. 2003).

Landauer, Laham, and Foltz (2003) claimed that student essays are scored by the LSA to detect the meaning of the contained words and compared with essays of known quality on relevant ideas and concepts. In addition, the engine of LSA scores a wide range of content-based essays as well as creative narratives, making it possible to use IEA as a tool to evaluate complex content about the analysis of a work about, for example, art. When the Florida Gulf Coast University (US) redesigned a General Education course entitled Understanding the Visual and Performing Arts, the faculty members decided to use IEA to score two short student essays in the course for reducing the cost of essay rating. Although the IEA system could not provide direct written feedback to students, the faculty concluded that the AES system provided a significantly more reliable scoring mechanism than did humans and that it was essential to use technology in appropriate places in teaching (Wohlpart, Lindsey, & Rademacher, 2008).

The training and calibration for the IEA engine requires obtaining about 300 or more representative essays for that topic/prompt that have been scored by humans, ideally by two independent human raters, and then by a third human rater whenever the first two did not agree. The scores from these human raters serve as input for training IEA. The engine of IEA assesses the content of an essay by comparing it against a set of essays previously scored by human readers. IEA assigns a score to each essay based in part on the similarity of the content of the essay to the training set essays (Pearson, 2009a). The software is most effective with very narrowly prescribed prompts and with essays that are between 100 and 500 words in length (Wohlpart et al., 2008). 30,000 essays were collected for a study of IEA, which was conducted by a US independent testing organisation. These written responses were scored on a number of traits including ideas, organization, conventions, sentence fluency, word choice and voice (Pearson, 2009a). The system gave a mathematical representation of the relations among words and passages by statistical computations which could be applied to large numbers of essays. Whether a given word does and does not appear in the aggregate of words provides constraints that determine the similarity of two words' meaning (Landauer et al., 2003). Moreover, IEA can be tuned to understand and evaluate text in any subject area, and includes built-in detectors for off-topic responses or other situations that might need to be referred to human raters (Pearson, 2009b).

Landauer et al. (2003) conducted an investigation of the inter-rater reliability of human essay raters by using IEA. The analysis was based on 3,296 essays in total: 2,263 from standardized tests and 1,033 from classroom tests. Each essay was rated by at least two human raters. The average score of each rater pair was used as the final score with a third rater to rate the essay in the event of the first two raters' disagreeing by more than one point. Then, the LSA system was trained using the resolved score of the raters. The results of the investigation revealed that IEA then generated its own scores while accurately matching the scores as determined by the human raters.

IEA has been used to assess essays in a variety of academic, professional and employment training domains, ranging from story and letter writing through biology and history to military leadership to medical patient interviews and legal knowledge tests (Pearson, 2009a), but it is not suitable to assess factual knowledge (after, Valenti et al. 2003).

The IEA system also claims to detect plagiarism in academic papers if two essays are very similar in the use of synonyms, paraphrasing or summarizing, particularly when grading large numbers of essays (Landauer et al., 2003). On the other hand, there are currently some limitations in using the LSA system to score essays. It does not directly analyze syntax, grammar, style or mechanics and is not applicable some particular genres, such as rhymes and poems.

The Lexile Framework for Reading and Writing

The Lexile Framework for Reading was developed over 20 years by MetaMetrics, a private educational measurement company in the United States. It is an assessment system that aims to determine precisely a student's level of reading comprehension and to match those students to texts of appropriate reading difficulty. The Lexile measure is obtained through analyzing the readability of a piece of text, the text is 'sliced' into standard 'paragraphs' of 125 words or to the end of the sentence. The software programme specially designed to evaluate the reading demand of text, which is based on two strong predictors of how difficult a text is to comprehend, word frequency and sentence length (MetaMetrics, Inc., 2008b). A word-frequency study by Carroll, Davies, and Richman (1971) examined the occurrences of words in a corpus of 5,088,721 words sampled from a variety of genres in a broad range of high school materials. They had some findings on the frequency of words. Approximately one-fourth of all spoken words do not appear in print, but spoken and printed words that are used in written and oral communication provide the best means of inferring the likelihood, that a spoken word would become part of that individual's receptive vocabulary. While sentence length has been shown to be a powerful proxy for the syntactic complexity of a passage, an important caveat is that sentence length is not the underlying causal influence (Chall, 1988). These two variables are used to operationalize the syntactic and semantic variables in the Lexile theory (MetaMetrics, Inc., 2008a). The Lexile Framework for Reading then uses a modification of the Rasch model in which a match is made between student reading ability and text difficulty on the same development scale.

The Lexile reading measure appears to be a useful starting point for teachers, parents and students to manage reading comprehension and to encourage students' reading progress. Its emphasis is to determine student reading ability directly, not to provide some sort of proxy for year or grade level - the most ubiquitous alternative reading level indicator. Students receive Lexile measures from performances on calibrated standardized tests and other assessment tools to help provide a more scientific basis for the choice of appropriate books from the libraries, bookstores and the on-line Lexile book database. That database contains more than 115,000 English and Spanish fiction and non-fiction titles available from more than 450 publishers. 80 million articles and 60,000 Web sites also have Lexile measures. Text measures in Lexiles (L) typically range from 0L to 1800L, but they can go below zero, for 'Beginning Reader' (BR), and above 2000L. The Lexile Framework for Reading is a tool for making decisions about where to place students and how to help them make the transition across different reading programmes. The Lexile measures of students are the indices to identify reading materials which they can comprehend with 75% accuracy. It is valid only for continuous prose, the measure does not fit poetry or other non-continuous prose, such as recipes, menus, or shopping lists (Stenner, 1999). More than 28 million Lexile measures are reported from reading assessment and classroom programmes every year, the measures are used at the school level in all 50 states of the United States (Smith, 2009).

The key focus for the current research is the Lexile Framework for Writing, a system based directly on that for Reading, developed for monitoring student growth in writing ability and for differentiating writing instruction on the basis of those results. The Lexile Framework for Writing was launched in 2007, and estimates students' status and growth in writing ability on the same Lexile scale as reading ability to monitor student performance in both reading and writing. It provides a scale to supplement reading assessment results that are derived from analytical and holistic scoring rubric systems. While scoring rubrics are useful for providing feedback related to critical areas of writing quality, the Lexile Framework for Writing supplements rubrics to describe status and growth in writing ability (Smith, 2009).

The Lexile Framework for Writing takes the holistic scoring process further in three crucial areas of the developing writer (Swart, Kroening, & Sanford-Moore, 2008):

a. writer ability, defined by the complexity of the words written (i.e., semantic) and sophistication of word combinations (i.e., syntactic) to communicate a meaningful message to an often unseen audience;
b. convention ability, illustrated by the ability to use the rules of written English language (i.e., capitalization, grammar and punctuation rules, and spelling errors); and

c. device fluency, the speed and accuracy with which students can produce authentic text either with a pencil or pen or keyboard. (p.7)

The Lexile Analyzer

MetaMetrics developed the Lexile Analyzer which is a newly developed component of the Lexile Framework for Writing. The Lexile Analyzer is a new AES engine which provides the results which are the focus of the current research. It was used to estimate writing ability, one of the key areas of developing writers in the Lexile Framework for Writing, for establishing Lexile writer measures. Unlike other AES systems, this free online scoring device does not need human-rated sample essays for training purposes because it is based on the same algorithm which is the basis of estimating student reading ability. The Lexile Analyzer system estimates the abilities of students to express language in writing based on factors related to semantic complexity (the frequency of words used) and syntactic sophistication (the combination of words into sentences). It is usual for students to receive the Lexile measure and instant feedback upon the electronic completion of an essay when they write to narrative, informative or persuasive prompts at Grades 3 to 12. The measure is independent of the topic, genre and punctuation use (MetaMetrics, Inc., 2008b).

The Rasch Model for Measurement

The Rasch model (Rasch, 1980) is a mathematical representation probabilistic of the relation between item difficulty and person ability. One aim of using the Rasch model is to provide social scientists with the means to produce genuine interval level measures and to monitor the adherence of those scales to scientific measurement principles, so that Rasch estimates of ability and difficulty become the data for statistical analysis. The basic Rasch model follows from a small set of assumptions. The basic Rasch assumptions are that a) each person is characterized by an ability, and b) each item by a difficulty which c) can be expressed by numbers along one line. Finally, d) from the difference between the numbers and the probability of observing any particular scored response can be computed (Bond & Fox, 2007).

The formula of the Rasch model can be expressed as follows:

$$\log [P_{nix} / P_{ni(x-1)}] = \mathrm{B}_n - \mathrm{D}_i$$

where

P_{nix}	=	the probability of person n having a correct response (x) to item i;
$P_{ni(x-1)}$	=	the probability of person n having an incorrect response (x-1) to item i;
B_n	=	the ability of person n; &
D_i	=	the difficulty of item i.

Each item is characterized by a difficulty estimate (D_i) and each person by ability estimate (B_n). The item difficulty (D_i) and person ability (B_n) estimates are then expressed on a scale of log odd ratios or logits. McNamara (1996) explained that the logit scale has the advantage that it is an interval scale, so it can tell the researcher not only that an item is more difficult than another, but also how much more difficult it is. The average item difficulty is conventionally set at 0 logit, with students positive logit scores have higher probabilities of success on items of that difficulty (Bond & Fox, 2007).

The Rasch model (Rasch, 1980) prevails over traditional psychometric analysis of measurement by transforming raw data into equal-interval measures that describe a single construct and provide for the theoretical independence of item difficulty and person ability scores from the particular samples which are used to calibrate them. Moreover, the analysis of the Rasch model calibrates persons and items on a unidimensional scale to emphasise the focus on one attribute or dimension at a time (Bond & Fox, 2007).

The estimation procedure in the analysis of the Rasch model has two distinct phases: calibration of the difficulties and abilities; and then estimation of fit. Rasch suggested the use of chi-square fit statistics to determine how well any set of empirical data met the requirements of his model. Rasch analysis programmes usually report fit statistics as two chi-square ratios, infit and outfit mean square statistics (Wright, 1984; Wright & Masters, 1981). These fit indicators along with the analysis of residual patterns give researchers to uncover the variable or have better control over the data collection. As a rule of thumb, values of infit and outfit mean squares between .70 and 1.30 are generally regarded as acceptable. Those values greater than 1.30 are termed underfitting and those less than .70 as overfitting. Underfitting performances are those that should prompt researchers to reflect on what went wrong and overfitting performances might mislead researchers into concluding that the quality of the measures is better than it really is (Bond & Fox, 2007).

In this research, the Rasch model is used to develop the essay scoring scales for AES systems and for human raters on essay writing. In the second stage of the research, the scales of AES engines were compared with the scales of human raters on the data set from students in the United States as well as for a small sample in Hong Kong.

Many-Facets Rasch Measurement (MFRM)

The Many-facets Rasch model (Linacre, 1989; Lunz & Linacre, 1998) is an extension to the Rasch model for fundamental measurement which is routinely used for writing and other human-judged assessment. Linacre (1989) claimed that the MFRM was appropriate for a wide range of performance assessment and Engelhard (1994, 1996) applied the model in high-stakes language performance tests to examine the quality of rater judgments of student essay writing ability. The application of Many-facets Rasch measurement to evaluate the validity of language performance assessment is valuable because of the complexity of factors related to the interaction of students, raters and prompts introduced by the performance requirement. Moreover, the MFRM has simply made it possible to routinely expose the extent of disagreement between raters and has the potential to help raters achieve self-consistency (McNamara, 1996).

Linacre (1989) designed a typical Many-facets Rasch model for an essay examination with three facets, the writing ability of the student (B), the difficulty of the prompt (D), the severity of the rater (C), and the difficulty of the rating scale step (F). Thus, the formula of the measurement model is:

$$\log (P_{nijk}/P_{nij(k-1)}) = \mathrm{B}_n - \mathrm{D}_i - \mathrm{C}_j - \mathrm{F}_k$$

where

P_{nijk} = the probability of student *n* being awarded on prompt *i* by rater *j* a rating of k;

$P_{nij(k-1)}$ = the probability of student n being awarded on prompt i by rater j a rating of k-1;

B_n = the ability of student n;

D_i = the difficulty of prompt i;

C_j = the severity of rater j; &

F_k = the difficulty of the step up from category k-1 to category k and k=1, M.

The student facet (B_n) is a measure of writing ability on a linear equal interval logistic scale in logits. The writing task (D_i) is the calibration of the difficulty of the writing tasks on the scale and provides an estimate of the relative difficulty of each writing task that should remain invariant over students and raters. The rater facet (C_j) is the calibration of the severity of the raters involved in the scoring process, which might influence the location of student ability estimates, and the rating scale step difficulties (F_k) are produced.

In the conceptualization of the Many-facets Rasch model, the human raters are regarded as independent experts who apply their understanding to rate student essay writing, rather than as behaving in a machine-like manner. The MFRM potentially solves an important problem of essay scoring in that raters can be linked together across common essays to reveal the differences in rater severity. Connectivity (i.e., linkage) among human raters is needed to ensure that the effects of rater severity on measures can be estimated. The data analysis system must account for various genres of writing to determine their effect on prompt difficulty and the dimensionality of the writing construct. Thereby, MFRM analyses will examine the fairness and consistency of raters and adjust students' final essay scores for the measured differences in prompt difficulty and rater severity in the essay writing process.

The MFRM is different from true score-based statistics. It does not require the potentially very costly procedure of obtaining a complete data set; only a sufficient density of data is required to permit the calibrations and to detect rater severity effects (Bond & Fox, 2007). Wright and Stone (1979) and Linacre (1997) claimed that the only requirement on the rating rotation roster is that there is enough linkage to create a network between all elements of all facets, so all parameters can be estimated without indeterminacy within one frame of reference. Of course, the saving in rating costs needs to be balanced against the cost of the consequent low measurement precision.

Linacre (1997) displayed three rating plans for ratings from the Advanced Placement Program of the College Board. The complete rating plan of 1152 ratings (32 students x 3 essays x 12 raters) required that every rater scored every essay. Then, the elements of essays, students and human raters could be compared directly and unambiguously with every other element. In extreme circumstances, a minimal rating plan can be devised in which each essay is scored only once: thirty-two students have written 3 essays. Each essay was rated by only one rater, and each rater rated 8 essays, including 2 or 3 of each essay type. In this minimal rating plan, the Rasch measures are less precise than those with complete data because 83% fewer observations are made. Linacre (1997) and Linacre and Wright (2002) claimed that (less precise) Rasch estimates can still be obtained because the facet-linking overlap is maintained.

Lunz, Wright, and Linacre (1990) conducted a study in which 15 histology slides submitted by 217 examinees in a US certification clinical examination were scored by 18 raters. For the rating plan, each rater scored each of the 15 slides, and each rater had at least one examinee in common with each of the other raters. They reported the successful implementation of this minimal rating plan in which every rater did not score every slide. In the MFRM analysis, rater severity was adjusted to improve the fairness of student measures and the subsequent pass-fail decisions because the raw scores favour student performances graded by lenient raters. The MFRM analysis demonstrates that items and raters can be calibrated to define a common linear variable and that student measures for this variable can be corrected for variation in rater severity. Moreover, the MFRM separates the estimates of rater severity, student ability and item difficulty and implies that each set of estimates is independent of the distributions of the other. In fact, the MFRM is similar to an analysis of variance, and its log-linear form produces linear estimates that are not affected by missing data.

In an illustrative application of the MFRM, Engelhard (1992, 1994) examined rater severity and other aspects of writing ability assessment in the high-stakes Eighth Grade (English language) Writing Test in the state of Georgia. There were 82 highly trained human raters to grade 1000 scripts selected from a particular cohort of examinees in 1992. All the essays were rated by two raters on the four-category analytic scoring scale with five scoring domains including content/organization, style, sentence formation, usage and mechanics. The rating pattern used to estimate student writing ability in the research was 10 ratings per student, i.e. 2 raters x 5 domains = 10 domains. The results of rater severity revealed that 60 out of 82 raters are spread across more than a 3.5 logits severity range on the scale even though rigorous formal training was provided to the raters. The variation in rater severity could still have a remarkable effect on student writing assessment scores even though remarkable training efforts had been made. The MFRM model was used to reveal differences in rater severity that have the potential to severely confound the writing ability estimates of these eighth-grade students (Bond & Fox, 2007). In a real essay writing assessment, the truth is that qualified examinees may fail due to unlucky encounters with severe raters or that unqualified examinees may pass due to lucky encounters with lenient raters (Lunz, Wright, & Linacre, 1990).

In this research, the double-marking rotation design was applied to have each essay rated by two teachers of English from Hong Kong and/or four professional raters from the United States, so each rater had at least one student in common with another rater. The MFRM's approach to monitoring the rater severity effect provides for the allocation of double-marking that is consonant with the calculation of rater effect estimations (Bond & Fox, 2007).

In the analysis of the Many-facets Rasch model, the fit statistics evaluate the suitability of the data used to construct the variable and its measures (Wright & Masters, 1982). The fit statistics are indices used to estimate the extent to which responses show adherence to the modelled expectations (Bond & Fox, 2007). Two kinds of fit statistics are reported as two chi-square ratios: infit and outfit mean square statistics (Wright, 1984). The infit statistic is an information weighted sum difference between observed and expected observations, which focuses on irregular inlying patterns. The outfit statistic is the usual unweighted estimates of the degree of fit of responses, which is particularly sensitive to outlying deviations from expectation. The infit and outfit statistics are reported as mean squares in the form of chi-square statistics divided by their degrees of freedom. The expected value for the infit and outfit mean squares is 1 (Bond & Fox, 2007; Lunz, Wright, & Linacre, 1990; Wright & Masters, 1982). The range for reasonable infit and outfit mean squares for rater behaviour is between .5 and 1.5 (Wright & Linacre, 1994). In this research, the range of acceptable fit adopted was a slightly more rigorous mean squares between .6 and 1.4. If items or persons with infit mean square values are greater than the acceptable values, they are considered noisy or misfitting. This indicates more erratic performance than the Rasch model predicts. On the other hand, if items or persons with infit mean square values are less than the acceptable values, they are considered muted or overfitting. That indicates less variability in the data than the Rasch model predicts (Bond & Fox, 2007, pp 312). Moreover, the fit statistics can be used to identify misfitting observations for diagnostic purposes and corrective actions can be taken when needed (Engelhard, 1992; Lunz, Wright, & Linacre, 1990). For example, misfitting items can be identified and deleted, misfitting raters can be identified for retraining or excused from the rating process and unfairly scored student performances can be identified by the infit and outfit statistics (Lunz, Wright, & Linacre, 1990).

The investigation of the effectiveness of the Lexile Analyzer in measuring a student's essay writing ability is the central focus of this research. It is necessary to undertake the estimation of rater, genre and rubric effects by applying a Many-facets Rasch model. After this application, the co-calibration of scorings among human raters, the Lexile Analyzer and IEA, will be undertaken. The next chapter will more fully outline the use of the Rasch model for this research.

Summary

Essay writing is a common performance assessment to demonstrate students' practical command of skills acquired during English language learning. In Hong Kong, the consequence of performance on high-stakes essay assessment impacts directly on university entrance. Thus, the issues of validity and reliability in essay marking and its potential problems are a source of concern for the stakeholders. The Hong Kong Examinations and Assessment Authority (HKEAA) is diligent in its pursuit of ensuring the maintenance of quality of marking in Hong Kong high-stakes public examinations by employing double marking, Onscreen Marking (OSM) and the Rasch model analysis.

On the other hand, numerous studies have been conducted on the reliability of Automated Essay Scoring (AES) systems over the last forty years since the first AES engine was produced in 1966. A number of different AES systems are currently in the market including PEG, e-rater, IntelliMetric and BETSY. A new developed AES system, the Lexile Analyzer and the widely accepted AES system in the US, IEA are used to score student essays in this research. The Many-Facets Rasch Measurement (MFRM) is a tool to analyse the data sets and provide results on the influence of human raters, AES systems and prompt characteristics on student ability estimates in English essay writing.

Research Questions

The discussion presented in this research sequentially describes the process to investigate how effective a new AES system, the Lexile Analyzer, is in measuring essay writing ability of the US and HK students as scored by trained, professional human raters from the US and HK and another AES system, IEA. The Many-facets Rasch Measurement (MFRM) is used to compare the psychometric properties of the ratings of student essays, the writing performance of students and the difficulty of the prompts.

Key Question 1:

Does the Lexile Analyzer perform well in assessing US students' English essay writing?

In order to answer that it is necessary to find out the extent to which the Rasch model can be used to develop:

a. a MFRM analysis of the US human scoring of the US student essays; and
b. a comparative scale for scoring of the Lexile Analyzer and IEA calibrated against the US students' MFRM judged scores by the US raters.

Key Question 2:

Does the Lexile Analyzer perform well in assessing HK students' English essay writing?

In order to answer that it is necessary to find out the extent to which the Rasch model can be used to develop:

a. a MFRM analysis of the HK human scoring of the HK student essays; and

b. a comparative scale for the Lexile Analyzer scoring calibrated against the HK students' MFRM judged scores by the HK raters.

Key Question 3:

Does the Lexile Analyzer perform well in assessing the selected sub-sample of the US students' English essay writing?

In order to answer that it is necessary to find out the extent to which the Rasch model can be used to develop:

a. a MFRM analysis of the US human scoring of the selected sub-sample of the US student essays;

b. a comparative scale for scoring of the Lexile Analyzer and IEA calibrated against the selected sub-sample of the US students' MFRM judged scores by the US raters;

c. a MFRM analysis of the HK human scoring of the selected sub-sample of the US student essays;

d. a comparative scale for scoring of the Lexile Analyzer and IEA calibrated against the selected sub-sample of the US students' MFRM judged scores by the HK raters;

e. a MFRM analysis of the US and the HK human scoring of the selected sub-sample of the US student essays; and

f. a comparative scale for the scoring of the Lexile Analyzer and IEA calibrated against the selected sub-sample of the US students' MFRM judged scores by the US and HK raters.

Consequential research questions:

4. Can these scales be used to compare:
   a. human scoring vs. scoring of the Lexile Analyzer of the essays of US and HK students?
   b. human scoring of the essays of US and HK students vs. the IEA scoring of the essays of US students?
   c. human scoring of the essays of US and HK students vs. the Lexile Analyzer scoring of the essays of US and HK students vs. the IEA scoring of the essays of US students?

5. What is the extent to which the MFRM can be used to detect and adjust for rater effects in human essay scoring?

6. Do/could the Lexile Analyzer and IEA provide meaningful error-detection feedback? If so, how do they detect/report? If not, can any suggestions be made?

Chapter Three

Methodology

To address the research questions outlined in Chapter Two, the research design focuses on structuring data for the Rasch model to develop scales to compare the psychometric properties of the ratings for the US and HK student essays performed by two AES systems, the Lexile Analyzer and IEA, and the US and HK human raters. The second part of this research is to apply the Many-facets Rasch Measurement (MFRM) to adjust the prompt difficulty for the student essays and to detect rater severity in the holistic scoring of the student essays. Success on those two sets of analyses should allow for the human and machine rated essay scores to be placed on a single linear interval measurement scale. Quantitative methods were applied in this research to investigate the effectiveness of a new AES system, the Lexile Analyzer, in measuring essay writing ability as scored by trained professional human raters.

Context

There are two sets of essay writing data used in this research: a set from the United States and a second from Hong Kong. The US essays were written by Grades 4, 6, 8, 10 and 12 students based on 18 different prompts. That research design required each student to respond to between 3 and 6 prompts, each prompt administered on a different day, each US essay scored by 4 paid independent experienced raters from an established US testing company. For the HK data, each essay was scored by 2 human raters from the HK primary school. Each essay in a sub-sample of US essays was scored by 4 US and 2 HK human raters. All variables including potentially important facets such as student, prompt, rater, day, grade and genre were included in the analytical design so that the effect of the facets could be measured and partialled out of the final essay score using the Rasch model.

The set of student essay writing data from the United States was provided by the MetaMetrics Inc. This private educational measurement company also paid for the human ratings and the IEA measures for the US essays and the Lexile Analyzer measures for both the US and HK essays.

This researcher has served as an advisor to teachers of English in a government funded research project in Hong Kong for about three years. She has supported the professional development of teachers implementing self-directed learning and assessment for learning in the English language subject at secondary and primary levels. She has a good working relationship with the principals and teachers of the schools. One of the schools, a prestigious school in the northwestern district of the New Territories of Hong Kong, accepted the invitation to become the HK partner school in this research. The school has previously joined various educational programmes, co-operating with the government, universities and other school bodies. The teachers in the school are regularly offered professional development opportunities in a range of academic areas. In 2010, the school was accredited the Quality Assessment Management Accreditation Scheme (QAMAS) by the Hong Kong Examinations and Assessment Authority (HKEAA). The QAMAS is awarded in recognition of the professionalism of the school. The accreditation is based on the training and professional development activities of the staff, the experiences of the examination personnel at the school and the physical provision of examination venues. Thus, the school that is so accredited is considered to have the capacity to provide quality examinations and assessment management. In this research, the principal, the deputy principal, the English teachers and the teaching assistants of the Hong Kong partner school provided ongoing support for HK data collection and processing, while four experienced English language teachers, including the deputy principal and the English panel chairperson, were the human raters for all of the Hong Kong student essays and for a selected sub-sample of the US student essays.

Participants

Students in the United States

Data from the US students consist of 3534 essays from 589 Grades 4, 6, 8, 10 and 12 students in one school district in north-central Mississippi of the United States, which were collected in November and December 2005.

Table 3.1 Demographic data of students in the United States

	Gender		Ethnicity					
Grade	Female	Male	White	African American	Asian	Hispanic	Other	Unknown
4 (n=141)	62	79	62	64	2	5	0	8
6 (n=129)	71	58	66	55	1	3	2	2
8 (n=129)	69	60	62	61	0	5	0	1
10 (n=92)	56	36	55	36	0	1	0	0
12 (n=98)	49	49	65	30	0	2	1	0
Total (N=589)	307	282	310	246	3	16	3	11

The sample consisted of 307 females and 282 males. The classes involved in this research were 7 Grade 4 classes, 7 Grade 6 classes, and one class each from Grades 8, 10, and 12. The Grades 4 and 6 students were from two elementary schools, the Grade 8 students were from one middle school, and the Grades 10 and 12 students were from one high school. All of the students were administered between four and six writing prompts for the essays in this research. After removing 81 problematic essays including 78 blank essays and 3 essays without IEA measures, 3453 essays were available for the final stage of analysis in this research. The details of the student demographics in the United States are summarised in Table 3.1.

A Selected Sub-sample of Students in the United States

A sub-sample of 150 essays from 50 students in the United States was later selected randomly as best matched to the levels of essay writing in the samples of Hong Kong students. These essays used the same prompts that were used in Hong Kong samples and allowed for the calibration of a sub-group that includes human scoring of US essays by Hong Kong teachers.

Students in Hong Kong

The data from the HK students were 408 essays written by 137 Grade 6 Hong Kong students who were from one local primary school, a government-subsidised primary school located in the New Territories of Hong Kong. There were 67 females and 70 males in the sample. Six of the total sample are students with Special Educational Needs (SEN), identified as having learning difficulties or speech and language problems during a territory-wide early identification and intervention programme that was conducted by the educational psychologists of the Hong Kong Education Bureau in Grade 1 (HKEDB, 2007b). The teachers' professional judgments suggest that 21 students in this sample pool had low English language proficiency and 8 students had extremely low English language proficiency. The HK students were administered the three writing prompts as part of their usual examination preparation.

Instruments

Prompts for Essay Writing by Students in the United States

Students responded to between four and six essay writing prompts that covered the narrative, informative and persuasive genres (USDOE, 1998a) (Appendix 1). Human raters scored the written responses against a holistic rubric providing four to six independent ratings for each essay.

Over a three-week period in the fall of 2005, students in each grade responded to a total of 6 prompts, writing 2 responses per week. The 18 prompts used in this research were released for public use on the website of the National Assessment of Educational Progress (USDOE, 1998b; USDOE, 2002). According to the NAEP Overview (USDOE, 2009), the National Assessment of Educational Progress (NAEP) is the only nationally representative and continuing assessment for the US students from Grade 4 to Grade 12 in various subject areas. It is carried out by the Commissioner of Education Statistics, who heads the National Center for Education Statistics in the United States Department of Education (USDOE). The assessment is periodically conducted in mathematics, reading, writing, science, the arts, civics, economics, geography and the US history. NAEP provides results on subject-matter achievement, instructional experiences, and the school environment for populations of students and sub-groups within those populations. The NAEP results are based on representative samples of students at Grades 4, 8 and 12 for the main assessment, or samples of students from 9 to 17 years of age for the long-term assessment.

The 18 prompts were selected for three grade levels (4, 8, 12), and were then randomly assigned to versions A and B, each version being assigned a narrative (N), an informative (I), and a persuasive (P) prompt (Appendices 2 and 3). Students were given a total of 30 minutes to respond to each prompt, 5 minutes to plan and 25 minutes to write.

Table 3.2 Prompt design by grade

Grade	Prompts					
4	4AN	4AI	4AP	4BN	4BI	4BP
6	4BN	4BI	4BP	8AN	8AI	8AP
8	8AN	8AI	8AP	8BN	8BI	8BP
10	8BN	8BI	8BP	12AN	12AI	12AP
12	12AN	12AI	12AP	12BN	12BI	12BP

Note: A and B are two different prompt indicators. N, I, and P stand for narrative, informative, and persuasive respectively.

Table 3.2 displays the prompt design by grade. The prompt design includes two extra grade levels, 6 and 10, which were included to provide linkage across the grades, each grade taking three prompts from the grade above and another three prompts from the grade below. The linkage prompts were used to construct a linked series for the essay writing scoring system. Each of the students in Grades 4, 6, 8, 10 and 12 was administered four to six prompts in the research. 524 students (89%) responded to 6 prompts, 49 students (8.3%) responded to 5 prompts and 16 students (2.7%) responded to 4 prompts. Three prompts (4AN, 4AI and 4AP) were only written by Grade 4 students, three (4BN, 4BI and 4BP) were written by both Grades 4 and 6, three (8AN, 8AI and 8AP) by Grades 6 and 8, three (8BN, 8BI and 8BP) by Grades 8 and 10, three (12AN, 12AI and 12AP) by Grades 10 and 12 and three (12BN, 12BI and 12BP) by Grade 12 only.

Table 3.3 Details of the essays by students in the United States

Grade	4						8		
				(6)					
Prompt	4AN	4AI	4AP	4BN	4BI	4BP	8AN	8AI	8AP
Female	58	61	58	61 *(71)*	61 *(70)*	60 *(70)*	*(71)* 67	*(69)* 64	*(71)* 65
Male	76	78	75	77 *(57)*	75 *(58)*	76 *(58)*	*(57)* 58	*(58)* 55	*(56)* 57
Unknown	0	0	0	0	0	0	*(0)* 1	0	*(0)* 1
Subtotal	134	139	133	138 *(128)*	136 *(128)*	136 *(128)*	*(128)* 126	*(127)* 119	*(127)* 123
Total	134	139	133	266	264	264	254	246	250

Grade	8			12						Total
	(10)									
Prompt	8BN	8BI	8BP	12AN	12AI	12AP	12BN	12BI	12BP	
Female	66 *(56)*	68 *(55)*	67 *(55)*	*(54)* 49	*(55)* 49	*(55)* 49	49	49	46	1799
Male	59 *(34)*	58 *(36)*	57 *(36)*	*(36)* 49	*(35)* 49	*(36)* 49	49	48	47	1649
Unknown	1 *(0)*	1 *(0)*	1 *(0)*	0	0	0	0	0	0	5
Subtotal	126 *(90)*	127 *(91)*	125 *(91)*	*(90)* 98	*(90)* 98	*(91)* 98	98	97	93	3453
Total	216	218	216	188	188	189	98	97	93	3453

Table 3.3 shows that 133 to 139 essays and 93 to 98 essays were written by Grade 4 and Grade 12 students respectively, and each pair of grades contains a maximum of 266 and minimum of 188 essays. There are a total of 3453 student essays available for developing scales in this research, 1799 essays (52.1%) were written by female students and 1649 essays (47.8%) were written by male students and 5 essays (0.1%) were written by one student whose gender is unknown.

Table 3.4 Distribution of prompts by day within grade

Class	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6
1	AN	AI	AP	BN	BI	BP
2	BN	BP	BI	AN	AP	AI
3	BP	BI	BN	AP	AI	AN
4	AP	AN	AI	BP	BN	BI
5	AI	AP	AN	BI	BP	BN
6	BI	BN	BP	AI	AN	AP

Note: In Grades 6 and 10, the lower grade prompt was treated as the A prompt in the randomization design.

To avoid ordering effects, students were randomly administered the prompts and each class responded to a different prompt on a given day, as displayed in Table 3.4. The performance of students should then not be affected by a learning effect.

Prompts for Essay Writing by Hong Kong Students

137 Grade 6 HK students responded to three writing prompts in the Grade 4 narrative, informative and persuasive genres from the National Assessment of Educational Progress (NAEP) over a period of four weeks in May and June 2009, as part of their usual examination preparation. They were given a total of 30 minutes to respond to each prompt, 5 minutes to plan and 25 minutes to write.

Table 3.5 Distribution of prompts by day

Class	Day 1	Day 2	Day 3
1	4BN	4BI	4BP
2	4BI	4BN	4BP
3	4BP	4BI	4BN
4	4BN	4BP	4BI

Table 3.6 Details of the Hong Kong student essays

Prompt	4BN	4BI	4BP	Total
Female	67	67	67	201
Male	70	69	68	207
Total	137	136	135	408

Three prompts were selected for the US students at the Grade 4 and Grade 6 levels (Appendix 4). Each of the Grade 6 student was administered the prompts of 4BN, 4BI and 4BP in random order to avoid ordering effects in this research, as displayed in Table 3.5. 134 students (97.8%) responded to 3 prompts, 3 students (2.2%) responded to 2 prompts. 201 of the 408 total essays (49.3%) were written by female students and 207 essays (50.7%) by male students, as presented in Table 3.6.

Scoring Procedures

McNamara's seminal text (1996, p.3) claimed that language performance assessment involves judgments and quality ratings against a rating scale. He introduced new features for the assessment setting including the following:

a. the raters themselves, who will vary the standards they use and the consistency of their application of those standards; and
b. the rating procedures they are required to implement.

Holistic rubrics were used to score the essays in this research. Scoring guides followed the NAEP testing design and the prompts in each grade level had unique rubrics (Appendix 5). In other words, all narrative essays were scored using a particular set of rubrics with criteria that were different from those used for informative or persuasive essays. The genres of the essays in Grades 4, 8 and 12 were the same. The overall essay ratings included 1- unsatisfactory, 2 - insufficient, 3 - uneven, 4 - sufficient, 5 - skillful and 6 - excellent.

Here is a holistic scoring rubric for Grade 4 informative prompt:

Score	Description
(6) Excellent	● Develops ideas well and uses specific, relevant details across the response. ● Well organized with clear transitions. ● Sustains varied sentence structure and exhibits specific word choices. ● Exhibits control over sentence boundaries; errors in grammar, spelling, and mechanics do not interfere with understanding.
(5) Skillful	● Develops ideas with some specific, relevant details. ● Clearly organized; information is presented in an orderly way, but response may lack transitions. ● Exhibits some variety in sentence structure and exhibits some specific word choices. ● Generally exhibits control over sentence boundaries; errors in grammar, spelling, and mechanics do not interfere with understanding.

(4) Sufficient	● Clear but sparsely developed; may have a few details. ● Provides a clear sequence of information; provides pieces of information that are generally related to each other. ● Generally has simple sentences and simple word choice; may exhibit uneven control over sentence boundaries. ● Has sentences that consist mostly of complete, clear, distinct thoughts; errors in grammar, spelling, and mechanics generally do not interfere with understanding.
(3) Uneven	May be characterized by one or more of the following: ● Provides limited or incomplete information; may be list-like or have the quality of an outline. ● Disorganized or provides a disjointed sequence of information. ● Exhibits uneven control over sentence boundaries and may have some inaccurate word choices. ● Errors in grammar, spelling, and mechanics sometimes interfere with understanding.
(2) Insufficient	May be characterized by one or more of the following: ● Provides little information and makes little attempt at development. ● Very disorganized OR too brief for reader to detect organization. ● Exhibits little control over sentence boundaries and sentence formation; word choice is inaccurate in much of the response. ● Characterized by misspellings, missing words, incorrect word order; errors in grammar, spelling, and mechanics are severe enough to make understanding very difficult in much of the response.
(1) Unsatisfactory	May be characterized by one or more of the following: ● Attempts a response, but may only paraphrase the prompt or be extremely brief. ● Exhibits no control over organization. ● Exhibits no control over sentence formation; word choice is inaccurate across the response. ● Characterized by misspellings, missing words, incorrect word order; errors in grammar, spelling, and mechanics severely impede understanding across the response.

Experienced raters from the United States and Hong Kong respectively scored the student essays in this research. The US student essays were all scored by nineteen raters with extensive essay scoring experience from an established United States testing company employed for that purpose by MetaMetrics Inc. Each US student essay was scored by four independent raters.

The Hong Kong student essays and a selected sub-sample of the US student essays (i.e., those best matched to the levels of essay writing of the Hong Kong samples) were scored by four experienced English teachers from the partner local primary school in Hong Kong. The Hong Kong raters all have ten to sixteen years of English language teaching experience at the elementary school level. Their mother tongue is Cantonese and they have all learnt English as a second language. One has a Master's degree in English Language and three received post-graduate training in Teaching English Language at the local universities. One is the school English panel chairperson and one is the deputy principal at the school. All of the Hong Kong raters have considerable experience in scoring essays with holistic and analytic scoring rubrics.

Before scoring the student essays, the Hong Kong raters were provided with a 5-hour training workshop by this researcher to establish a common understanding of the prompts and the criteria of the holistic scoring rubrics and to develop agreement on the given ratings. A brief description of the scoring rubrics was given by the researcher at the beginning of the workshop and 10 sample essays in the narrative, informative and persuasive genres from the US and HK student essays were rated in the workshop. This researcher was involved in the discussion with the four raters to minimise any differences between their ratings on those sample essays. The goal of the training workshop was not only to obtain agreement on the scores of sample papers and on the usefulness of the scoring rubrics, but also to help the raters internalise every criterion of the scoring scale by combining descriptions with examples. This intensive training on the criteria allowed the raters to score accurately and quickly by only referring to the scoring rubrics or the sample paper occasionally (White, 1984). Each essay written by a HK student was scored by two HK raters against the appropriate holistic NAEP rubric.

After scoring the students' essays, the Hong Kong raters met with this researcher for their essay scoring. The four raters were provided 5 essays with a maximum discrepancy of 4 out of 6 points on the scale between two of them. They rescored the essays, and answered several questions about the scoring process and individually reflected on the scoring system in the HK education field.

In 2006, two AES engines, the Lexile Analyzer and IEA, were used to score the transcriptions of English essays which were written by the students in the United States. In 2010, the transcriptions of Hong Kong student essays were scored for this research by the Lexile Analyzer only. Although the MetaMetrics Inc. tried to have the HK essays scored using IEA, that company declined the invitation to be paid to do so.

Rating Plans

Although the ideal way to have a complete data set would be to have every rater score every student on every essay, it is too expensive and time-consuming to use this rating plan. Linacre (1997) claimed that the Rasch estimates could be calculated, as long as overlap is maintained. The key requirement of the rating plan is that there is enough linkage between all of the elements of all the facets, so that all of the parameters can be estimated without any indeterminacy within that one frame of reference. A reduced network of observations can retain the connection between examinees, raters and items. The parameters are linked into one frame of reference which shares pairs of parameters (common essays, common examinees or common raters). One constraint is that each rater must avoid scoring any one type of essay too frequently (Linacre, 1989). In this research, the pairings of raters for each essay were changed and the repeated pattern was applied in the plans for both the United States and Hong Kong essays.

For the data set for students in the United States, every essay was scored by four of the nineteen independently paid trained raters. 16 out of 19 raters scored 607 to 1100 essays, and 3 of them scored 18, 19 and 287 essays. In total, 13812 essay writing scores for the students in the United States were produced for the research analysis. For the data set of Hong Kong student essays, each of four raters scored between 203 and 206 of the 408 essays in the Hong Kong sample. Additionally, each Hong Kong rater scored 75 essays in the sub-sample of the US student essays so that a greater number of common essays, common examinees and common raters could be established for building the scales for the research. This design of the rating plan provided a connected network that linked the students, prompts and raters to enable all measures and calibrations estimated from the observations to be placed on a common scale (Linacre & Wright, 2002).

Informed Consent

The researcher approached the principal and the deputy principal of the Hong Kong partner school for approval to collect the student essay writing data. Then, four teachers of English, including the English panel chairperson and the deputy principal, were invited to be raters for this research. The teachers understood the aims of the research and took part in the research voluntarily on the understanding that their names would be kept strictly confidential. They were also reminded that they could withdraw from the research project at any time without explanation or prejudice. All of the raters received an Information Sheet and signed an Informed Consent Form (Appendix 6) and the principal signed the Data Use Agreement (Appendix 7) for the research.

Data Collection

The essay writing data set from the students in one school district in the north-central Mississippi of the United States was provided by the MetaMetrics Inc. to allow for the construction of comparative scales. The researcher was given the signed Data Use Agreement (Appendix 7) for the MetaMetrics data. The scanned student essays and their transcriptions were provided in a text file, with all of the students' names removed to keep their identities confidential.

The relatively small set of essay data from the Hong Kong students was collected by the researcher from a local primary school. These essays were used to develop the Rasch comparative scales and to implement the MFRM to adjust for prompt difficulty and to detect the severity of the Hong Kong human raters. Four classes of Grade 6 students were selected to be the participants in the research. Because Hong Kong students learn English as a second language, the three prompts (4BI, 4BN and 4BP) administered to them were the same three prompts used for Grade 4 and Grade 6 students in the United States. The HK partner school and this researcher prepared the transcriptions of essays in a text file and the measures given by the Lexile Analyzer for the Hong Kong student essays were provided by the MetaMetrics Inc. in May 2010.

The design of the Hong Kong student essay data provided a convenient opportunity for the researcher to compare the performance of the students in the United States and Hong Kong on three particular prompts. A sub-sample of essay writing from the students in the United States was then selected as the best matched to the levels of essay writing of the sample of Hong Kong students. This subset of the US student essays were then scored by the four English teachers from the Hong Kong partner school.

In other words, there was a small set of essay data that had been scored by two AES systems and by human raters from the United States and Hong Kong. The data set for the MFRM analysis was therefore structured to analyse for prompt difficulty, rater severity and whole-instrument replication so that an appropriate assessment of the validity and precision of the writing assessment could be performed.

Table 3.7 Data summary

Essay writings	Students in the United States				Hong Kong students	
Month/Year	2006	2006	Nov & Dec 2005	Oct 2010	May 2010	Nov 2009
Rated by AES/human raters	Lexile Analyzer	IEA	Human raters in the United States	Human raters in Hong Kong	Lexile Analyzer	Human raters in Hong Kong
Grade 4	●	●	●	○		
Grade 6	●	●	●	○	●	●
Grade 8	●	●	●			
Grade 10	●	●	●			
Grade 12	●	●	●			

○ Human raters in Hong Kong rated a selected sub-sample of essays from the US data set.

Data Analyses

Software

The software package used for the orginal Rasch analyses was WINSTEPS Version 3.70.0.4 (Linacre, 2010d). WINSTEPS is a software programme used for many applications of the Rasch model, particularly in the areas of educational testing, attitude surveys and rating scale analysis. WINSTEPS is designed to construct Rasch measurement from the responses of a set of persons to a set of items. WINSTEPS has the data capacity for up to 1,000,000 persons and 30,000 items and allows each item to have a rating scale of up to 255 categories (Linacre, 2010b).

The software package FACETS Version 3.67.0 was used for the Many-facets Rasch Measurement (MFRM) analysis (Linacre, 2010c). The computer programme Facets implements MFRM in the areas of performance assessment and paired comparisons. Facets provides measure (linear quantity), standard error (precision) and fit statistics (statistical validity), and calibrations of response format structures, such as rating scale, partial credit items, letter grades and rank. Facets can quantify discrepant interactions between the elements of the different facets. Once measures have been estimated from a data set, differential facet functioning, equivalent to differential item functioning or item bias, can be investigated automatically. A rater's bias on one item, or an item's bias against a group of persons can be identified and its size and statistical significance can be estimated. Facets is ideally suited for essay grading, portfolio assessment and other kinds of judged performances (Linacre, 2010a).

Model

In the current research, the six defined facets for the analysis: the writing ability of the student (B), the difficulty of the prompt (D), the severity of the rater (C), the severity of the AES systems (A), and the difficulty of the rating scale step (F). Because a holistic rubric (rather than analytic) was used to score the essays, the domain facet is not appropriate in this model (Linacre, 1989). Thus, the formula of this measurement model is:

$$\log (P_{nijklmx} / P_{nijklm(x-1)}) = B_n - D_i - C_j - C_k - A_l - A_m - F_x$$

where

$P_{nijklmx}$	=	the probability of student n being rated, on prompt i by rater j and/or rater k and AES l and/or AES m, a rating of x;
$P_{nijklm(x-1)}$	=	the probability of student n being rated, on prompt i by rater j and/or rater k and AES l and/or AES m, a rating of $x-1$;
B_n	=	the writing ability of student n;
D_i	=	the difficulty of prompt i;
C_j	=	the severity of rater j from the United States;
C_k	=	the severity of rater k from Hong Kong;
A_l	=	the severity of the Lexile Analyzer;
A_m	=	the severity of IEA; &
F_x	=	the difficulty of the rating step up from category $x-1$ to category x and x=1, M.

In this Many-facets Rasch model, facets of the student essays can be modeled and their effects on the scoring of the raters and the AES systems are estimated. There is no limit to the number of facets that can interact to produce a rating, but each facet should provide a particular meaning in the research (Linacre, 1989). The elements of each facet are summarised by their measure mean, their standard deviation, the reliability of element separation, and the corresponding chi-square for homogeneity (Lunz, Wright, & Linacre, 1990).

Because the essays in this research were scored by two groups of human raters and two AES engines, the design of this Many-facets Rasch model is fairly typical and has an individual facet for each group of raters (C_j & C_k) and for each AES system (A_l & A_m) to produce more meaningful bias analyses between the facets. The bias analyses provide further analysis of the residuals to demonstrate whether additional sub-patterns emerge in the data sets and whether the data have significant potential for use in this research on essay assessment. Moreover, the study of bias/interaction effects reveals unexplained random variation in the analyses (McNamara, 1996).

In this research, the results of the ultimate MFRM analysis of all student essays with scoring of the US and HK human raters and two AES systems, the Lexile Analyzer and IEA, were produced following a series of MFRM analyses:

a. US student essays with scoring of US human raters;
b. US student essays with scoring of US human raters, the Lexile Analyzer and IEA;
c. HK student essays with scoring of HK human raters;
d. HK student essays with scoring of HK human raters and the Lexile Analyzer;
e. the selected sub-sample of the US student essays with scoring of US human raters;
f. the selected sub-sample of the US student essays with scoring of US human raters, the Lexile Analyzer and IEA;
g. the selected sub-sample of the US student essays with scoring of HK human raters;
h. the selected sub-sample of the US student essays with scoring of HK human raters, the Lexile Analyzer and IEA;
i. the selected sub-sample of the US student essays with scoring of US and HK human raters;
j. the selected sub-sample of the US student essays with scoring of US and HK human raters, the Lexile Analyzer and IEA;
k. all student essays including US student essays, HK student essays and a sub-sample of the US student essays with scoring of US and HK human raters

and the Lexile Analyzer; and, finally,

l. all student essays including US student essays, HK student essays and a sub-sample of the US student essays with scoring of US and HK human raters and IEA.

Ethics

All human interactions, including the interactions involved in human research, have ethical dimensions (National Health and Medical Research Council [NHMRC], 2007). According to the James Cook University Statement and Guidelines on Research Practice (2001), researchers at the University who conduct research involving any form of human participation must have approval from the James Cook University Ethics Review Committee along with the Human Research Ethics Committee to ensure that ethical and legal responsibilities are appropriately addressed in the research.

This research has been approved by the Human Research Ethics Committee of the James Cook University on 7 October 2009 (Approval Number H3477, Appendix 8).

Chapter Four

Results

The prime purpose of this research was to investigate the effectiveness of a new AES system, the Lexile Analyzer, for measuring student essay writing ability as scored by trained human raters. The scales for the scoring of the Lexile Analyzer and IEA of the US and HK students' essays were developed by applying the rigorous measurement principles of the Rasch model. The results produced by the Many-facets Rasch Measurement (MFRM) are presented in this chapter. The scales of the comparison reveal how a new AES system, the Lexile Analyzer, produces the scores for student essays in this research.

Categorisation of Raw Scores of the Lexile Analyzer and IEA

Because the rating scale analysis of the Rasch model using WINSTEPS software provides for up to 255 response categories, and the range of scores for each of the Lexile Analyzer and IEA is greater than 2,000, this is far more than the number of categories that can be accommodated by WINSTEPS. Accordingly, a Poisson logarithmic transformation was applied to the data to ensure functional rating scale categorisation in this research. This transformation is used as an early step in the effort to construct scales for the scores of the AES systems with WINSTEPS and to create a more reasonable range of category responses (Wright & Masters, 1982).

The transformation can be expressed as

$$\text{Scored category} = 1 + n * \frac{\log(\text{observation}+1) - \log(L+1)}{\log(H+1) - \log(L+1)} \qquad (1)$$

where L is the lowest value of the observations, and H is the highest value of the observations. For example, '8' was chosen as n when the intended transformation target was to a 9-category structure.

This Poisson logarithmic transformation separated the observations into intervals, with all observations in the interval classified into the same category level (Yan, 2009). Thus, the raw scores of the student essays were first divided into 50 categories, an progressively reduced to 9 categories, with the intention of providing more even distributions and more meaningful category structures at the first stage of analyses in this research.

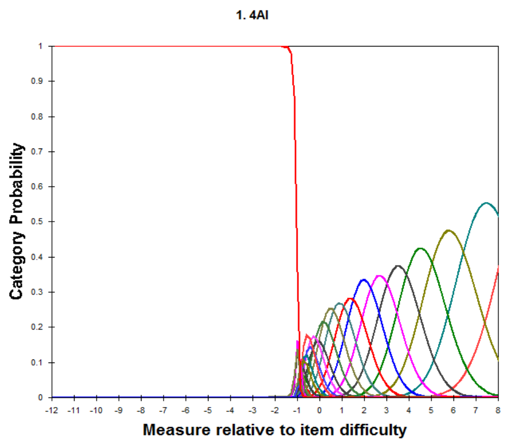


Figure 4.1 Category Probability Curves for the scores of the Lexile Analyzer of US student essays on the 4AI prompt (50 categories)

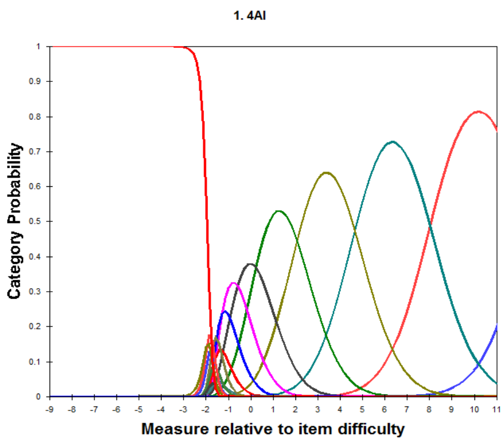


Figure 4.2 Category Probability Curves for the scores of the Lexile Analyzer of US student essays on the 4AI prompt (23 categories)

However, the analysis results in Figure 4.1 for the 50 categories after the Poisson logarithmic transformation have 86% of all responses in categories 40 to 47 (i.e., only eight categories). Categories 1 to 39 and categories 48 to 50 are completely collapsed and have no distinct peaks, implying that, practically, these categories are redundant (Appendix 9). Categories with low frequencies are problematic because they do not provide enough observations to allow for an estimation of stable threshold values (Bond & Fox, 2007; Linacre, 2002).

The results with 23 categories (shown in Figure 4.2) or 9 categories for the Poisson logarithmic transformation are similar. It is impossible to find peaks in some categories, and they are totally submerged by others, i.e., some categories were rarely performed by the students in this sample. The observations in each category are not even (Appendix 9).

Linacre (1999, 2002) stated that a regular uniform distribution across categories should provide satisfactory Rasch estimates. The transformation of the raw data using the Lexile Analyzer from 50 categories to 9 categories was conducted on a trial basis, and similar results were obtained for the transformation of the scores using IEA for the US student essays. Thus, it is necessary to reconstruct the category structure, combining and splitting categories in the analysis.

Collapsing of Categories

Because the best categorisation might not be obtained directly from raw data, even after applying the process of the Poisson logarithmic transformation, Bond and Fox (2007) suggested that the collapsing of categories can be used to investigate the best match between diagnostics and respondent use, with the optimal categorisation being used in the final phase of the instrument development.

Linacre (1999, 2002) provided guidelines for combining categories to improve overall measurement quality. According to the following guidelines for collapsing categories, a uniform frequency distribution should be created to produce a measurement process under the condition of an equal contribution for each category (Linacre, 1995, 1999 & 2002):

Preliminary Guideline: All items oriented with latent variable;

Guideline #1: At least 10 observations of each category;

Guideline #2: Regular observation distribution;

Guideline #3: Average measure advance monotonically with category;

Guideline #4: Outfit mean square less than 2.0;

Guideline #5: Step calibrations advance;

Guideline #6: Ratings imply measures, and measures imply ratings;

Guideline #7: Step difficulties advance by at least 1.4 logits; and

Guideline #8: Step difficulties advance by less than 5.0 logits.

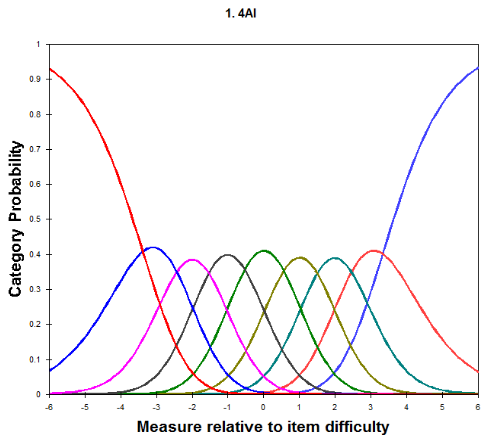


Figure 4.3 Category Probability Curves for the scores of the Lexile Analyzer of US student essays on the 4AI prompt (new 9 categories)

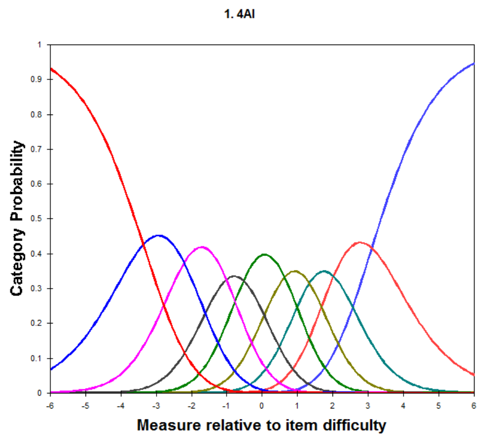


Figure 4.4 Category Probability Curves for the scores of IEA of US student essays on the 4AI prompt (new 9 categories)

Figures 4.3 and 4.4 present the results from the re-categorisation of the AES data into the 9 new categories. The category probability curves for the scores of the Lexile Analyzer and IEA of the US student essays show that the categories are evenly distributed and have their own distinct peaks, displaying the overall quality required for analyses. All categories fulfilled the guidelines for collapsing categories (Appendix 10), and this optimal categorisation was used in the following analyses of the Many-facets Rasch Measurement (MFRM).

Sample Essays in this Research

Some student essays are reproduced here in order to illustrate the products which the students provided during their usual examination preparations for data collection of this research. Students were given 30 minutes in which to freely express ideas instructed by the prompts. Some sample essays from both the United States and Hong Kong are included below.

a. The 4BN prompt

One day you wake up and go down to breakfast. You eat what you normally eat. Your breakfast is the last normal thing that happens to you all day. Write a story called 'The Very Unusual Day' about what happens that day, from right after breakfast until you go the bed again.

The following essay was written by a HK Grade 6 male student, Student HK135, it was given a very high score of 6 by HKR002 and HKR003, and the Lexile Analyzer measure is 1301.80 which is the highest measure in the HK student essays.

The Very Unusual Day

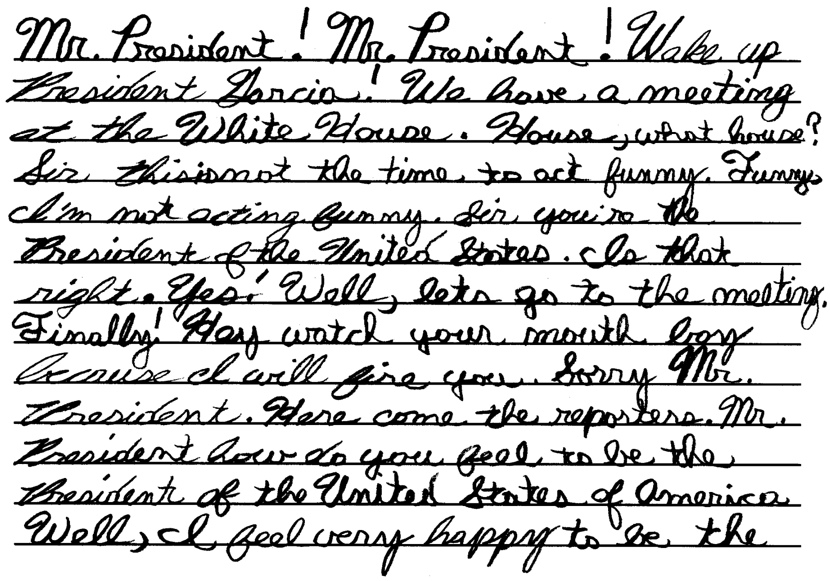
DreeeeeeeeeeEEEEEEEE.....

'Okay, I'm waking up...' I mumbled, pressing out the switch on the alarm clock. That was the start of another boring day ~~an~~ at the ~~orphanage~~ government 'emergency' ~~cen~~ shelter. My name's Michael, Michael Yeung Wang Pong. My parents died in a car crash. Right now I'm considered to be an illegal ~~immigrant~~ immigrant just because my ID card was burnt to ashes. Typical English ignorance. And now they're offering me an 'emergency' shelter that seems more like a prison cell. I'm only allowed to go out for a walk in the afternoon. Rest of the

b. The 8BN prompt

Imagine that you wake up one morning to discover that you have become the President of the United States. Write a story about your first day as President.

A Hispanic Grade 8 male student in the sample of US data, Student US086, wrote the following essay. It was given a score of 3 by USR101 and USR102, and a score of 2 by USR104 and US105. The Lexile Analyzer and IEA measures of this essay are 331.29 and 61 respectively.



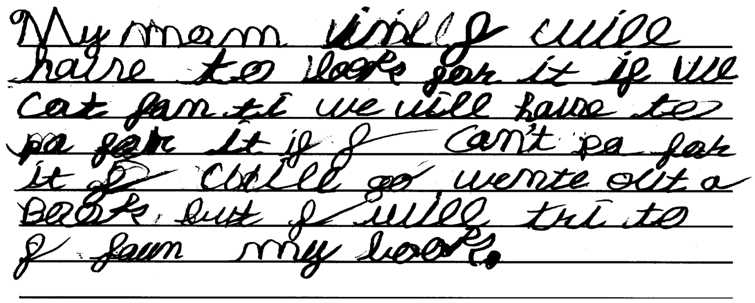
Mr. President! Mr. President! Wake up
President Garcia! We have a meeting
at the White House. House, what house?
Sir, this is not the time to act funny. Funny,
I'm not acting funny. Sir, you're the
President of the United States. Is that
right. Yes! Well, lets go to the meeting.
Finally! Hey watch your mouth boy
because I will fire you. Sorry Mr.
President. Here come the reporters. Mr.
President how do you feel to be the
President of the United States of America
Well, I feel very happy to be the

c. The 4BI prompt

Imagine this situation: Your favorite book is missing from your school library. It might be a book that you like to read over and over again. Or it might be a book that your teacher or parent has read to you. Some of your friends also like to read this book. The school librarian is not sure she wants to buy the book again. Write a letter to convince your school librarian to buy the book again. In your letter, give lots of reasons why the book should be is in your school library.

Some students from both the United States and Hong Kong responded to the prompts with words that did not make complete sense. Below is an example.

Student US256, an African American Grade 4 male student, wrote the following essay. It was given a consistently low score of 0 by USR201, USR206, USR207 and USR210. The Lexile Analyzer measure is -182.76, and the IEA measure is 1.


My mom will I will
haire to look for it if we
Cat fam. ti we will haive to
pa for it if I can't pa for
it I will go wente out a
Break, but I will tri to
I fam my look.

The MFRM Analyses Results of US Student Essays

Key Question 1:

Does the Lexile Analyzer perform well in assessing US students' English essay writing?

In order to answer that it is necessary to find out the extent to which the Rasch model can be used to develop:

a. a MFRM analysis of the US human scoring of the US student essays; and
b. a comparative scale for scoring of the Lexile Analyzer and IEA calibrated against the US students' MFRM judged scores by the US raters.

In the analysis of the US human scoring of the US student essays using the Many-facets Rasch Measurement (MFRM), the results are satisfactory. Eighteen prompts and nineteen US human raters fit the Rasch model very well, although 85 out of 589 US students show misfit in the analysis, with the range of the infit mean squares between 1.41 and 5.6 (Appendix 11).

```
+-----------------------------------------------------------------------------------------------------------------------+
|Measr|-Prompt               |-US Rater                                        |-Lexile Analyzer|-IEA|+Students   |R6SCA|
|-----+----------------------+-------------------------------------------------+----------------+----+------------+-----|
|   4 +                      +                                                 +                +    +            + (6) |
|     |                      |                                                 |                |    | .          |     |
|     |                      |                                                 |                |    |            |     |
|     |                      |                                                 |                | 1  |            |     |
|   3 +                      +                                                 +                +    + .          + --- |
|     |                      |                                                 |                |    | .          |     |
|     |                      |                                                 |                |    | .          |     |
|     | 13-12AI  15-12AP     |                                                 |                | 2  | *.         |     |
|   2 + 16-12BI              +                                                 +                +    + *.         +     |
|     | 9-8AP                |                                                 |                |    | ***.       |     |
|     | 17-12BN              |                                                 |                | 3  | ****.      |  4  |
|     | 11-8BN               | US200                                           | 1              |    | ******     |     |
|   1 + 18-12BP              +                                                 + 2              + 4  + *******.   +     |
|     | 14-12AN  12-8BP      | US209  US211                                    |                |    | *********. |     |
|     |                      | US103  US104  US206                             | 3              |    | ******     |     |
|     | 10-8BI   8-8AN       |                                                 | 4              |    | *********  |     |
*   0 * 7-8AI                * US102  US105  US201  US203  US204  US205  US210 * 5              * 5  * ********.  * --- *
|     |                      | US101                                           | 6              |    | ******.    |     |
|     |                      | US202  US207  US208                             |                |    | ******.    |     |
|     |                      | US106                                           | 7 8            | 6  | ***.       |     |
|  -1 +                      +                                                 +                +    + **.        +     |
|     |                      |                                                 | 9              |    | **.        |     |
|     | 5-4BN                |                                                 |                | 7  | *          |  3  |
|     |                      | US108                                           |                |    | *.         |     |
|  -2 +                      +                                                 +                +    + .          +     |
|     | 4-4BI   2-4AN   6-4BP|                                                 |                | 8  | .          |     |
|     |                      |                                                 |                |    | .          |     |
|     | 1-4AI                |                                                 |                |    | .          | --- |
|  -3 + 3-4AP                +                                                 +                +    +            +     |
|     |                      |                                                 |                | 9  | .          |     |
|     |                      |                                                 |                |    | .          |     |
|     |                      |                                                 |                |    |            |     |
|  -4 +                      +                                                 +                +    +            +     |
|     |                      |                                                 |                |    |            |  2  |
|     |                      |                                                 |                |    |            |     |
|     |                      |                                                 |                |    |            |     |
|  -5 +                      +                                                 +                +    +            +     |
|     |                      |                                                 |                |    |            |     |
|     |                      |                                                 |                |    |            |     |
|     |                      |                                                 |                |    |            | --- |
|  -6 +                      +                                                 +                +    + .          +     |
|     |                      |                                                 |                |    |            |     |
|     |                      |                                                 |                |    |            |     |
|     |                      |                                                 |                |    |            |     |
|  -7 +                      +                                                 +                +    +            + (1) |
|-----+----------------------+-------------------------------------------------+----------------+----+------------+-----|
|Measr|-Prompt               |-US Rater                                        |-Lexile Analyzer|-IEA| * = 7      |R6SCA|
+-----------------------------------------------------------------------------------------------------------------------+
```

Figure 4.5 Vertical map for the US student essays with scoring of US raters, Lexile Analyzer and IEA

Figure 4.5 illustrates a continuum of the difficulty of 18 prompts, the severity of 19 US raters, the Lexile Analyzer and IEA response category thresholds, and the ability of 589 US students that was constructed by analysing US student essays. In this MFRM analysis of the US student essays, column 1 shows the units of measurement on the logit scale, which extends from -7 to 4 logits. Column 7 presents the human ratings, which range from 1 to 6. The averages of prompt difficulty, rater severity and student ability have been set at 0.00 logits by default in columns 2 through 6.

In column 2, the most difficult prompt, 12AP, is located at the top of the scale (2.18 logits). Grade 10 and Grade 12 students responded to this persuasive prompt, which was for Grade 12 students from NAEP. The easiest prompt is 4AP (-2.92 logits), and it was administered to Grade 4 students only. Higher measures on the scale in column 3 indicate that the raters are more severe on essay scoring. In column 3, USR200 is the most severe rater (1.25 logits), and USR108 is the most lenient rater (-1.71 logits).

In columns 4 and 5, for both the scales of the Lexile Analyzer and IEA, the lowest score is Category 1 and the highest score is Category 9 and the categories are ordered appropriately. The range of logit measures of the Lexile Analyzer is between -2 and 2, and the IEA category measures range from -4 to 4 logits.

The essay writing ability measures of 589 US students range from -7 to 4 logits (column 6), with the majority of the students' writing abilities falling between -2 and 2 logits.

Table 4.1 Calibration of the students for US student essays using US raters, Lexile Analyzer and IEA (20 of the 76 misfitting students)

Student			Observed	Measure	S.E.	Infit		Outfit	
Number		Group	score	(Logit)		MnSq	ZStd	MnSq	ZStd
21	US018	104	3.8	-0.61	0.35	1.41	1.3	1.41	1.3
110	US095	108	3.3	0.52	0.35	1.41	1.3	1.39	1.2
44	US435	106	4.3	0.84	0.33	1.42	1.4	1.39	1.3
162	US035	108	4.4	1.30	0.37	1.42	1.2	1.43	1.2
470	US147	104	3.0	-1.12	0.35	1.42	1.3	1.42	1.3
383	US586	110	4.8	2.19	0.33	1.43	1.4	1.44	1.5
635	US353	104	2.9	-0.12	0.35	1.43	1.4	1.46	1.4
195	US291	108	3.5	1.12	0.35	1.44	1.3	1.46	1.4
321	US180	106	3.3	0.24	0.35	1.44	1.4	1.44	1.4
400	US370	110	3.6	0.48	0.35	1.44	1.3	1.42	1.3
...	...	...	...	...	...	...	...	...	...
329	US299	106	3.8	0.32	0.37	2.38	3.3	2.56	3.6
336	US581	106	3.0	-1.05	0.43	2.42	2.9	2.50	3.0
630	US306	104	3.1	-1.41	0.35	2.42	3.5	2.42	3.5
396	US366	110	3.9	0.29	0.35	2.80	4.0	2.75	3.9

59	US049	104	2.6	-1.85	0.35	2.89	4.4	2.90	4.4
620	US572	104	3.2	-0.89	0.35	3.21	5.1	3.19	5.1
511	US465	106	2.3	-2.69	0.37	3.30	5.1	3.21	4.6
625	US577	104	3.0	-1.77	0.35	4.59	6.9	4.62	6.9
299	US271	104	3.2	-1.36	0.38	4.92	6.7	4.94	6.7
628	US579	104	3.8	-1.11	0.35	8.56	9.0	7.79	9.0

Separation: 2.80 Reliability: 0.89

Fixed (all same) chi-square: 4912.1 significance *p*: .00

Note: 104 = Grade 4 US student, 106 = Grade 6 US student, 108 = Grade 8 US student, 110 = Grade 10 US student, 112 = Grade 12 US student

The measurement report describing the ability of Grade 4 to Grade 12 US students yields a fixed chi-square of 4912.1 (p = .00) to show students have different writing abilities. A high reliability of separation index of .89 indicates that a line of inquiry has been developed in which some students' abilities are statistically significantly higher and some are statistically significantly lower in the analysis. The standard errors of the 589 US students' scored essays are between .32 and .61. The lowest writing ability is at -6.12 logits and the highest writing ability is at 3.81 logits, showing a wide range of 9.93 logits between Grade 4 and Grade 12 students (Appendix 12).

In this analysis, 76 misfitting students out of 589 total students are found with infit mean squares greater than 1.40 (Appendix 12). This number includes 29 Grade 4 students (38.2%), 19 Grade 6 students (25%), 13 Grade 8 students (17.1%), 12 Grade 10 students (15.8%) and 3 Grade 12 students (3.9%). The analysis results for 20 of the 76 misfitting students are shown in Table 4.1.

Table 4.2 Calibration of the prompts for US student essays using US raters, Lexile Analyzer and IEA

Prompt	Observed score	Measure (Logit)	S.E.	Infit MnSq	Infit ZStd	Outfit MnSq	Outfit ZStd

4AI	3.6	-2.80	0.07	0.84	-2.8	0.83	-3.0
4AN	3.7	-2.23	0.07	0.84	-2.7	0.90	-1.5
4AP	3.6	-2.92	0.07	1.82	9.0	1.79	9.0
4BI	3.7	-2.37	0.05	1.08	1.7	1.07	1.6
4BN	3.8	-1.60	0.05	0.74	-6.6	0.75	-5.8
4BP	3.6	-2.18	0.05	1.53	9.0	1.53	9.0
8AI	3.5	0.06	0.05	1.12	2.6	1.13	2.6
8AN	3.7	0.34	0.05	0.73	-6.8	0.72	-6.9
8AP	2.9	1.79	0.05	0.96	-0.9	0.96	-0.9
8BI	3.8	0.16	0.06	0.97	-0.6	0.97	-0.7
8BN	3.7	1.15	0.06	0.70	-6.9	0.70	-6.8
8BP	3.6	0.81	0.06	1.24	4.7	1.24	4.7
12AI	3.8	2.17	0.06	0.86	-2.8	0.85	-2.9
12AN	3.9	0.85	0.06	0.78	-4.6	0.78	-4.5
12AP	3.4	2.18	0.06	0.91	-1.8	0.91	-1.7
12BI	3.5	2.08	0.09	0.92	-1.1	0.92	-1.1
12BN	4.0	1.47	0.09	0.82	-2.7	0.81	-2.8
12BP	4.0	1.04	0.09	1.11	1.4	1.11	1.4
Separation: 27.26				Reliability: 1.00			
Fixed (all same) chi-square: 14041.7				significance *p*: .00			

In Table 4.2, the calibration for the 18 prompts is presented. The span of prompt difficulties ranges from 4AP (-2.92 logits), the easiest prompt to 12AP (2.18 logits), the most difficult prompt is 5.1 logits. The fixed chi-square is 14041.7 (p = .00) to show that the prompt difficulties are not the same. A high reliability of 1.00 indicates that some prompts are statistically significantly more difficult and some are statistically significantly easier. The standard errors range from .05 to .09 showing high precision of measurement. Two misfitting prompts are found in this analysis: 4AP (infit mn sq 1.82; outfit mn sq 1.79), and 4BP (infit and outfit mn sqs 1.53). In this research, the two persuasive prompts 4AP and 4BP are used for Grade 4 students and Grades 4 and 6 students respectively. They are for Grade 4 students in NAEP. Students are expected to write letters to convince a friend to become visible and to convince the school librarian to buy their favourite book (Appendix 3).

Table 4.3 Calibration of the US raters for US student essays using US raters, Lexile Analyzer and IEA

US rater	Observed score	Measure (Logit)	S.E.	Infit		Outfit	
				MnSq	ZStd	MnSq	ZStd
USR101	3.9	-0.27	0.06	0.78	-4.9	0.82	-4.0
USR102	3.8	-0.03	0.05	0.74	-6.4	0.74	-6.2
USR103	3.6	0.54	0.10	0.70	-3.9	0.69	-3.8
USR104	3.6	0.59	0.06	0.70	-7.2	0.72	-6.7
USR105	3.8	-0.02	0.06	0.90	-2.1	0.90	-2.1
USR106	4.0	-0.75	0.07	0.63	-7.5	0.63	-7.2
USR108	3.3	-1.71	0.40	0.61	-1.2	0.60	-1.2
USR200	3.2	1.25	0.07	1.27	4.3	1.26	4.1
USR201	3.6	-0.11	0.05	1.22	4.6	1.22	4.4
USR202	3.6	-0.44	0.07	1.19	3.2	1.18	3.0
USR203	3.6	0.03	0.05	1.17	3.8	1.17	3.8
USR204	3.7	0.04	0.05	0.95	-1.2	0.95	-1.1
USR205	3.5	0.01	0.06	1.06	1.2	1.06	1.2
USR206	3.4	0.39	0.05	0.88	-2.6	0.88	-2.6
USR207	3.7	-0.45	0.06	1.07	1.3	1.06	1.2
USR208	3.7	-0.38	0.06	1.04	0.8	1.04	0.7
USR209	3.4	0.73	0.06	1.12	2.2	1.11	2.1
USR210	3.7	-0.09	0.06	1.38	7.3	1.37	7.1
USR211	3.3	0.68	0.41	0.73	-0.8	0.69	-0.9
Separation: 4.27				Reliability: 0.95			
Fixed (all same) chi-square: 935.8				significance *p*: .00			

In Table 4.3, the severity span across US raters (USR200 to USR108) is 2.96 logits. A fixed chi-square shows the hypothesis that the raters all have the same leniency, apart from measurement error; but the fixed chi-square is 935.8 (p = .00), so this hypothesis is unlikely to be true. In other words, the raters have different severity in the analysis. A high reliability index of .95 indicates that some raters scored statistically significantly higher and some scored statistically significantly lower. The standard errors of USR108 and USR211 are .4 and .41 respectively. These are relatively high and show a low precision of measurement because these two raters only scored 19 and

18 out of 3453 essays in the data set. The 19 US raters generally interpreted the scale consistently, with infit and outfit mean squares between .61 and 1.38. All of them fit well to the Rasch model.

Table 4.4 Calibration of the Lexile Analyzer for US student essays using US raters, Lexile Analyzer and IEA

Lexile Analyzer	Observed score	Measure (Logit)	S.E.	Infit MnSq	Infit ZStd	Outfit MnSq	Outfit ZStd
Category 1	2.5	1.33	0.07	1.11	1.9	1.13	2.1
Category 2	2.9	1.00	0.05	1.11	2.3	1.14	2.9
Category 3	3.1	0.57	0.04	0.93	-2.0	0.93	-1.9
Category 4	3.4	0.29	0.04	0.99	-0.4	0.99	-0.4
Category 5	3.6	-0.04	0.03	1.04	1.3	1.03	0.9
Category 6	3.8	-0.29	0.04	0.91	-3.1	0.91	-3.1
Category 7	4.1	-0.63	0.04	0.98	-0.4	0.98	-0.5
Category 8	4.5	-0.86	0.05	0.95	-1.1	0.94	-1.3
Category 9	4.9	-1.37	0.06	1.14	2.5	1.14	2.5

Separation: 17.05 Reliability: 1.00
Fixed (all same) chi-square: 2007.2 significance *p*: .00

The calibration of the categories from the Lexile Analyzer scoring is presented in Table 4.4. The facet of the Lexile Analyzer has been set by default as negative in the model specifications, so higher catergory implies lower measure; thus, the span of categories is 2.7 logits, from 1.33 logits for Category 1, the lowest category, to -1.37 logits for Category 9, the highest category. The separation index for the Lexile Analyzer is 17.05, which corresponds to a high reliability of separation index (R=1.00); it reveals that some categories are statistically significantly higher and some are statistically significantly lower. The fixed (all same) chi-square is 2007.2 (p = .00) to show categories have different measures. The standard errors range between .03 and .07, showing a high precision of measurement. The infit and outfit statistics of Categories 1 to 9 are between .91 and 1.14, well within the reasonable mean square ranges of .7 to 1.4 for the infit and outfit statistics.

Table 4.5 Calibration of IEA for US student essays using US raters, Lexile Analyzer and IEA

IEA	Observed score	Measure (Logit)	S.E.	Infit		Outfit	
				MnSq	ZStd	MnSq	ZStd
Category 1	2.2	3.22	0.07	1.04	0.8	1.11	1.7
Category 2	2.7	2.25	0.05	0.90	-2.3	0.90	-2.4
Category 3	3.0	1.55	0.04	0.99	-0.2	0.99	-0.2
Category 4	3.3	0.94	0.04	1.04	1.2	1.04	1.2
Category 5	3.6	-0.04	0.03	0.92	-3.1	0.91	-3.2
Category 6	3.9	-0.87	0.04	0.98	-0.6	0.98	-0.5
Category 7	4.2	-1.46	0.04	1.08	2.3	1.07	2.0
Category 8	4.6	-2.32	0.05	1.05	1.3	1.05	1.3
Category 9	5.0	-3.28	0.07	1.00	0	1.01	0.1
Separation: 41.77				Reliability: 1.00			
Fixed (all same) chi-square: 12593.0				significance *p*: .00			

Table 4.5 presents the calibrations for IEA scoring. Category 1 is at 3.22 logits, and Category 9 is at -3.28 logits, showing a wide range of 6.5 logits between the lowest and highest categories. The fixed chi-square is 12593 (p = .00) to show categories have different measures. A high reliability of separation index of 1.00 indicates that some categories are statistically significantly higher and some are statistically significantly lower in the analysis. The standard errors are between .03 and .07 and show a high precision of measurement. There is no evidence of misfit, with all of the infit and outfit mean squares between .9 and 1.08, i.e., close to the expected value of 1.0 (Bond & Fox, 2007).

The MFRM Analyses Results of HK Student Essays

Key Question 2:

Does the Lexile Analyzer perform well in assessing HK students' English essay writing?

In order to answer that it is necessary to find out the extent to which the Rasch model can be used to develop:

a. a MFRM analysis of the HK human scoring of the HK student essays; and
b. a comparative scale for the Lexile Analyzer scoring calibrated against the HK students' MFRM judged scores by the HK raters.

In the MFRM analysis of the HK human scoring of 408 HK student essays, the results are satisfactory. Three prompts (4AI, 4AN and 4AP) and four HK human raters fit well to the Rasch model. However, 34 out of 137 HK students show misfit with a range of the infit mean squares between 1.41 and 7.49 (Appendix 13).

```
+------------------------------------------------------------+
|Measr|-Prompt|-HK Rater|-Lexile Analyzer|+Student    |R6SCA|
|-----+-------+---------+----------------+------------+-----|
|  11 +       +         +                +            + (6) |
|     |       |         |                |            |     |
|     |       |         |                | .          |     |
|  10 +       +         +                +            +     |
|     |       |         |                |            |     |
|     |       |         |                |            |     |
|   9 +       +         +                +            +     |
|     |       |         |                |            |     |
|     |       |         |                |            |     |
|   8 +       +         +                +            +     |
|     |       |         |                |            |     |
|     |       |         |                |            |     |
|   7 +       +         +                +            +     |
|     |       |         |                |            |     |
|     |       |         |                |            |     |
|   6 +       +         +                +            +     |
|     |       |         |                |            |     |
|     |       |         |                |            |     |
|   5 +       +         + 8              +            + --- |
|     |       |         |                |            |     |
|     |       |         |                | .          |     |
|   4 +       +         +                + .          +     |
|     |       |         |                | **.        |  5  |
|     |       |         |                | .          |     |
|   3 +       +         +                + **         +     |
|     |       |         |                |            | --- |
|     |       |         |                | **         |     |
|   2 +       +         +                + *****.     +     |
|     |       |         |                | *******    |  4  |
|     |       |         |                | *********  |     |
|   1 +       +         + 1              + ******     + --- |
|     |       | HK004   | 2              | *****      |     |
|     | 4-4BI | HK001   | 3              | *******    |     |
*   0 * 5-4BN * HK002   * 4              * *****      *  3  *
|     | 6-4BP |         | 5              | **.        |     |
|     |       |         | 6              | *.         | --- |
|  -1 +       + HK003   +                + ***        +     |
|     |       |         |                | **.        |     |
|     |       |         | 9              | *.         |  2  |
|  -2 +       +         +                + *.         +     |
|     |       |         |                | .          |     |
|     |       |         |                | **         | --- |
|  -3 +       +         +                + .          +     |
|     |       |         |                |            |     |
|     |       |         |                |            |  1  |
|  -4 +       +         +                +            +     |
|     |       |         | 7              |            |     |
|     |       |         |                |            |     |
|  -5 +       +         +                +            + (0) |
|-----+-------+---------+----------------+------------+-----|
|Measr|-Prompt|-HK Rater|-Lexile Analyzer| * = 2      |R6SCA|
+------------------------------------------------------------+
```

Figure 4.6 Vertical map for the Hong Kong student essays with scoring of HK raters and Lexile Analyzer

Figure 4.6 shows the calibrations for 3 prompts, 4 HK raters, the Lexile Analyzer and 137 HK student essays.

As indicated in column 2, the most difficult prompt is 4BI (.22 logits), and the easiest prompt is 4BP (-.31 logits), yielding a range across the three prompts of less than 1 logit.

The severity of HK raters is displayed in column 3. The most severe rater is HKR004 (.71 logits) and the most lenient is HKR003 (-1.06 logits). The range of category measures of the Lexile Analyzer is between -5 and 6 logits. However, Categories 7, 8 and 9 are in unexpected locations on the scale. Category 8 is on top of Category 1 and much higher than Category 1. Category 9 is located between Categories 6 and 7.

The essay writing ability measures of all HK students range from -4 to 5 logits in column 5, with the exception of one student who has approximately 10 logits on the scale.

Table 4.6 Calibration of the HK students for HK student essays using HK raters and Lexile Analyzer (32 misfitting students)

Student		Observed Score	Measure (Logit)	S.E.	Infit MnSq	Infit ZStd	Outfit MnSq	Outfit ZStd
733	HK094	4.0	1.28	0.53	1.43	0.8	1.41	0.8
712	HK073	2.8	0.60	0.55	1.44	0.8	1.46	0.8
734	HK095	5.0	3.60	0.59	1.48	0.9	1.42	0.8
767	HK128	2.7	0.35	0.55	1.57	1.0	1.52	0.9
713	HK074	2.0	-1.11	0.57	1.58	1.0	1.52	0.9
760	HK121	3.0	0.77	0.54	1.59	1.0	1.56	1.0
693	HK054	3.8	1.91	0.54	1.66	1.1	1.71	1.2
698	HK059	3.3	1.28	0.54	1.70	1.2	1.71	1.2
697	HK058	3.3	1.05	0.54	1.76	1.2	1.73	1.2
704	HK065	4.0	2.13	0.54	1.76	1.2	1.89	1.4
688	HK049	3.7	1.70	0.53	1.82	1.3	1.82	1.3
756	HK117	3.0	0.63	0.54	1.83	1.3	1.85	1.3
650	HK011	2.3	-0.06	0.56	1.87	1.3	1.84	1.3
716	HK077	1.8	-1.63	0.57	1.94	1.4	1.96	1.5
751	HK112	3.3	1.17	0.54	1.99	1.5	2.03	1.5
683	HK063	3.8	1.90	0.53	2.20	1.7	2.21	1.8
702	HK044	4.2	1.67	0.54	2.20	1.8	2.21	1.8

658	HK019	3.8	1.52	0.54	2.22	1.8	2.19	1.7
719	HK080	3.0	0.46	0.54	2.27	1.8	2.32	1.8
729	HK090	3.3	0.97	0.54	2.28	1.8	2.27	1.8
640	HK001	1.0	-2.68	0.75	2.30	1.6	2.12	1.5
671	HK032	4.8	1.46	0.61	2.35	1.8	2.81	2.3
701	HK078	2.7	-0.28	0.55	2.37	1.9	2.27	1.8
717	HK062	2.0	-1.59	0.57	2.37	1.9	2.43	1.9
711	HK072	1.8	-1.11	0.70	2.48	1.7	2.45	1.7
743	HK104	2.2	-0.86	0.56	2.50	2.0	2.46	2.0
687	HK048	5.2	4.30	0.63	2.55	2.0	2.98	2.4
753	HK114	3.8	1.77	0.53	2.78	2.3	2.76	2.3
723	HK084	4.3	1.99	0.67	2.84	2.0	2.73	1.9
675	HK036	3.2	-1.07	0.56	3.04	2.5	3.02	2.5
725	HK086	1.8	-1.44	0.57	5.20	4.0	4.96	3.8
660	HK021	3.8	1.25	0.53	5.67	4.3	5.76	4.4
Separation: 2.76				Reliability: 0.88				
Fixed (all same) chi-square: 871.6				significance *p*: .00				

Table 4.6 displays the measurement report describing the 32 misfitting students (infit mean squares 1.43 to 5.67). The logit measures of 136 out of 137 HK students range from -2.96 to 4.3, with outlying Student 774 (HK135; 10.46 logits), demonstrating the highest writing ability in the analysis. The standard errors of 137 HK students range between .53 and 1.21. The fixed chi-square of 871.6 (p = .00) shows that students have different writing abilities and an acceptable reliability of separation index of .88 indicates a line of inquiry has been developed where some students received statistically significantly higher scores and some received lower scores (Appendix 14).

Table 4.7 Calibration of the prompts for HK student essays using HK raters and Lexile Analyzer

Prompt	Observed score	Measure (Logit)	S.E.	Infit		Outfit	
				MnSq	ZStd	MnSq	ZStd
4BI	3.1	0.22	0.08	0.98	-0.2	0.97	-0.3
4BN	3.3	0.09	0.08	0.86	-1.7	0.88	-0.5

4BP	3.4	-0.31	0.08	1.14	1.5	1.14	0.9
Separation: 2.55			Reliability: 0.87				
Fixed (all same) chi-square: 22.8			significance *p*: .00				

Table 4.7 presents the calibration for the three prompts. The logit measure of the most difficult prompt, 4BI, is .22 logits (SE .08), and the logit measure of the easiest prompt, 4BP, is -.31 logits (SE .08). The span of difficulty is .53 logits.

The fixed chi-square is 22.8 (p = .00) and an acceptable reliability of .87 indicate that two prompts are not measurably different, and one is just measurably different. The standard errors are .08, showing a high precision of measurement in this analysis. The infit and outfit mean squares of the three prompts range from .86 to 1.14, revealing that the prompts fit very well to the Rasch model.

Table 4.8 Calibration of the HK raters for HK student essays using HK raters and Lexile Analyzer

HK rater	Observed score	Measure (Logit)	S.E.	Infit		Outfit	
				MnSq	ZStd	MnSq	ZStd
HKR001	3.0	0.38	0.10	0.86	-1.4	0.88	-1.2
HKR002	3.3	-0.02	0.10	0.81	-2.0	0.81	-0.6
HKR003	3.8	-1.06	0.10	1.37	3.3	1.39	3.5
HKR004	2.9	0.71	0.10	0.94	-0.6	0.91	-0.6
Separation: 6.87				Reliability: 0.98			
Fixed (all same) chi-square: 191.3				significance *p*: .00			

In Table 4.8, the severity span ranges from the most lenient rater, HKR004 (.71 logits) to the most severe rater, HKR003 (-1.06 logits) is 1.77 logits. The standard errors are .10. The separation index for HK raters is 6.87, which corresponds to a high reliability of separation index (R=.98) to show that some raters scored statistically significantly severer and some scored statistically significantly more lenient. The fixed (all same) chi-square is 191.3 (p = .00) to indicate that the raters do not have the same leniency. The infit and outfit statistics of four HK raters range from .81 to 1.39, revealing that the HK raters fit the Rasch model well.

Table 4.9 Calibration of the Lexile Analyzer for HK student essays using HK raters and Lexile Analyzer

Lexile Analyzer	Observed score	Measure (Logit)	S.E.	Infit		Outfit	
				MnSq	ZStd	MnSq	ZStd
Category 1	1.9	1.16	0.13	1.02	0.1	0.99	0
Category 2	2.9	0.54	0.09	1.02	0.2	1.03	0.3
Category 3	3.3	0.34	0.10	0.89	-1.0	0.91	-0.8
Category 4	3.7	0.10	0.10	0.84	-1.5	0.85	-1.4
Category 5	4.2	-0.43	0.15	1.25	1.5	1.19	0.5
Category 6	4.6	-0.81	0.27	1.28	1.0	1.27	1.0
Category 7	5.3	-4.18	0.81	3.95	2.6	4.13	2.7
Category 8	5.5	4.86	1.21	0.91	0.2	0.91	0.2
Category 9	5.8	-1.56	1.21	0.55	-0.2	0.28	3.8
Separation: 3.37			Reliability: 0.92				
Fixed (all same) chi-square: 141.3			significance *p*: .00				

The calibration of the Lexile Analyzer categories is shown in Table 4.9. The span of categories is 9.04 logits, from 1.16 logits (SE .13) for Category 1 to -.81 logits (SE .27) for Category 6, with the standard errors between .09 and .27 showing less precision of measurement.

However, Categories 7 (-4.18 logits), 8 (4.86 logits) and 9 (-1.56 logits) are in unexpected locations on the scale because the number of observations in Category 7 is only one and the number of observations in Categories 8 and 9 are two. The standard errors of Categories 7 to 9 are between .81 and 1.21 showing an inadequate number of observations in those categories.

The categories of the Lexile Analyzer generally interpreted the scale consistently with infit and outfit mean squares between .84 and 1.28, but the exceptions are Category 9 (infit mn sq .55; outfit mn sq .28) and Category 7 (infit mn sq 3.95; outfit mn sq 4.13). An infit mean square of Category 7 of the Lexile Analyzer is very large. It reveals that there is a large amount of unexpected variance (i.e., erratic performance/noise) in the observations for Lexile Category 7. It is likely that Lexile Category 7 is highly unpredictable, and that inferences based on Lexile Category 7 are very uncertain.

The separation index for the Lexile Analyzer is 3.37, which corresponds to a high reliability of separation index (R=.92) to show some categories are statistically significantly higher and some are statistically significantly lower in the analysis. The fixed (all same) chi-square is 141.3 (p = .00) to indicate that categories have different measures, except for measurement error.

Selection of a Sub-sample of the US Student Essays

In this part of the research, a sub-sample of 150 essays from 50 Grades 4 and 6 US students was selected to be scored by four HK human raters. The analyses results of student measures of Rasch model for the US students and HK students on the same three prompts were compared: the mean of the 270 Grades 4 and 6 US students' scores as 14.5 (SD=6.3), very close to the distribution of the 137 Grade 6 HK students' scores (mean=14.7; and SD=6.5) (Appendix 15). Then, 50 students were randomly chosen from that sample of 270 Grade 4 and Grade 6 US students, i.e., best-matched to the levels of HK student essay writing. The students were assigned a computer-generated integral number from 1 to 270 with an online random numbers service, www.random.org, developed by Mads Haahr in 1998 and operated by the School of Computer Science and Statistics at Trinity College in Dublin (Mara, Maskova, Fucikova, Kuzel, Belsan, & Sosna, 2008). This provides 150 essays (i.e., three per US student) to be rated by Hong Kong teachers in order to develop well-linked scales.

The MFRM Analyses Results of the Selected Sub-sample of the US Student Essays

Key Question 3:

Does the Lexile Analyzer perform well in assessing the selected sub-sample of the US students' English essay writing?

In order to answer that it is necessary to find out the extent to which the Rasch model can be used to develop:

a. a MFRM analysis of the US human scoring of the selected sub-sample of the US student essays;
b. a comparative scale for scoring of the Lexile Analyzer and IEA calibrated against the selected sub-sample of the US students' MFRM judged scores by the US raters;
c. a MFRM analysis of the HK human scoring of the selected sub-sample of the US student essays;
d. a comparative scale for scoring of the Lexile Analyzer and IEA calibrated against the selected sub-sample of the US students' MFRM judged scores by the HK raters;
e. a MFRM analysis of the US and the HK human scoring of the selected sub-sample of the US student essays; and
f. a comparative scale for the scoring of the Lexile Analyzer and IEA calibrated against the selected sub-sample of the US students' MFRM judged scores by the US and HK raters.

The MFRM analyses results of the a) scoring of US raters only (Appendix 16), b) scoring of US raters, the Lexile Analyzer and IEA (Appendix 16), c) scoring of HK raters (Appendix 17), d) scoring of HK raters, and the Lexile Analyzer and IEA (Appendix 17) and e) scoring of both US and HK human raters of a sub-sample of the US student essays are satisfactory (Appendix 18).

In these five analyses, 3 prompts fit the Rasch model very well, and 18 US human raters fit the model, with the exception of one misfitting US rater who was identified in the analyses with the scoring of the US human raters. However, 8 to 13 US students show misfit in the analyses (Appendices 16 - 18).

In the MFRM analysis of the sub-sample of the US students using the scoring by the US and HK human raters (only), the 4 Hong Kong raters showed misfit to the model with infit mean squares between 1.53 and 1.84 (Appendix 18).

```
+-----------------------------------------------------------------------------------------------+
|Measr|-Prompts      |-US Rater                 |-HK Rater |-Lexile Analyzer|-IEA|+Student   |R6SCA|
|-----+--------------+-------------------------+---------+----------------+----+-----------+-----|
|   3 +              +                         +         +                +    +           + (6) |
|     |              |                         |         |                |    |           |  5  |
|     |              |                         |         |                |    |           |     |
|     |              |                         |         |                |    |           |     |
|     |              |                         |         |                |    |           |     |
|     |              |                         |         |                |    | **        |     |
|     |              |                         |         |                |    | *         |     |
|     |              |                         |         |                |    |           |     |
|   2 +              +                         +         +                +    +           +     |
|     |              |                         |         |                |    | *         | --- |
|     |              |                         |         |                |    | **        |     |
|     |              |                         |         | 1              |    | *         |     |
|     |              |                         |         |                |    | *         |     |
|     |              |                         |         |                |    |           |     |
|     |              |                         |         |                |    | **        |     |
|     |              | US200                   |         | 2              |    | **        |     |
|   1 +              +                         +         +                + 1  + *         +  4  |
|     |              | US211                   |         |                | 2  | ***       |     |
|     |              |                         |         |                |    | ********  |     |
|     |              | US104    US209          |         |                | 4  | ****      |     |
|     |              |                         | HK001   |                |    | *         |     |
|     |              | US103                   |         |                | 3  | ****      |     |
|     | 5-4BN        | US203                   | HK004   |                |    |           |     |
|     |              | US204    US208          |         | 3 8            |    | *         |     |
*   0 *              * US105    US207          *         *                *    * **        * --- *
|     | 4-4BI  6-4BP | US206                   |         |                | 5  | *         |     |
|     |              | US102    US106          |         |                | 7  | **        |     |
|     |              | US101    US201    US210 |         |                |    | ***       |     |
|     |              |                         |         | 4 5            |    |           |     |
|     |              | US202                   | HK002   |                |    | *         |     |
|     |              |                         |         | 9              |    | ***       |     |
|     |              | US205                   | HK003   |                | 6  |           |     |
|  -1 +              +                         +         +                +    + *         +     |
|     |              |                         |         | 6              |    |           |  3  |
|     |              |                         |         |                |    | *         |     |
|     |              |                         |         |                |    | *         |     |
|     |              |                         |         |                |    |           |     |
|     |              |                         |         |                | 8  |           |     |
|     |              |                         |         |                |    |           |     |
|     |              |                         |         |                |    |           |     |
|  -2 +              +                         +         +                +    +           + --- |
|     |              |                         |         |                |    |           |     |
|     |              |                         |         |                |    |           |     |
|     |              |                         |         |                |    |           |     |
|     |              |                         |         |                |    |           |     |
|     |              |                         |         |                |    |           |     |
|     |              |                         |         |                |    |           |  2  |
|  -3 +              +                         +         +                +    + *         + (1) |
|-----+--------------+-------------------------+---------+----------------+----+-----------+-----|
|Measr|-Prompts      |-US Rater                 |-HK Rater |-Lexile Analyzer|-IEA| * = 1     |R6SCA|
+-----------------------------------------------------------------------------------------------+
```

Figure 4.7 Vertical map for a sub-sample of the US student essays with scoring of US and HK raters, Lexile Analyzer and IEA

Figure 4.7 illustrates the calibrations for 3 prompts, 18 US raters and 4 HK raters, the Lexile Analyzer and IEA, and essays of the selected sub-sample of 50 US students when all of their data are analysed together.

In column 2, it shows that the most difficult prompt is 4BN (.22 logits) in the analysis. For the severity of the 18 US raters, USR200 (1.09 logits) is the most severe rater, and USR205 (-.85 logits) is the most lenient rater. For the severity of HK raters, column 4 displays that the most severe rater is HKR001 (.44 logits), and that the most lenient rater is HKR003 (-.86 logits).

The range of logit measures of the two AES engines is between -2 and 2. However, some unexpected locations of categories are displayed on the scales of the Lexile Analyzer and IEA. In column 5, Category 8 of the Lexile Analyzer is located very close to Category 3, and Category 9 is between Categories 4 and 6. In column 6, Category 7 is between Categories 5 and 6 on the scale of IEA.

The logit measures of the 50 US students range from -3 to 3 in column 7, with a slightly top-heavy distribution.

Table 4.10 Calibration of the students for a sub-sample of US student essays using US and HK raters, Lexile Analyzer and IEA

Student			Observed	Measure	S.E.	Infit		Outfit	
Number		Group	score	(Logit)		MnSq	ZStd	MnSq	ZStd
7	US006	304	3.9	0.73	0.23	0.46	-2.8	0.45	-2.9
15	US013	304	4.3	1.28	0.23	0.79	-0.9	0.81	-0.8
*20	US017	304	1.5	-1.34	0.27	1.57	2.0	1.51	1.7
25	US021	304	3.1	-0.32	0.23	0.88	-0.4	0.88	-0.4
27	US022	306	3.2	-0.67	0.24	0.56	-1.9	0.55	-2.0
32	US026	306	3.9	0.39	0.23	0.55	-2.2	0.55	-2.2
34	US027	306	4.8	1.92	0.24	1.21	0.9	1.15	0.7

*38	US030	306	3.5	0.64	0.24	1.70	2.4	1.68	2.3
*43	US034	306	3.3	0.74	0.23	1.66	2.3	1.66	2.3
56	US046	304	2.7	-0.15	0.22	0.72	-1.4	0.73	-1.3
*60	US049	304	2.7	-0.74	0.23	1.48	1.9	1.40	1.6
*69	US057	304	3.5	1.14	0.29	1.67	1.9	1.67	1.9
73	US060	304	2.0	-1.19	0.23	0.73	-1.3	0.69	-1.4
78	US064	306	4.2	0.60	0.23	1.11	0.5	1.12	0.5
85	US070	306	2.4	-0.42	0.22	1.25	1.1	1.27	1.2
228	US212	304	3.0	0	0.23	1.31	1.2	1.29	1.1
232	US215	304	3.5	0.84	0.24	0.78	-0.8	0.78	-0.8
*236	US218	304	4.2	1.62	0.23	1.41	1.6	1.42	1.7
246	US227	304	2.6	0.06	0.27	0.44	-2.6	0.42	-2.6
255	US235	306	2.6	0.79	0.22	1.02	0.1	1.00	0
257	US236	306	2.6	-0.77	0.23	1.19	0.8	1.17	0.7
*259	US237	306	3.4	0.33	0.24	1.46	1.7	1.45	1.6
264	US241	306	4.3	0.70	0.23	0.44	-3.1	0.45	-3.0
267	US243	306	4.4	1.18	0.23	0.86	-0.6	0.85	-0.6
271	US246	306	3.6	0.02	0.24	0.93	-0.2	0.93	-0.2
274	US248	306	4.9	2.35	0.23	1.21	1.0	1.20	0.9
283	US256	304	1.8	-0.73	0.29	1.29	1.1	1.22	0.8
294	US266	304	2.1	-1.01	0.23	1.06	0.3	1.05	0.3
302	US273	304	4.1	0.76	0.23	0.60	-1.9	0.60	-1.9
315	US285	304	1.1	-3.00	0.72	1.03	0.2	1.65	0.9
435	US404	306	2.7	0.50	0.23	0.49	-2.6	0.50	-2.5
451	US419	304	2.8	0.88	0.23	0.38	-3.4	0.39	-3.3
453	US420	304	2.2	-0.27	0.22	0.39	-3.9	0.40	-3.8
461	US427	304	3.3	0.74	0.24	1.33	1.2	1.28	1.0
464	US429	304	2.6	0.32	0.23	0.81	-0.7	0.82	-0.7
472	US436	304	4.4	0.89	0.23	1.37	1.5	1.38	1.5
476	US439	306	3.4	0.57	0.23	0.78	-0.8	0.81	-0.7
478	US440	306	5.4	2.29	0.26	0.68	-1.5	0.72	-1.2
482	US443	306	4.8	2.36	0.24	1.00	0	0.96	-0.1
488	US448	306	2.6	-0.40	0.22	0.85	-0.6	0.89	-0.4
491	US450	306	3.7	0.81	0.24	0.85	-0.5	0.87	-0.4
*496	US454	306	4.3	-0.23	0.23	1.43	1.7	1.42	1.7
*502	US459	306	3.1	0.94	0.24	1.69	2.3	1.74	2.4
*504	US460	306	3.4	1.20	0.24	1.72	2.4	1.70	2.4
507	US462	306	4.2	0.73	0.23	0.82	-0.7	0.83	-0.6

509	US463	306	4.3	1.69	0.23	0.42	-3.2	0.42	-3.2
513	US466	306	4.6	1.55	0.23	1.34	1.4	1.35	1.4
518	US470	306	4.4	1.76	0.23	1.06	0.3	1.08	0.4
626	US577	304	3.3	0.33	0.25	1.00	0	1.01	0.1
636	US586	304	2.7	0.61	0.23	0.77	-0.9	0.80	-0.7

Separation: 3.86 Reliability: 0.94

Fixed (all same) chi-square: 725.0 significance *p*: .00

Note: 304 = Grade 4 US sub-sample student, 306 = Grade 6 US sub-sample student *misfitting students

Table 4.10 gives the measurement report for the 50 sub-sample US students' abilities. The standard errors of the 50 US students in the data set of sub-sample are between .22 and .72. The range between the lowest writing ability (-3 logits) and the highest writing ability (2.36 logits) is 5.36 logits. The fixed chi-square of 725.0 (*p* = .00) shows that students have different writing abilities, and a high reliability of separation index at .94 indicates that some students score statistically significantly higher and some score statistically significantly lower.

Ten misfitting students (4 Grade 4; 6 Grade 6) out of the total 50 students are shown in Table 4.10 (infit mean squares range between 1.41 and 1.72) in the analysis.

Table 4.11 Calibration of the prompts for a sub-sample of US student essays using US and HK raters, Lexile Analyzer and IEA

Prompt	Observed Score	Measure (Logit)	S.E.	Infit		Outfit	
				MnSq	ZStd	MnSq	ZStd
4BI	3.5	-0.14	0.06	1.02	0.4	1.02	0.4
4BN	3.5	0.22	0.06	0.83	-3.0	0.87	-2.2
4BP	3.3	-0.07	0.06	1.14	2.3	1.14	2.2

Separation: 2.48 Reliability: 0.86

Fixed (all same) chi-square: 21.6 significance *p*: .00

The calibration for the 3 prompts is presented in Table 4.11. The difficulty of the most difficult prompt, 4BN, is .22 logits, and the prompt difficulty of the easiest prompt, 4BI, is -.14 logits. The fixed chi-square is 21.6 (p = .00) and a reliability of .86 to show the difficulties of 2 prompts are very close, and the 4BN prompt is slightly harder. The standard error of prompts is .06, displaying a high precision of measurement. The infit and outfit statistics of 3 prompts are between .83 and 1.14, which fit the Rasch model well. They are in the reasonable mean square ranges of .7 to 1.4 for infit and outfit statistics.

Table 4.12 Calibration of the US and HK raters for a sub-sample of US student essays using US and HK raters, Lexile Analyzer and IEA

US rater	Observed score	Measure (Logit)	S.E.	Infit		Outfit	
				MnSq	ZStd	MnSq	ZStd
USR101	3.6	-0.39	0.26	0.59	-1.7	0.79	-0.7
USR102	3.5	-0.24	0.23	0.47	-2.8	0.51	-2.5
USR103	3.4	0.36	0.26	0.34	-3.3	0.33	-3.0
USR104	3.1	0.56	0.25	0.38	-3.2	0.42	-2.8
USR105	3.4	0.02	0.22	0.38	-3.6	0.37	-3.4
USR106	3.7	-0.25	0.26	0.57	-1.9	0.58	-1.8
USR200	3.1	1.09	0.24	0.92	-0.2	0.90	-0.3
USR201	3.5	-0.41	0.24	0.74	-1.1	0.72	-1.2
USR202	3.2	-0.62	0.34	0.64	-1.0	0.64	-1.0
USR203	3.3	0.21	0.17	0.76	-1.4	0.77	-1.3
USR204	3.5	0.11	0.21	0.84	-0.7	0.85	-0.6
USR205	3.8	-0.85	0.28	0.57	-1.7	0.56	-1.8
USR206	3.3	-0.07	0.20	0.77	-1.1	0.77	-1.1
USR207	3.1	-0.06	0.26	0.51	-2.2	0.51	-2.2
USR208	3.1	0.18	0.34	0.79	-0.5	0.76	-0.5
USR209	3.0	0.62	0.28	1.01	0.1	1.00	0.1
USR210	3.7	-0.39	0.20	0.64	-2.0	0.65	-1.9
USR211	2.5	0.92	0.76	0.38	-0.8	0.38	-0.7
HKR001	3.1	0.44	0.16	1.72	3.7	1.87	4.0
HKR002	3.7	-0.62	0.17	1.97	4.6	1.94	4.4
HKR003	3.8	-0.86	0.17	1.49	2.6	1.45	2.3
HKR004	3.4	0.24	0.17	1.69	3.5	1.70	3.5

Separation: 1.55	Reliability: 0.71
Fixed (all same) chi-square: 110.4	significance *p*: .00

Table 4.12 shows that the severity span between the most severe rater, USR200 (1.09 logits) and the most lenient rater, HKR003 (-.86 logits) is 1.95 logits; but the most severe raters, USR200, USR211 (.92 logits), USR209 (.62 logits) and USR104 (.56 logits), are all from the US.

The fixed chi-square of 110.4 (p = .00) and a reliability index of .71 showing the raters do not have the same leniency. The standard errors of the raters are between .17 and .76, which are relatively high, exhibiting a low precision of measurement.

In terms of the infit and outfit statistics, the 18 US raters generally interpreted the scale consistently in the reasonable mean square ranges, but 9 out of 18 US raters were overconsistent who had muted rating patterns with infit and outfit statistics less than .6. On the other hand, all the HK raters have erratic rating patterns, with all infit and outfit statistics (i.e., mn sqs and Zstds) greater than the usually accepted values.

Table 4.13 Calibration of the Lexile Analyzer for a sub-sample of US student essays using US and HK raters, Lexile Analyzer and IEA

Lexile Analyzer	Observed score	Measure (Logit)	S.E.	Infit MnSq	Infit ZStd	Outfit MnSq	Outfit ZStd
Category 1	2.1	1.61	0.08	0.88	-1.5	0.94	-0.7
Category 2	2.9	1.15	0.07	1.20	2.5	1.20	2.5
Category 3	3.6	0.15	0.07	1.16	2.0	1.17	2.2
Category 4	3.9	-0.56	0.07	0.74	-3.8	0.73	-3.9
Category 5	4.3	-0.55	0.10	0.98	-0.1	0.99	0
Category 6	4.7	-1.16	0.15	0.94	-0.3	0.92	-0.5
Category 8	4.9	0.13	0.24	1.21	0.9	1.13	0.6
Category 9	5.5	-0.77	0.49	0.76	-0.4	0.88	-0.1
Separation: 4.18				Reliability: 0.95			
Fixed (all same) chi-square: 727.0				significance *p*: .00			

The calibration of the categories from the Lexile Analyzer scoring is presented in Table 4.13. The span of categories is 2.77 logits, from Category 1 (1.61 logits) to Category 6 (-1.16 logits).

In this analysis, Category 8 is located very close to Category 3, and Category 9 is between Categories 4 and 6 because the numbers of observations in Categories 8 and 9 are three and one respectively. The standard errors of Categories 1 to 6 range between .07 and .15, showing a high precision of measurement. However, the standard errors of Categories 8 and 9 are .24 and .49, indicating inadequate observations in the measurement.

The fixed (all same) chi-square is 727.0 (p = .00) to show categories do not have the same measures. The separation index for the Lexile Analyzer is 4.18, which corresponds to a high reliability of .95 to indicate some categories are statistically significantly higher and some are statistically significantly lower.

The infit and outfit statistics of Categories 1 to 9 range between .74 and 1.21, revealing that all categories fit the Rasch model well.

Table 4.14 Calibration of IEA for a sub-sample of US student essays using US and HK raters, Lexile Analyzer and IEA

IEA	Observed score	Measure (Logit)	S.E.	Infit		Outfit	
				MnSq	ZStd	MnSq	ZStd
Category 1	2.2	0.96	0.08	0.96	-0.4	1.00	0
Category 2	2.8	0.90	0.08	0.78	-2.8	0.79	-2.6
Category 3	3.3	0.32	0.08	1.17	2.1	1.16	1.9
Category 4	3.6	0.68	0.08	1.19	2.2	1.19	2.2
Category 5	4.4	-0.07	0.08	0.88	-1.5	0.88	-1.5
Category 6	4.8	-0.90	0.12	1.13	1.1	1.12	1.0

Category 7	4.8	-0.24	0.30	0.48	-2.2	0.49	-2.1
Category 8	4.9	-1.66	0.30	0.41	-2.6	0.43	-2.5
Separation: 5.06			Reliability: 0.96				
Fixed (all same) chi-square: 301.6			significance *p*: .00				

Table 4.14 displays the calibrations for IEA scoring. Category 1 is at .96 logits and Category 8 is at -1.66 logits, showing a range of 2.62 logits between the lowest and highest categories. The fixed chi-square is 301.6 (p = .00) to show categories have different measures. A high reliability of separation index of .96 shows that some categories are higher, and some are lower, and that the consistency of these inferences could be expected in the analysis.

The standard errors of Categories 1 to 6 are between .08 and .12, signifying a high precision of measurement. The observations in Categories 7 and 8 are only 2, and no observation occurs in Category 9; the standard errors of Categories 7 and 8 are .30, indicating the inadequate number of observations in those categories.

The infit and outfit mean squares of Categories 1 to 6 are between .78 and 1.19, which fit well to the Rasch model. The infit and outfit mean squares of Categories 7 and 8 are between .41 and .49, which are considered overfitting.

Consequential Research Question:

The MFRM Analyses Results of US and HK Student Essays (all student essays)

4. Can these scales be used to compare:
   a. human scoring vs. scoring of the Lexile Analyzer of the essays of US and HK students?
   b. human scoring of the essays of US and HK students vs. the IEA scoring of the essays of US students?
   c. human scoring of the essays of US and HK students vs. the Lexile Analyzer scoring of the essays of US and HK students vs. the IEA scoring of the essays of US students?

In the MFRM analyses of all student essays with scoring of US and HK human raters and the Lexile Analyzer, and all student essays with scoring of the US and HK human raters and IEA: the results are satisfactory and similar.

The Lexile Analyzer, IEA and 19 US human raters fit the Rasch model very well, although approximately 130 students show misfit in the analyses, with the range of infit mean squares between 1.41 and 8.79; two prompts, 4AP and 4BP, and 3 to 4 HK human raters show misfit in the analyses (Appendices 19 & 20).

```
+-------------------------------------------------------------------------------------------------------------+
|Measr|-Prompt           |-US Rater                       |-HK Rater |-Lexile Analyzer|-IEA|+Student    |R6SCA|
|-----+------------------+--------------------------------+----------+----------------+----+------------+-----|
|   4 +                  +                                +          +                +    + .          + (6) |
|     |                  |                                |          |                |    |            |  5  |
|     |                  |                                |          |                |    | .          |     |
|     |                  |                                |          |                |    |            |     |
|   3 +                  +                                +          +                +    +            +     |
|     |                  |                                |          |                |    | .          | --- |
|     |                  |                                |          |                | 1  | .          |     |
|     | 13-12AI   15-12AP|                                |          |                |    | .          |     |
|   2 + 16-12BI          +                                +          +                + 2  + **.        +     |
|     | 9-8AP            |                                |          |                |    | **.        |     |
|     |                  |                                |          | 1              |    | ***.       |     |
|     | 17-12BN   18-12BP|                                |          |                | 3  | *****.     |  4  |
|   1 +                  + US200                          +          + 2              + 4  + ******.    +     |
|     | 11-8BN    12-8BP | US103   US104   US211          | HK004    |                |    | ********.  |     |
|     | 14-12AN          | US209                          |          | 3              |    | *******.   |     |
|     | 7-8AI     10-8BI | US102   US105   US206          | HK001    | 4              |    | *******.   |     |
*   0 * 8-8AN            * US101   US203   US204   US205  *          *                * 5  * *********. * --- *
|     |                  | US201   US210                  |          | 5 6            |    | ********   |     |
|     |                  | US106   US202   US207   US208  | HK002    |                |    | *****.     |     |
|     |                  |                                |          | 7 8            | 6  | ****.      |     |
|  -1 +                  +                                +          +                +    + ***.       +     |
|     |                  |                                |          | 9              | 7  | ***.       |  3  |
|     |                  |                                |          |                |    | *.         |     |
|     | 5-4BN            | US108                          | HK003    |                |    | **         |     |
|  -2 + 4-4BI     6-4BP  +                                +          +                + 8  + *.         +     |
|     | 2-4AN            |                                |          |                |    | *.         |     |
|     | 1-4AI     3-4AP  |                                |          |                |    | .          | --- |
|     |                  |                                |          |                |    | *.         |     |
|  -3 +                  +                                +          +                + 9  + *          +     |
|     |                  |                                |          |                |    | .          |     |
|     |                  |                                |          |                |    | .          |     |
|     |                  |                                |          |                |    | .          |  2  |
|  -4 +                  +                                +          +                +    + .          +     |
|     |                  |                                |          |                |    | .          |     |
|     |                  |                                |          |                |    | .          |     |
|     |                  |                                |          |                |    | .          |     |
|  -5 +                  +                                +          +                +    + .          + --- |
|     |                  |                                |          |                |    | .          |     |
|     |                  |                                |          |                |    | .          |     |
|     |                  |                                |          |                |    | .          |     |
|  -6 +                  +                                +          +                +    +            +     |
|     |                  |                                |          |                |    | .          |     |
|     |                  |                                |          |                |    |            |     |
|     |                  |                                |          |                |    | .          |     |
|  -7 +                  +                                +          +                +    +            +     |
|     |                  |                                |          |                |    |            |     |
|     |                  |                                |          |                |    | .          |     |
|     |                  |                                |          |                |    |            |     |
|  -8 +                  +                                +          +                +    + .          + (1) |
|-----+------------------+--------------------------------+----------+----------------+----+------------+-----|
|Measr|-Prompt           |-US Rater                       |-HK Rater |-Lexile Analyzer|-IEA| * = 8      |R6SCA|
+-------------------------------------------------------------------------------------------------------------+
```

Figure 4.8 Vertical map for all student essays with scoring of US and HK raters, Lexile Analyzer and IEA

Considering the similarities of the MFRM analyses results of the student essays in this research, a detailed description of the overall picture only is presented below:

Figure 4.8 shows a set of scales of prompt difficulty, rater severity and student ability that was constructed by analysing every student essay in this research. In this analysis, column 1 shows the logit scale is from -8 to 4 logits. Column 8 presents the human ratings, which range from categories 1 to 6. All the means for prompt difficulty, rater severity and student ability (columns 2 to 7) have been set by default at 0.00 logits.

In column 2, the top of the scale is the most difficult prompt, 12AP (2.15 logits). Grade 10 and Grade 12 students responded to this persuasive prompt, which was for Grade 12 students from the NAEP. The easiest prompt is 4AP (-2.59 logits), and it was administered to Grade 4 students only.

In column 3, USR200 is the most severe US rater (1.05 logits), and USR108 is the most lenient US rater (-1.68 logits). For the severity of HK raters, column 4 displays that the most severe rater is HKR004 (0.64 logits), and the most lenient one is HKR003 (-1.71 logits).

In columns 5 and 6, the Lexile Analyzer and IEA categories are ordered appropriately. The range of logit measures of the Lexile Analyzer is between -2 and 2, which is close to that of the human raters. On the other hand, the IEA category measures range from -3 to 3 logits, which is close to the measurement of the majority of the students' writing ability.

Students are more spread out along the essay writing ability measure in column 7, with the range from -8 to 3 logits. The majority of students' writing abilities fall between -3 and 3 logits,

Table 4.15 Calibration of the students for all student essays using US and HK raters, Lexile Analyzer and IEA (20 of the 127 misfitting students)

Student			Observed	Measure	S.E.	Infit		Outfit	
Number		Group	Score	(Logit)		MnSq	ZStd	MnSq	ZStd
730	HK091	206	2.2	-4.97	0.65	1.42	0.8	1.42	0.8
487	US448	106	2.5	-1.30	0.29	1.43	1.5	1.42	1.5
227	US212	104	3.3	-0.88	0.30	1.43	1.5	1.43	1.5
769	HK130	206	4.2	-0.58	0.65	1.43	0.8	1.51	0.9

417	US387	110	4.6	1.57	0.32	1.44	1.5	1.44	1.5
152	US137	108	4.9	1.78	0.32	1.44	1.5	1.47	1.6
498	US456	106	3.4	-0.30	0.33	1.45	1.4	1.48	1.5
309	US280	104	4.4	0.02	0.39	1.45	1.2	1.43	1.1
307	US278	104	4.8	1.19	0.31	1.45	1.5	1.42	1.4
201	US186	108	4.2	1.35	0.33	1.45	1.4	1.46	1.4
...	...	...	...	...	...	...	...	...	...
729	HK090	206	3.3	-2.07	0.66	3.34	2.6	3.34	2.6
751	HK112	206	3.3	-1.92	0.66	3.78	2.9	3.75	2.9
723	HK084	206	4.3	-0.40	0.79	4.05	2.7	4.33	2.9
625	US577	104	3.1	-1.24	0.31	4.11	6.6	4.19	6.7
299	US271	104	3.2	-1.43	0.36	4.30	5.9	4.34	6.0
753	HK114	206	3.8	-1.03	0.66	5.10	3.9	5.11	3.9
725	HK086	206	2.2	-4.64	0.71	7.48	5.0	7.45	5.0
628	US579	104	3.8	-1.12	0.33	7.70	9.0	7.08	9.0
675	HK036	206	3.8	-1.72	0.72	8.05	4.9	8.10	5.0
660	HK021	206	3.8	-1.43	0.66	9.00	5.8	8.79	5.7

Separation: 3.68 Reliability: 0.93

Fixed (all same) chi-square: 8841.3 significance *p*: .00

Note: 104 = Grade 4 US student, 106 = Grade 6 US student, 108 = Grade 8 US student, 110 = Grade 10 US student, 112 = Grade 12 US student, 206 = Grade 6 HK student, 304 = Grade 4 US sub-sample student, 306 = Grade 6 US sub-sample student

The measurement report describing the ability of 776 Grade 4 to Grade 12 students indicates that the standard errors of 776 students range between .26 and 1.17. The wide range between the lowest writing ability (-8.97 logits) and the highest writing ability (3.92 logits) is 12.89 logits.The fixed chi-square value of 8841.3 (p = .00) shows that students have different writing abilities and a high reliability of separation index of .93 signify that some students' writing abilities are statistically significantly higher and some are statistically significantly lower (Appendix 21).

In Table 4.15, 127 misfitting students out of the 776 students are found with infit mean squares ranging from 1.42 to 9 (Appendix 21). These 127 students include 58

US students (45.7%), most of whom are Grade 4 and Grade 6 students; 45 HK students (35.4%); and 24 US sub-sample students (18.9%). The analysis results for 20 of the 127 misfitting students are shown in Table 4.15.

Table 4.16 Calibration of the prompts for all student essays using US and HK raters, Lexile Analyzer and IEA

Prompt	Observed score	Measure (Logit)	S.E.	Infit MnSq	Infit ZStd	Outfit MnSq	Outfit ZStd
4AI	3.6	-2.46	0.07	0.73	-4.9	0.72	-5.1
4AN	3.7	-2.32	0.07	0.76	-4.2	0.81	-3.1
4AP	3.6	-2.59	0.07	1.66	9.0	1.64	8.7
4BI	3.6	-2.00	0.04	1.24	7.0	1.24	6.8
4BN	3.6	-1.76	0.04	0.96	-1.3	0.98	-0.6
4BP	3.5	-1.98	0.04	1.55	9.0	1.55	9.0
8AI	3.5	0.16	0.05	1.01	0.1	1.01	0.3
8AN	3.7	0.04	0.05	0.65	-8.9	0.65	-9.0
8AP	2.9	1.70	0.05	0.85	-3.6	0.85	-3.4
8BI	3.8	0.30	0.05	0.87	-2.7	0.87	-2.8
8BN	3.7	0.81	0.05	0.63	-8.9	0.63	-8.7
8BP	3.6	0.87	0.05	1.12	2.4	1.12	2.4
12AI	3.8	2.13	0.06	0.77	-4.8	0.77	-4.9
12AN	3.9	0.60	0.06	0.70	-6.4	0.71	-6.3
12AP	3.4	2.15	0.06	0.81	-3.9	0.81	-3.8
12BI	3.5	2.07	0.08	0.82	-2.6	0.81	-2.7
12BN	4.0	1.16	0.08	0.74	-4.0	0.73	-4.1
12BP	4.0	1.13	0.08	1.00	0	1.00	0
Separation: 27.66				Reliability: 1.00			
Fixed (all same) chi-square: 16753.8				significance *p*: .00			

Table 4.16 presents the calibration for the 18 prompts. The span of prompt difficulties ranges from 4AP (-2.59 logits), the easiest prompt to 12AP (2.15 logits), the most difficult prompt is 4.74 logits. The fixed chi-square is 16753.8 (p = .00) to show that the prompts are in various levels of difficulty. A high reliability of 1.00 indicates that some prompts are statistically significantly more difficult and some are statistically

significantly easier. The standard errors range from .04 to .08 showing high precision of measurement. Two misfitting prompts are found in this analysis: 4AP (infit mn sq 1.66; outfit mn sq 1.64), and 4BP (infit and outfit mn sqs 1.55) (Appendix 3).

Table 4.17 Calibration of the US raters for all student essays using US and HK raters, Lexile Analyzer and IEA

US rater	Observed score	Measure (Logit)	S.E.	Infit		Outfit	
				MnSq	ZStd	MnSq	ZStd
USR101	3.8	0.02	0.05	0.71	-6.9	0.75	-5.6
USR102	3.8	0.24	0.05	0.66	-8.6	0.67	-8.3
USR103	3.6	0.76	0.09	0.61	-5.5	0.61	-5.3
USR104	3.6	0.80	0.05	0.63	-9.0	0.64	-9.0
USR105	3.8	0.25	0.05	0.79	-4.8	0.79	-4.9
USR106	4.0	-0.39	0.06	0.59	-8.8	0.59	-8.6
USR108	3.3	-1.68	0.37	0.54	-1.5	0.53	-1.6
USR200	3.2	1.05	0.06	1.14	2.4	1.14	2.4
USR201	3.6	-0.20	0.05	1.09	2.0	1.09	1.9
USR202	3.6	-0.48	0.06	1.09	1.5	1.08	1.3
USR203	3.5	-0.04	0.05	1.06	1.3	1.06	1.3
USR204	3.6	-0.04	0.05	0.85	-3.4	0.86	-3.3
USR205	3.5	-0.10	0.06	0.95	-0.8	0.96	-0.7
USR206	3.4	0.26	0.05	0.82	-4.3	0.82	-4.4
USR207	3.6	-0.47	0.06	0.95	-0.9	0.95	-0.9
USR208	3.7	-0.40	0.06	0.95	-1.0	0.95	-1.0
USR209	3.4	0.58	0.06	1.00	0	1.00	0
USR210	3.7	-0.18	0.05	1.22	4.5	1.22	4.4
USR211	3.2	0.78	0.35	0.64	-1.2	0.61	-1.4
Separation: 4.85				Reliability: 0.96			
Fixed (all same) chi-square: 1130.5				significance *p*: .00			

Table 4.17 shows that the severity span across US raters (USR200 to USR108) is 2.73 logits. The fixed chi-square of 1130.5 (p = .00) and a high reliability index of .96 show that the US raters have different severity on scoring essays.

The standard errors of USR108 (.37) and USR211 (.35) are relatively high, signifying a low precision of measurement because these raters scored only 19 and 18 respectively of 3453 essays in the data set. The US raters generally interpreted the scale consistently, with infit and outfit mean squares between .61 and 1.22. The exception was two overfitting raters, USR106 and USR108, who had overconsistent rating patterns with infit and outfit statistics between .53 and .59.

Table 4.18 Calibration of the HK raters for all student essays using US and HK raters, Lexile Analyzer and IEA

HK rater	Observed score	Measure (Logit)	S.E.	Infit		Outfit	
				MnSq	ZStd	MnSq	ZStd
HKR001	3.1	0.23	0.10	1.62	6.0	1.65	6.1
HKR002	3.5	-0.57	0.10	1.46	4.7	1.45	4.5
HKR003	3.8	-1.71	0.10	1.99	9.0	1.98	8.5
HKR004	3.0	0.64	0.10	1.70	6.8	1.68	6.5
Separation: 2.09				Reliability: 0.81			
Fixed (all same) chi-square: 421.7				significance *p*: .00			

In Table 4.18, the range of severity span is from the harshest rater, HKR004 (.64 logits) to the most lenient rater, HKR003 (-1.71 logits) is 2.35 logits.The standard errors are .10. The separation index for HK raters is 2.09, which corresponds to an acceptable reliability of separation index (R=.81) to show that some raters scored (statistically significantly) more severely and some scored more leniently. The fixed (all same) chi-square was 421.7 (p = .00) to reveal that raters have different severity. All four HK raters have erratic rating patterns, with all infit and outfit statistics greater than the usually accepted values.

Table 4.19 Calibration of the Lexile Analyzer for all student essays using US and HK raters, Lexile Analyzer and IEA

Lexile Analyzer	Observed score	Measure (Logit)	S.E.	Infit		Outfit	
				MnSq	ZStd	MnSq	ZStd
Category 1	2.3	1.51	0.05	1.16	3.6	1.18	3.8

Category 2	2.9	1.05	0.04	1.28	7.3	1.29	7.6
Category 3	3.2	0.46	0.03	1.06	1.9	1.07	2.2
Category 4	3.5	0.15	0.03	0.94	-2.1	0.95	-2.0
Category 5	3.6	-0.14	0.03	1.01	0.2	1.00	0
Category 6	3.9	-0.36	0.03	0.84	-5.7	0.84	-5.6
Category 7	4.1	-0.64	0.04	0.90	-3.1	0.89	-3.2
Category 8	4.5	-0.77	0.05	0.92	-1.9	0.92	-2.0
Category 9	4.9	-1.26	0.06	1.05	0.8	1.06	1.0
Separation: 20.20			Reliability: 1.00				
Fixed (all same) chi-square: 2936.7			significance *p*: .00				

The calibration of the Lexile Analyzer categories is presented in Table 4.19. The span of categories is 2.77 logits, from 1.51 logits (SE .05) for Category 1, the lowest category, to -1.26 logits (SE .06) for Category 9, the highest category. The separation index for the Lexile Analyzer is 20.20, which corresponds to a high reliability of separation index (R=1.00) to show some categories are higher and some are lower. The fixed (all same) chi-square was 2936.7 (p = .00), indicating that categories have different measures. The standard errors range between .03 and .06, displaying a high precision of measurement. The infit and outfit statistics of Categories 1 to 9 are between .84 and 1.28, well within the reasonable mean square ranges of .7 to 1.4 for the infit and outfit statistics.

Table 4.20 Calibration of IEA for all student essays using US and HK raters, Lexile Analyzer and IEA

IEA	Observed score	Measure (Logit)	S.E.	Infit MnSq	Infit ZStd	Outfit MnSq	Outfit ZStd
Category 1	2.2	2.57	0.05	1.11	2.5	1.16	3.2
Category 2	2.8	1.94	0.04	0.91	-2.4	0.91	-2.4
Category 3	3.1	1.33	0.03	1.07	2.1	1.07	2.2
Category 4	3.3	0.92	0.03	1.10	3.1	1.10	3.4
Category 5	3.7	0.05	0.03	0.88	-4.7	0.88	-4.9
Category 6	4.0	-0.70	0.03	0.93	-2.2	0.93	-2.3
Category 7	4.2	-1.21	0.04	0.97	-0.8	0.96	-1.0

Category 8	4.6	-2.00	0.04	0.96	-1.1	0.95	-1.2
Category 9	5.0	-2.89	0.06	0.93	-1.4	0.93	-1.2
Separation: 40.80			Reliability: 1.00				
Fixed (all same) chi-square: 12062.6			significance *p*: .00				

Table 4.20 presents the calibrations for IEA scoring. Category 1 is at 2.57 logits, and Category 9 is at -2.89 logits, indicating a wide range of 5.46 logits between the lowest and highest categories. The fixed chi-square of 12062.6 (p = .00) shows that categories do not have the same measures. A high reliability of separation index at 1.00 signifies that some categories are higher and some are lower in the analysis, and that the consistency of these inferences could be expected. The standard errors are between .03 and .06, showing a high precision of measurement. There is no evidence of misfit with all of the infit and outfit mean squares between .88 and 1.11.

Bias Analyses

Bias analyses can show interactions between facets in the model that have a significant influence affecting the research on essay ratings. The analyses help improve the validity of student scores as the biases that are due to the differences in severity of particular raters are removed in the Many-facets Rasch analysis (Lunz, Wright, & Linacre, 1990).

The interaction involved a rater and other aspects of the rating situation. Bias analyses, the identification of these systematic sub-patterns of behavior, are achieved in the MFRM (McNamara, 1996). Differential facet functioning in the MFRM model can be examined between various groups to examine bias issues. In other words, the facets are calibrated separately within relevant groups, and the relative difficulty of the elements is examined. Interactions between facets can be examined as a potential source of bias in a writing assessment (Engelhard, 1992; Linacre, Engelhard, Tatum, & Myford, 1994).

For the bias/interactions, the t-statistic is obtained by dividing the bias measure by its standard error. With more than 30 observations, a t-statistic is approximately normally distributed, i.e., a z-statistic. The number of z-score values outside the range of approximately +2 to -2 (i.e., the range +1.96 to -1.96) suggests significant bias (Linacre, 1989; McNamara, 1996).

In the bias analysis for the US raters and students, 244 of the 8670 bias terms with z-scores outside of or equal to the particular range are found. However, while 723 bias terms are included in the bias analysis for the HK raters and students, only 65 bias terms with z-scores are outside of or equal to +/-1.96 (Appendix 22).

The results of a bias analysis on the interaction between the raters and prompts provide information on each rater's unusual patterns of leniency or severity for each prompt (McNamara, 1996). In the bias analysis for the US raters and prompts, 69 significant US rater-by-prompt interactions of 176 bias terms are indicated. The results of a bias analysis on the interaction with the HK raters and prompts demonstrate that only 8 significant HK rater-by-prompt interactions out of 12 bias terms are indicated by z-scores outside the range. Half of the significant bias interactions are negative, showing lenient ratings, and the other half are positive, showing severe ratings by the HK raters in the analysis (Appendix 22). All the HK raters show equally severe or lenient bias towards the three prompts. Obviously, for a particular prompt, the rater is sometimes more lenient or severe than expected. There is no stable rating pattern of bias in one direction or another, although the rater may be consistently more lenient or severe for that prompt (McNamara, 1996).

Consequential Research Question:

5. What is the extent to which the MFRM can be used to detect and adjust for rater effects in human essay scoring?

Wright (1996) stated that separation in the analysis is the number of significantly different performance strata that can be identified by the test. Separation can be expressed with the range 0 to infinity. (i.e., Separation = True Standard Deviation / Error Standard Deviation)

In the MFRM analyses, the results of separation and Standard Deviation (S.D.) indicate a model to detect and adjust for rater effects. The 'True' S.D. is often called the 'Adjusted' (for measurement) S.D. In fact, the MFRM analyses of all facets of the essay data set can also be examined individually. Linacre (1989) stated that the Many-facets Rasch Measurement (MFRM) provides for multiple facets that can be calibrated simultaneously, so they can be examined separately. A sound theoretical framework of the MFRM is provided by adjusting for the differences in raters and writing tasks in assessment. The objectivity and fairness of the measurement of writing competence is improved by adjustments for rater severity and prompt difficulty. Raw scores may lead to under- or over-estimates of writing competence when students are rated by different raters on different writing tasks (Engelhard, 1992; Linacre, Engelhard, Tatum, & Myford, 1994).

Summary

The best categorisation for the measures of two AES systems, the Lexile Analyzer and IEA, was obtained from the collapsing of categories, following the guidelines of Linacre (1999, 2002). Thus, optimal categorisation was used in the MFRM analyses in this research.

The purpose of analysing different sets of data, including the US student essays, the HK student essays and the sub-sample of the US student essays with scoring of US raters and/or HK raters and/or the Lexile Analyzer and/or IEA is to generate the MFRM analysis results of all student essays with scoring of the US and HK human raters, the Lexile Analyzer and IEA.

In the analysis of the student essays in this research, 127 students, 2 prompts (4AP and 4BP) and 4 HK human raters show misfit. The analyses results show that the Lexile Analyzer and IEA fit very well with the Rasch model, and no misfitting US human raters are found.

Chapter Five

Discussion

Overview

This research investigated the effectiveness of a new AES system, the Lexile Analyzer, through the Many-facets Rasch Measurement (MFRM) analyses using FACETS software. The analyses focused on essays from US students, HK students and a sub-sample of the US students. The students were administered writing prompts from the US National Assessment of Educational Progress (NAEP) and essays were scored by human raters using the NAEP holistic essay marking rubrics covering the narrative, informative and persuasive genres. 3453 essays were collected from 589 Grades 4 to 12 US students and each essay was rated by four of the nineteen paid trained raters from an established US testing company. 408 essays were written by 137 Grade 6 HK students from a single local primary school as the HK raters and each essay was scored against the US holistic rubrics by four of the HK raters. Moreover, a sub-sample of the US student essays was scored by two of the four HK raters. The US essays were scored using the Lexile Analyzer and IEA scoring engines; the HK essays by the Lexile Analyzer only.

Main Findings

The results of a series of analyses adopting the Many-facets Rasch model, showing the effect of various facets including student, prompt, rater and two AES systems, the Lexile Analyzer and IEA, could be measured with the Rasch model which were presented in Chapter Four.

The main findings of the MFRM analysis of all US and HK student essays with scoring of the US and HK human raters, the Lexile Analyzer and IEA in this research can be summarised as follows:

1. Students

127 misfitting students of total 776 students were included in the analysis. Most of the misfitting cases were students who performed erratically, produced off-topic essays or wrote essays that were scored by the most severe or lenient rater(s) in this research.

Three misfitting examples, Student HK021, Student US579 and Student US057, are discussed - below to show how they performed haphazardly when responding to prompts. The selection criteria are quantitative (Rasch fit statistics), but the evidence below is qualitative.

The infit and outfit statistics of Student HK021 are infit mn sq 9.0 and Zstd 5.8; outfit mn sq 8.79 and Zstd 5.7. This student responded to three prompts, and the rating pattern of three different combinations of 2 HK raters for each prompt is 1,2 (very low) for the 4BI prompt, 5,5 (high) for the 4BN prompt and 4,6 (mixed) for the 4BP prompt. The main reason for this discrepancy is that Student HK021 seems to have misinterpretation of the instructions in the 4BI prompt, which asked students to describe ‘what lunch time is like on a school day’. Student HK021 described the day he could not find his lunch box at school (Appendix 23).

Student US579 exhibits a noisy rating pattern, with an infit mean square of 7.7 and Zstd of 9.0; an outfit mean square of 7.08 and Zstd of 9.0. This student wrote 6 essays, and the rating pattern of six different combinations of four US raters for these essays is 4,4,3,4 (moderate) for the 4AI prompt, 5,4,4,4 (high) for the 4AN prompt, 1,1,1,1 (consistently low) for the 4AP prompt, 4,4,4,4 (moderate) for the 4BI prompt, 5,5,4,5 (high) for the 4BN prompt and 6,5,5,5 (very high) for the 4BP prompt. The instructions in the 4AP prompt seem similarly misinterpreted by Student US579; this student described a happy family life rather than responding to the essay prompt which required the construction of a letter to convince a friend to be visible (Appendix 23).

The infit and outfit mean squares of Student US057 are both 2.27. This student wrote 3 essays. For Student US057, the rating pattern of three different combinations of 4 US and 2 HK human raters on the 3 prompts is 5,3,4,4,3,6 (mixed) for the 4BI prompt, 3,3,3,3,3,2 (moderate) for the 4BN prompt and 1,1,1,1,1,1 (consistently low) for the 4BP prompt. Coincidentally, the scores of the 4BP prompt from both the US and HK human raters for Student US057 are the lowest score on the rating scale. For the 4BP prompt, students were asked to write a letter to convince the school librarian to buy their missing favourite book, but Student US057 wrote a letter to his parents to describe how much he loves them and how he tries to be a good boy in school. This student misinterpreted the instructions of the 4BP prompt (Appendix 23).

Students HK021, US579 and US057 exhibit unexpectedly erratic performances on one prompt to which they responded. The ratings for that prompt are far worse than their average performance on the other prompts.

Another misfitting student, US366, (infit mean square 2.53; Zstd 3.6; outfit mean square 2.49; Zstd 3.5), produced one off-topic essay when he responded to the 12AN prompt. He wrote 6 essays in total. For Student US366, the rating pattern of six different combinations of 4 US raters on the 6 prompts is 5,6,6,5 (very high) (Lexile

Analyzer measure: 1471.02; IEA measure: 64) for the 8BI prompt; 5,5,4,5 (high) (Lexile: 1365.69; IEA: 66) for the 8BN prompt; 5,5,4,3 (mixed) (Lexile: 1284.27; IEA: 59) for the 8BP prompt; 4,4,4,4 (moderate) (Lexile: 1173.02; IEA: 68) for the 12AI prompt; 2,1,1,1 (very low) (Lexile: 751.92; IEA: 33) for the 12AN prompt; and 4,4,4,2 (mixed) (Lexile: 1427.62; IEA: 67) for the 12AP prompt (Appendix 23). The 12AN prompt requires a story about a special object. Students were expected to describe the main character's first encounter with the object and why the object is so important to the character. The following excerpt from Student US366 on the 12AN prompt illustrates the response:

Once upon a time I felt very sad during the particular day on which we do these Lexile papers. Then I thought to myself, "Wow, I've done pretty well so far, but today's paper is just not gonna happen. I only need enough writing here to fill up this and part of the second to seem like I'm accomplished for the day." So in conclusion, if you've read this far, I'm just gonna do it right now.

Student US366's performance including human ratings and the Lexile Analyzer and IEA measures on the 12AN prompt was much worse than his above-average performances on the other 5 prompts. In responding to this prompt, he did try to show that he produced this off-topic essay on purpose.

On the other hand, the severity of raters impacts directly on the ratings of student essays. Each rater's particular patterns of harshness or leniency impacts the scores for the student essays (McNamara, 1996). For example, the infit and outfit statistics of Student HK086 are infit mn sq 7.48 and Zstd 5.0; outfit mn sq 7.45 and Zstd 5.0. This student has a relatively consistent performance responding to three prompts, and he wrote only a few words on the given paper or responded to the prompt without a complete ending. The rating pattern of three different combinations of 2 HK raters for Student HK086 is 1,1 (very low) for the 4BI prompt, 1,1 (very low) for the 4BN prompt and 2,6 (mixed) for the 4BP prompt (Appendix 23). For the 4BP prompt, the most lenient HK rater, HKR003, gave a rating of 6 to Student HK086's obviously unfinished prompt:

Dear school librarian:

I was missing a favourite book from your school library, this book name is 'FULLMETAL ALCHEMIST' this book have many people to reading, this book my teachers like reading too, so you wants to buy the book again.

The score for Student HK086 from one rater is different by four points from the rating of another rater for the 4BP prompt. In other words, the students' performance would be affected by both writing ability and the severity of human raters.

2. Prompts

The MFRM results show two misfitting prompts, 4AP and 4BP. Both were Grade 4 persuasive prompts, and students were requested to write letters to convince a) a friend to become visible and b) the school librarian to buy their missing favourite book for the school library, with details, examples or reasons.

However, the MFRM analyses of both the US human ratings of the US student essays and the HK human ratings of the HK student essays reveal that all prompts fit very well with the Rasch model. A similar result can be found by analyzing the scoring of the Lexile Analyzer of the HK student essays, which shows that the prompts fit well with the model.

Surprisingly, two same prompts, 4AP and 4BP, show misfit in the MFRM analyses results of scoring of the Lexile Analyzer and IEA on the US student essays. These two misfitting prompts impact only the AES scoring of the US student essays. This is unusual for both the Lexile Analyzer and IEA engines because they could not assess directly the content of an essay. In other words, the Grade 4 persuasive prompts fit the model when only human ratings of the US essays were included in the analysis. Human raters might make professional adjustment before scoring these two persuasive prompts in this research as they understood that persuasive prompts were difficult for Grade 4 students.

3. US Raters

All 19 US raters scored all student essays consistently; two out of 19, USR106 and USR108, are over-consistent, having constrained or muted rating patterns with infit and outfit statistics between .53 and .59.

Other than that, the MFRM analysis result of the US student essays shows that all 19 US human raters are in the acceptable range of infit and outfit mean squares; all of the US raters scored the US student essays consistently.

4. HK Raters

In the MFRM analysis result of all student essays including 3453 US student essays, 408 HK student essays and a sub-sample of 150 US student essays, all 4 HK raters show misfit in the infit and outfit statistics.

On the other hand, the analysis result of the HK human scoring of the 150 sub-sample of the US student essays shows that four HK raters fit the Rasch model very well

(Appendix 17), i.e., they scored the US student essays consistently. A similar result can be found in the analysis of the HK human scoring of the 408 HK essays: four HK raters fit well with the model (Appendix 13), showing that they scored the HK student essays consistently. Moreover, the analysis results of the HK human scoring of the 408 HK essays and the 150 sub-sample of the US student essays show that four HK raters scored the US and HK essays consistently (Appendix 24).

Consequently, these analyses reveal that the 4 HK raters scored the US essays and/or the HK essays consistently, but that they scored essays differently from the 19 US raters in this research.

5. The Lexile Analyzer

All the categories of the Lexile Analyzer fit well with the Rasch model in the analysis of all student essays with scoring of the US and HK human raters, the Lexile Analyzer and IEA. The infit and outfit statistics fall in the range of .84 to 1.28, well within the acceptable range of .6 to 1.4. Thus, the new AES system, the Lexile Analyzer, can score student essays as consistently as the trained professional human raters in this research. Most importantly, the Lexile Analyzer scores consistently with HK human raters.

6. Intelligent Essay Assessor (IEA)

The infit and outfit mean squares of Categories 1 to 9 on the scale of IEA are very close to 1.0: evidence that IEA scored student essays as consistently as did the human raters in this research.

Rescaling the Lexile Analyzer and IEA Measures

Correlations estimate merely the degree of agreement between grades across the rating procedures, whereas calibration establishes agreement between the actual scores assigned with the possibility of converting from one system to another - as in the calibrated Celsius, Kelvin and Fahrenheit systems for measuring temperature (Bond & Fox, 2007). The same logic can be applied to the relationship between the Lexile Analyzer and IEA scales in this research.

Linacre and Wright (1989) claimed that a step is necessary to adjust the relative lengths of the logits constructed by two tests by using the ratio of the observed standard deviations of the measures common to those tests, so that the two tests measure in the same substantive units. This researcher has rescaled the measures of the Lexile Analyzer and IEA to the range of each other using the formulae as follows:

New IEA measure = (IEA measure - IEA mean) * Lexile Analyzer S.D. / IEA S.D. + Lexile Analyzer mean; and

New Lexile Analyzer measure = (Lexile Analyzer measure - Lexile Analyzer mean) * IEA S.D. / Lexile Analyzer S.D. + IEA mean

Table 5.1 Rescaling the IEA and Lexile Analyzer measures

Original IEA measures	New IEA measures	Original Lexile Analyzer measures	New Lexile Analyzer measures
1	-700.90	-196.09	20
100	1894.62	2039.68	106

In Table 5.1, the IEA measures were rescaled from 1 to 100 to be the new IEA measures between -700.90 and 1894.62. The Lexile Analyzer measures were then rescaled the range from -196.09 to 2039.68 to be the new Lexile Analyzer measures between 20 and 106.

In the WINSTEPS analyses, the best 9 categories for the new measures of the Lexile Analyzer and IEA were used and all response categories showed good functioning in the results. However, the MFRM results of the new measures of two AES systems showed only minor differences from the analyses already conducted in Chapter 4 (Appendix 25).

Implications for Practice

In the MFRM analyses, all facets of the rating situation are modeled simultaneously so their effect on scores can be estimated. McNamara (1996) claimed that raw scores can be a misleading guide to student ability. Many-facets measurement has made routine exposure of the extent of (dis)agreement among raters possible, and has the potential to offer a way out where the technology is available. MFRM adjustments for rater severity improve the objectivity and fairness of the measurement of writing ability (Engelhard, 1992). Given that raters are consistently more or less lenient than each other, it would be patently unfair not to adjust students' final grades according to the allocation of rater pairs (Bond & Fox, 2007).

Exemplars from the US Student Essay Sample - Unfair Raw Score Cases

As expected, the writing ability estimates of the 589 US students in the MFRM analysis varied considerably: raw scores from 4 to 24 out of a possible 24 yielded estimates that varied from a low of -6.12 logits to a high of 3.81 logits.

The range of variation in severity across the 19 raters is approximately 3 logits (see Table 4.3), with the most severe rater USR200 estimated at 1.25 logits (SE .07) and the most lenient rater USR108 estimated at -1.71 logits (SE .40). This range of rater severity could have a remarkable effect on raw scores of the US student essay writing, and has the potential to confound severely the essay writing ability estimates of students.

Many cases in the MFRM analysis showed that students received the same raw scores from raters, but they had significant differences in Rasch abilities.

Table 5.2 Observed and Expected Ratings for Grades 4, 8 and 12 Selected US Students

Grade	Student	Prompt	R101	R102	*R103	*R104	R105	R106	R108	*R200	R201	R202	*R203	*R204	*R205	*R206	R207	R208	*R209	R210	*R211	Raw Score	Infit MnSq	Outfit MnSq	Rasch Ability
4	14 (US013)	1-4AI								4	4		5	5								104	0.76	0.78	1.35
		2-4AN		4	3		4	4																	
		3-4AP								4	5		5	4											
		4-4BI								5	5		5			5									
		5-4BN	5	5		4	5																		
		6-4BP										4				3	4				3				
	471 (US436)	1-4AI										4	4		4			4				104	1	0.99	0.08
		2-4AN		5	4	4	5																		
		3-4AP												4			5		4	5					
		4-4BI									5		6			5	5								
		5-4BN	4	5			4	4																	
		6-4BP								2		4	4				4								
8	133 (US118)	7-8AI									3			3	2			2				55	0.72	0.71	-2.17
		8-8AN	3	3		3		3																	
		9-8AP									1		1					2		1					
		10-8BI												2		2			2	2					
		11-8BN	3	3		2	3																		
		12-8BP								2	2			2				3							
	140 (US125)	7-8AI												2	2	2				3		55	0.86	0.87	-0.32
		8-8AN	2	2		2		2																	
		9-8AP								1							2		2	3					
		10-8BI								1			2		2			2							
		11-8BN	3	3		3		4																	
		12-8BP										2	3	2			3								
12	556 (US508)	13-12AI											3			3		4	3			82	0.79	0.8	0.34
		14-12AN	4	4		4		4																	
		15-12AP										2	3			2	2								
		16-12BI									3	2					3		4						
		17-12BN	5	5		4		5																	
		18-12BP											3				3		3	4					
	562 (US514)	13-12AI											3			4		4	4			82	0.77	0.78	1.23
		14-12AN	3	3		3		3																	
		15-12AP										4	3			3	4								
		16-12BI									3			3	3		3								
		17-12BN		4		4	4	4																	
		18-12BP											3	4		3				3					

*severe rater (logit measure > 0.00)

Table 5.2 displays raw scores of two US students from each of Grades 4, 8 and 12 respectively to show unfair cases.

Grade 4 Students US013 and US436 received equal total raw scores of 104 out of a possible 144 from different combinations of 4 raters on the 6 prompts. Student US013's essays were scored by 7 severe raters to produce 13 ratings and 7 lenient raters to produce 11 ratings. However, Student US436's total raw score of 104 came from 8 severe raters who produced 10 ratings and 9 lenient raters who produced 14 ratings. Now, the MFRM estimates of student writing ability are based directly on the identical raw scores (104 for Students US013 and US436), but the unequal severity of the rater pairs, as derived from the modeled estimates of rater severity in the MFRM analysis. This was unfair for Student US013 because relying on raw scores alone and ignoring rater severity would have underestimated that student's writing ability by more than 1 logit in this case (1.35 - 0.08 = 1.27 logits).

Similarly, the total raw scores for 6 essays of Grade 8 Students US118 and US125 are 55 out of a possible 144. The essays were scored by 4 US human raters in different combinations. The essays of Student US118 were graded by 7 severe raters to produce 10 ratings and 7 lenient raters to produce 14 ratings. On the other hand, Student US125 received 55 from 7 severe raters who produced 12 ratings and 7 lenient raters who produced 12 ratings. This would have been unfair for Student US125 because his raw scores would have underestimated that student's writing ability by approximately 2 logits in this research (-0.32 - (-2.17) = 1.85 logits).

For the highest level of students in this research, examples of the same potentially unfair situation were found. Grade 12 Students US508 and US514 received raw score totals of 82 out of a possible 144 each. The 6 essays of Student US508 were scored by 4 severe US human raters to produce 10 ratings and 8 lenient US human raters to produce 14 ratings. However, the raw score of 82 of Student US514 was from 6 severe raters who produced 12 ratings and 9 lenient raters who produced 12 ratings. The assessment of Student US514's writing ability relied on raw scores alone, and it caused the score to be underestimated by a considerable amount (1.23 - 0.34 = 0.89 logits).

On the other hand, there were no unusual human raw score cases with obvious discrepancy from their logit measures in the data set of HK student essays because the arithmetic means of raw scores for the HK student essays already contained all of the relevant rating information: every student's results contained at least one rating from each HK human rater in the rating plan. Thus, their average raw scores approximate their fair average facets scores in the MFRM analysis.

Rating Discrepancies between HK Raters

In essay writing assessment, it is a common training practice for raters to achieve agreement in understanding the descriptions or criteria of scoring rubrics before scoring essays. Seldom does the researcher have a post-scoring meeting with raters to share and discuss their scorings. For the purpose of this research, it is worthwhile to talk to raters about how they judged essays to understand their underlying criteria about particular prompts.

Fortunately, the HK raters were enthusiastic to provide their ideas and professional reflections on essay scoring for this research. This researcher conducted a post-scoring meeting with the four HK raters. Each rater was provided 2-3 essays with a maximum discrepancy of 4 out of 6 points in the scoring rubrics from another rater who scored the same student essay. After rescoring some of those essays, HK raters were asked the following questions to reflect how they would proceed to finalise, qualitatively, fair scores for the essays and how the results of essay scoring might affect students' development.

Questions asked of the HK raters included:

1. Here is an essay you already marked. What did you think when you rated the essay at first? Why?
2. You gave it a score of __. What does that make you think now? Why did you give this score to this student's essay?
3. You remember that every essay was scored by 2 raters from your school. What would you think if I told you that the essay was given a discrepant score of __ by another teacher?

4BI My Lunch Time

Describe what lunch time is like for you on a school day. Be sure to tell about your lunch time so that someone who has never had lunch with you on a school day can understand where you have lunch and what lunch time is like.

Scoring Rubrics

Score	Description
(6) Excellent	• Develops ideas well and uses specific, relevant details across the response. • Well organized with clear transitions. • Sustains varied sentence structure and exhibits specific word choices. • Exhibits control over sentence boundaries; errors in grammar, spelling, and mechanics do not interfere with understanding.
(5) Skillful	• Develops ideas with some specific, relevant details. • Clearly organized; information is presented in an orderly way, but response may lack transitions. • Exhibits some variety in sentence structure and exhibits some specific word choices. • Generally exhibits control over sentence boundaries; errors in grammar, spelling, and mechanics do not interfere with understanding.
(4) Sufficient	• Clear but sparsely developed; may have a few details. • Provides a clear sequence of information; provides pieces of information that are generally related to each other. • Generally has simple sentences and simple word choice; may exhibit uneven control over sentence boundaries. • Has sentences that consist mostly of complete, clear, distinct thoughts; errors in grammar, spelling, and mechanics generally do not interfere with understanding.
(3) Uneven	May be characterized by one or more of the following: • Provides limited or incomplete information; may be list-like or have the quality of an outline. • Disorganized or provides a disjointed sequence of information. • Exhibits uneven control over sentence boundaries and may have some inaccurate word choices. • Errors in grammar, spelling, and mechanics sometimes interfere with understanding.
(2) Insufficient	May be characterized by one or more of the following: • Provides little information and makes little attempt at development. • Very disorganized OR too brief for reader to detect organization. • Exhibits little control over sentence boundaries and sentence formation; word choice is inaccurate in much of the response. • Characterized by misspellings, missing words, incorrect word order; errors in grammar, spelling, and mechanics are severe enough to make understanding very difficult in much of the response.
(1) Unsatisfactory	May be characterized by one or more of the following: • Attempts a response, but may only paraphrase the prompt or be extremely brief. • Exhibits no control over organization. • Exhibits no control over sentence formation; word choice is inaccurate across the response. • Characterized by misspellings, missing words, incorrect word order; errors in grammar, spelling, and mechanics severely impede understanding across the response.

Essay 1 Student 648 (HK009)

The bells rings, it time to lunch, I sit it on the chair and wait for my lunchbox. I start to eat my lunch box! I will eat pizza less. Sometimes, I will eat noodles. I always eat rice! I never eat congee. The food are not delicious. I eat a few and put the lunch box in a big box. After has already eatten my lunch box. I will do my homeworks because I want to watch TV at night. When I doing my homeworks, some of my classmathes will talk to their friend. It is very noisy! When the ring bell agins, It time to recult time!

HKR003's rating: 5	**HKR004's rating: 2**
1. This student provides a clear sequence of information in the essay, so it's worth a score of 4. 2. *R: You gave it a score of 5.* Again, this essay is clearly organised. I think this essay is worth 4 or 5 out of 6. 3. *R: The essay was given a score of 2 by another rater.* I think another rater's interpretation of the word 'little' might be different from mine, so from her perspective, this student only provides 'little'	1. The performance of this student fulfilled the requirements of the prompt to describe what lunch time is like on a school day and provided relevant details. For example, he mentioned 'put the lunch box in a big box...', it is a key step to tidy up the empty lunch box after having lunch. I can see what exactly happened at lunch time in school. So, I give this student a score of 5. 2. *R: You gave it a score of 2.* I might compare this student's performance with the other students.

information and makes 'little' attempt at development. And it also might be affected by our personal essay rating experiences.	There must have been something that distracted my focus when I scored this essay, but I have forgotten what it was. 3. *R: The essay was given a score of 5 by another rater.* I agree with that rater now.

Note: R= this researcher

Essay 2 Student 663 (HK024)

Do you like having lunch in your school? What will you do during your lunch time? I like having lunch in my school very much, it's really fun. During lunch time, I will have lunch with my classmate. Sometimes, teacher will play some interesting videos, it is very fun! Sometimes, I will talk to my friends. Sometimes, my friends will tell us some jokes too, it's very funny! After having lunch, if I still have some times, I will do my homework. If I finish all my homework, I will read books or talk to my friends. Some students will play in the classroom and some students will talk to teachers too! Having lunch in my school is really fun! It is also an enjoyable experience too!

HKR002's rating: 5	HKR004's rating: 3
1. I did enjoy reading this essay because it is interesting. I want to know the ending of the essay. I give it a score of 6, the highest score in the rubrics. I use	1. I give this essay a score of 4 because it shows a few details and provides a clear sequence of information, but it can't show some specific, relevant

the word ‘enjoy’, as this is the only essay I need to score; in this case, I have more free time to read through every detail of the story, which is completely different from the feeling when I needed to score a hundred essays. 2. *R: You gave it a score of 5.* I’m not sure why I gave it 5, but that score is not very different from what I think it should be. 3. *R: The essay was given a score of 3 by another rater.* Really??!! I can’t understand why another rater gave a score of 3. It’s such a low score. This student uses varied sentence structures and specific details. I love its ending.	details in the level of skillful (the score of 5); e.g., This student doesn’t mention any food they have for lunch. 2. *R: You gave it a score of 3.* I might treat ‘food items’ to be an essential element in this essay. I thought this essay only provides limited or incomplete information, so I gave it only a score of 3. 3. *R: The essay was given a score of 5 by another rater.* I don’t agree with her because I couldn’t find specific and relevant details in this essay.

Note: R= this researcher

4BN The Very Unusual Day

One day you wake up and go down to breakfast. You eat what you normally eat. Your breakfast is the last normal thing that happens to you all day. Write a story called 'The Very Unusual Day' about what happens that day, from right after breakfast until you go to bed again.

Scoring Rubrics

Score	Description
(6) Excellent	● Tells a well-developed story with relevant descriptive details across the response. ● Events are well connected and tie the story together with transitions across the response. ● Sustains varied sentence structure and exhibits specific word choices. ● Exhibits control over sentence boundaries; errors in grammar, spelling, and mechanics do not interfere with understanding.
(5) Skillful	● Tells a clear story with some development, including some relevant descriptive details. ● Events are connected in much of the response; may lack some transitions. ● Exhibits some variety in sentence structure and exhibits some specific word choices. ● Generally exhibits control over sentence boundaries; errors in grammar, spelling, and mechanics do not interfere with understanding.
(4) Sufficient	● Tells a clear story with little development; has few details. ● Events are generally related; may contain brief digressions or inconsistencies. ● Generally has simple sentences and simple word choice; may exhibit uneven control over sentence boundaries. ● Has sentences that consist mostly of complete, clear, distinct thoughts; errors in grammar, spelling, and mechanics generally do not interfere with understanding.
(3) Uneven	May be characterized by one or more of the following: ● Attempts to tell a story, but tells only part of a story, gives a plan for a story, or is list-like. ● Lacks a clear progression of events; elements may not fit together or be in sequence. ● Exhibits uneven control over sentence boundaries and may have some inaccurate word choices. ● Errors in grammar, spelling, and mechanics sometimes interfere with understanding.
(2) Insufficient	May be characterized by one or more of the following: ● Attempts a response, but is no more than a fragment or the beginning of a story OR is very repetitive. ● Very disorganized OR too brief to detect organization. ● Exhibits little control over sentence boundaries and sentence formation; word choice is inaccurate in much of the response. ● Characterized by misspellings, missing words, incorrect word order, errors in grammar, spelling, and mechanics are severe enough to make understanding very difficult in much of the response.

(1) Unsatisfactory	May be characterized by one or more of the following: ● Attempts a response but may only paraphrase the prompt or be extremely brief. ● Exhibits no control over organization. ● Exhibits no control over sentence formation; word choice is inaccurate across the response. ● Characterized by misspellings, missing words, incorrect word order; errors in grammar, spelling, and mechanics severely impede understanding across the response.

Essay 3 Student 698 (HK059)

After I finsh my breakfast, I go back my home. Then, I swich on the TV. Suddenly, I feel so tired, so I want sleep. When I wake up, I am in a rainforest. I see the animals playing with the people, such as monkey, sharks, crocodiles...They playing happily. I think 'Why do the animals will play happily with the people?'

One of the child say to me 'Why you not play with us?' After that I play with they. I want everyday can play happily with the animals. Then, I see angle, and I ask 'Why in the world the people can play happily with the animals?' 'Because in the, the people will not kill the animals and eat the animals, so that they can play happily with the animals.' The angle tall me.

'Wake up! Wake up!' My mum say. I found me is sleeping in my home. After that, I tall the government we should not kill the animals anymore.

Form them on, we play happily with the animals.

HKR001's rating: 5	HKR004's rating: 2
1. This essay fulfils the level 4 (sufficient) criteria including simple sentences and events are generally related.	1. This student tells a clear story with some relevant details. There is an advice to the government at the end of the story. It is meaningful. So I give this essay a score of 5.

2. *R: You gave it a score of 5.* I am one of the teachers who teaches civic-moral education. The sentence '...the government we should not kill the animals anymore' impressed me because this student cares about animals. I might compare the performance of other students to adjust the scoring, so it is difficult for me to assess only one essay. 3. *R: The essay was given a score of 2 by another rater.* Another rater might focus on the student's grammatical mistakes. She might find that the mistakes interfere with understanding.	2. *R: You gave it a score of 2.* This prompt is about 'the very unusual day'. I remembered that the time for students to write this story while they were learning about how to protect animals, I might find that the content of this essay was not unusual at that moment; it is exactly the same content that the students have learnt before. It involves their previous knowledge, so I only gave it a score of 2. 3. *R: The essay was given a score of 5 by another rater.* I agree with that rater now.

Note: R= this researcher

4BP My Favourite Book

Your favourite book is missing from your school library. It might be a book that you like to read over and over again. Or it might be a book that your teacher or parent has read to you. Some of your friends also like to read this book. The school librarian is not sure she wants to buy the book again. Write a letter to convince your school librarian to buy the book again. In your letter, give lots of reasons why the book should be in your school library.

Scoring rubrics

Score	Description
(6) Excellent	● Takes a clear position and develops support with well-chosen details, reasons, or examples across the response. ● Well organized; maintains focus. ● Sustains varied sentence structure and exhibits specific word choices. ● Exhibits control over sentence boundaries; errors in grammar, spelling, and mechanics do not interfere with understanding.
(5) Skillful	● Takes a clear position and develops support with some specific details, reasons, or examples. ● Provides some organization of ideas by, for example, using contrast or building to a point. ● Exhibits some variety in sentence structure and exhibits some specific word choices. ● Generally exhibits control over sentence boundaries; errors in grammar, spelling, and mechanics do not interfere with understanding.
(4) Sufficient	● Takes a clear position with support that is clear and generally related to the issue. ● Generally organized. ● Generally has simple sentences and simple word choice; may exhibit uneven control over sentence boundaries. ● Has sentences that consist mostly of complete, clear, distinct thoughts; errors in grammar, spelling, and mechanics generally do not interfere with understanding.
(3) Uneven	May be characterized by one or more of the following: ● Takes a position and offers limited or incomplete support; some reasons may not be clear or related to the issue. ● Disorganized OR provides a disjointed sequence of information. ● Exhibits uneven control over sentence boundaries and may have some inaccurate word choices. ● Errors in grammar, spelling, and mechanics sometimes interfere with understanding.
(2) Insufficient	May be characterized by one or more of the following: ● Takes a position, but provides only minimal support (generalizations or a specific reason or example); OR attempts to take a position but the position is unclear. ● Very disorganized or too brief to detect organization. ● May exhibit little control over sentence boundaries and sentence

	formation; word choice is inaccurate in much of the response. ● Characterized by misspellings, missing words, incorrect word order; errors in grammar, spelling, and mechanics may be severe enough to ● make understanding very difficult in much of the response.
(1) Unsatisfactory	May be characterized by one or more of the following: ● Takes a position but provides no support, OR attempts to take a position (is on topic) but the position is very unclear; may only paraphrase the prompt. ● Exhibits no control over organization. ● Exhibits no control over sentence formation; word choice is inaccurate across the response. ● Characterized by misspellings, missing words, incorrect word order; errors in grammar, spelling, and mechanics severely impede understanding across the response.

Essay 4 Student 725 (HK086)

Dear school librarian:

I was missing a favourite book from your school library, this book name is 'FULLMETAL ALCHEMIST' this book have many people to reading, this book my teachers like reading too, so you wants to buy ths book again.

HKR001's rating: 2	HKR003's rating: 6
1. This essay fulfils most of the level 3 criteria, so I give it a score of 3. 2. *R: You gave it a score of 2.* I eliminated the possibility for this student to get a score of 4. I think I struggled between giving this student 2 or 3 when I scored this essay last time. 3. *R: The essay was given a score of 6 by*	1. I give this essay a score of 3 because this student provides a disjointed sequence of information in the essay. 2. *R: You gave it a score of 6.* I thought the name of this student's favourite book, FULLMETAL ALCHEMIST, impressed me, as it fits the description of level of excellent. This student takes a clear

another rater. I do want to know how another rater interprets the criteria of rubrics.	position and develops support with well-chosen details in the story. 3. *R: The essay was given a score of 2 by another rater.* That rater might emphasise the criterion of organisation as the focus when she scored this essay.

Note: R= this researcher

Essay 5 Student 764 (HK125)

Dear school librarian,

I am P.6E student, Jacky Ng. My favourit book Q-spy is missing from our our school library. I am very sad because I like to read over and over again. Some of my classmates also like to read this book. I am very enjoy reading this book in the library. My parent has read this book to me. So, I hope you can buy this book again. Thanks.

Love,

Jacky Ng.

HKR002's rating: 2	**HKR003's rating: 4**
1. This student states a clear position with adequate support, e.g., 'I like to read over and over again. Some of my classmates also like to read this book. My parent has read this book to me.' Thus, I give this essay a score of 4. 2. *R: You gave it a score of 2.* I thought I treated those three points I mentioned to be only one point when I scored this essay the last time, so I might think this student provides only minimal support in this essay. 3. *R: The essay was given a score of 4 by*	1. This essay shows a generally organised structure. The words 'sad' and 'enjoy' convince me as a librarian to buy the book again. I think it warranted a score of 4. 2. *R: You gave it a score of 4.* I have a consistent rating for this essay. 3. *R: The essay was given a score of 2 by another rater.* She might care about the number of words of the essay, but this is not what I considered when I scored it.

another rater. Now I think I understand why she gave a score of 4 to this essay.	

Note: R= this researcher

Questions asked of the HK raters - Ideas from them are summarised below:

4. As a professional rater/teacher, what do you do when you are given rubrics and essays? How do you prepare for scoring?
   - I read through the details of the scoring rubrics and understand them.
   - If I have an opportunity to contact the other raters, I will try to compromise with them about the meaning of the rubrics.
   - I will have to adjust the scores after reading all the essays I need to score.
   - I will try to understand the background/previous knowledge of the students.
   - For the school examinations, I will read the essays at least twice. Then, I will focus on one area of criteria for each reading, e.g., content or language.

5. What are the reasons for the discrepancy between the other three raters and you? What do you think about your severity? Why?
   - The discrepancy may be affected by our scoring experiences or educational background.
   - The standards of good writing may differ among us.
   - The interpretation of the scoring rubrics is the key factor that causes discrepancy.
   - If content or wording of the essays is directly taken from books or supplementary exercises, the essays seldom earn high marks from me.
   - Raters may have discrepancy with the others or with themselves on scoring essays in the different time of a day or different period of a year.
   - Raters' scoring seriousness may be affected by the purpose of essay assessment.
   - Raters' judgement may be affected by the number of essays.
   - I don't have any ideas about my severity because it is hard to compare

among teachers or prompts, and I don’t know the ways to compare my severity with others.

6. What would you like to tell me now about the way of HK teachers score English essays?
    - Teachers seldom refer to the rubrics to score essays.
    - HK teachers may have a process of moderation especially for examinations.
    - Teachers design tailor-made scoring rubrics for various prompts/genres according to the rubrics of high-stakes assessment.
    - They may emphasise the correctness of grammar usage.
    - They provide feedback on students’ strengthens and weaknesses.
    - Teachers take turns to score all essays in the examination, but there is still a potential problem of consistency.

7. How would you think if the large difference of score is in a high-stakes examination?
    - It cannot detect the students’ writing abilities.
    - It is unfair for students, but it is commonplace in examinations.

The qualitative approach helps understand raters’ behaviour through interviews with 4 HK raters. e.g., The instructions in the 4BN prompt requires students to describe a very unusual day, HKR004 claimed that she seldom gave a high score to a student who responded to this prompt with a particular topic that the student was learning at that period. HKR001 teaches civic-moral education in school, so this rater would be easily impressed by students’ positive attitude of treating animals and gave them high scores.

From the teachers' responses, they showed that the main factor for the discrepancy is raters have various interpretations of the criteria in the rubrics with themselves or others over different periods of time. On the other hand, this problem has never happened by using the Lexile Analyzer or IEA to score essays because they can score essays consistently in this research so these two AES systems might help reduce teachers' scoring workload, especially in classroom practices.

Most importantly, HKR004 highlighted that teachers never know their comparative essay scoring severity because they do not have an appropriate method to assess it. The MFRM provides a rigorous analysis to reveal the severity of raters in a writing assessment. It provides practical and reliable information for teachers to understand their rating severity on essay scoring.

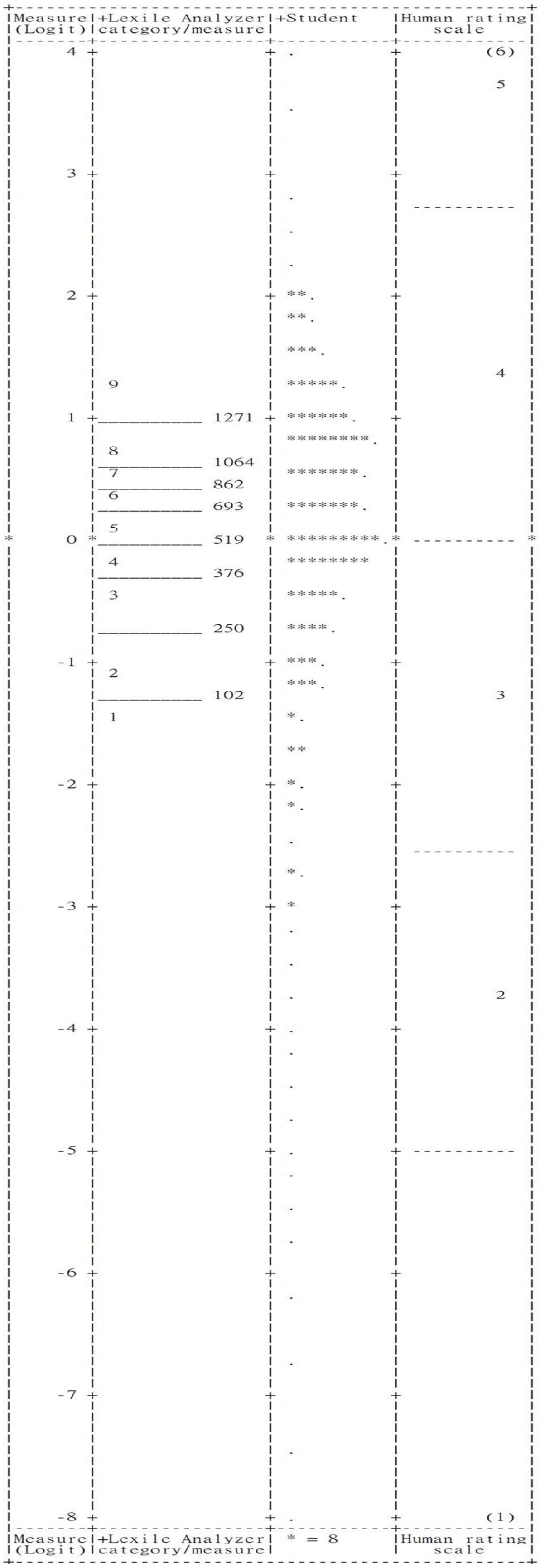


Figure 5.1 Vertical map for all student essays with scoring of Lexile Analyzer

Calibration of the Lexile Analyzer Measures against Human Ratings

From the analyses in this research, the calibration of the Lexile Analyzer against human ratings can be established for HK teachers' reference on scoring essays. Figure 5.1 shows the cut-off points of 9 categories developed for the Lexile Analyzer and the relationship between the Lexile Analyzer measures and human ratings. The facet for the Lexile Analyzer was positive in this FACETS specification file; in contrast to the analyses in Chapter Four (i.e., in the analyses in Chapter Four, the Lexile Analyzer facet has been set as negative by default, and only the student facet was positive to present a higher score implied a higher measure on the scale).

If the essay is scored by the Lexile Analyzer, then an approximate teacher's rating could be derived from Figure 5.1. For example, if a student's essay receives the Lexile measure of 690, its human score is inferred to be 4 out of 6.

Consequential Research Question:

6. Do/could the Lexile Analyzer and IEA provide meaningful error-detection feedback? If so, how do they detect/report? If not, can any suggestions be made?

In English essay writing, two aspects of feedback can be provided: students' strengths and weaknesses of essay writing on particular prompts and growth of student writing ability. The Lexile Analyzer provides only a Lexile measure as feedback. It is appropriate for anyone who submits his/her essay to determine the appropriate Lexile level of reading materials to improve their English abilities in the future.

IEA provides an overall score and feedback on spelling, grammar and redundancy. Moreover, reference scores (without detailed descriptions) are given for six areas including ideas and content, organisation, sentence fluency, word choice, conventions and voice (Pearson, 2010).

The Lexile Analyzer does not function as do other AES systems, which often require hundreds of scripts on a single topic to develop their own potentially reliable rating metric (Dikli, 2006). The Lexile Analyzer does not require exposure to any essays for training purposes, so it could easily be used by classroom teachers as an adjunct essay scoring device used for classroom writing assessment. If the software was used to score some classroom writing tasks, teachers would be free to concentrate on more important aspects of teaching and learning English (Breland, 1996). For example, teachers could focus on providing individual feedback to students on their performance in particular writing tasks.

Moreover, the Lexile Analyzer can provide a reliable record/feedback on the growth of student writing ability over time from one grade level to the next. This is usually difficult to obtain because students' English language teachers change from year to year and teachers do not used a calibrated rating scale for marking essays over time. Computer ratings could be expected to be psychometrically superior to the usual classroom human ratings. Essays could be described statistically in many different ways and used to study group differences, yearly trends, and a host of other important research questions (Page & Petersen, 1995).

Limitations and Suggestions for Future Research

This research involved hundreds students from the United States and Hong Kong, nineteen human raters from the United States and four human raters from only one Hong Kong primary school. The writing and rating sample size of the students and raters who learn English as second language was limited. The essays were not given to larger groups of students and teachers (who are raters) from different schools or school systems in Hong Kong; that would be essential to enable the establishment of a standard scale to measure the relative rating efficiency of the various school/assessment systems. If such a scale was generated, it would help determine the factors that influence student writing ability and essay rating criteria and teacher severity under the various existing conditions in Hong Kong schools.

In Chapter Two, six AES systems in the field of education including the Lexile Analyzer are introduced. However, the Lexile Analyzer and the Bayesian Essay Test Scoring sYstem (BETSY) are two free AES engines to educators for non-commercial use. The ideal data set for large-scale research would have different AES measures for the essays. In fact because the measures of the Lexile Analyzer are the only AES measures for the Hong Kong student essays in this research, it seems that it is not easy to have another set of AES measures for the 408 Hong Kong student essays, even though the MetaMetrics Inc. had spent more than two years communicating with various relevant companies/organisations in the United States to try to achieve this. Otherwise, the findings of such a large-scale research project must serve as a reference guide for educators, teachers, students or anyone involved in the writing assessment systems to choose the most suitable AES system(s) for their needs.

Hillegas (1912) claimed that it is possible to develop a perfect scale if the service of many more raters could be required to account for the variation among existing or previous raters. The calibration of essays and raters is improved when the network has every rater to grade every essay in the design (Lunz, Wright, & Linacre, 1990). Future research could involve the US raters scoring the HK student essays. If the HK student essays are scored by the US raters, the comparison of human raters would be well-rounded to serve as a reference guide for policy makers, educators, teachers, students or anyone involved in the writing assessment systems in both learning English as a first and second language.

Conclusion

Effective language learning outcomes remain a function of interactions coupled with sensitive and compassionate language teachers and students within a creative learning environment. Of paramount importance are the educational philosophy and practices associated with the use of technology (Chavez, 1990). AES is now accepted as a tool to complement, but not replace, expert human raters, no matter how closely the automated scores approximate the scores by expert human raters (Rudner, Garcia, & Welch, 2006; Warschauer & Grimes, 2008).

Although AES systems have substantial blind spots, including insensitivity to connotation and context, they could be used in different ways in different settings. Teachers can deploy AES systems to best meet their students' and their own needs (Warschauer & Grimes, 2008). Because the Lexile Analyzer can help ensure the maintenance of fairness and consistency while reducing teachers' workload, it provides insights for the stakeholders involved in English essay assessment in the educational systems, especially for Hong Kong English language teachers.

Using the Lexile Analyzer for classroom assessment or for students' self-learning practices in language rooms in Hong Kong schools, perhaps 30%-50% of all formative assessment essays could be scored by the Lexile Analyzer only. This AES system might be introduced on a trial basis for one semester/year in Grade 4 because students in this level do not face high-stakes assessment, other than school-based summative assessment, and the proportion of essays in classroom assessment could be gradually increased and other year levels could be involved in the trial in the next academic year. Teachers would then save time scoring essays, and they could spend more time conferencing with students individually and providing verbal feedback on the students' strengths and weaknesses of the teacher-graded classroom essay assessment. Moreover, conferencing with students on these writing pieces could have a positive impact on the motivation for students' English language learning.

A format of essay double marking could be used in school-based examinations: each essay would be scored by the Lexile Analyzer and one English teacher. The Lexile Analyzer measures could be a reference point for moderation across classes. For each essay, the mean of differences between the Lexile Analyzer measures and scores from teachers could be obtained; if essays have far more than the average of discrepancies between the Lexile Analyzer measure and human rating, the essays would be rescored by another English teacher. This moderation strategy is much more objective and effective than having teachers to take turns to score all student essays in the same year level in the examinations. English teachers would be provided more reliable information on their severity of essay scoring in assessment, helping to support English teachers' professional development on essay scoring and establishing a fairer English essay scoring system for high-stake assessment in school.

Teachers might register at the website of MetaMetrics Inc. to access to the Lexile Analyzer. The English subject plan would suggest number of essays for classroom practices to be scored by the Lexile Analyzer only in any semester/school year. Then, a column could be provided in each student's individual school report to show that student's Lexile Analyzer measures for each essay, and the means of the student measures, the whole class measures and the whole year level measures could be reference for the year growth of students' writing abilities. Based on student annual reports of the Lexile Analyzer measures from Grades 1 to 6 at primary level or Grades 7 to 13 at secondary level, English teachers (in different grade levels) or parents might more readily monitor the progress of student writing abilities. Moreover, schools could also make use of the Lexile Analyzer measures as reference points for students' reading abilities to guide them in choosing English reading materials appropriately difficult for the development of their English reading skills.

If scores of student essays could be provided by a free online AES system, the Lexile Analyzer, it could be a tool to lighten Hong Kong English language teachers' grading load in classroom rountine practices or low-stakes assessment, and it may help improve the quality of moderation between teachers in school-based examinations. To receive the Lexile Analyzer measures, no extra effort for English teachers and no human scored essays for system training purposes are needed; although the system only counts words and gives a rating to an essay, the Lexile Analyzer measures could be given a piece of extra useful information about student writing abilities to teachers or parents for tracking their performance from year to year. With more information about students' writing and reading levels from the Lexile Analyzer measures, Hong Kong schools could have well development of the school-based English curriculum, and the MFRM analyses of essay assessment are beneficial to English teachers' professional development and the design of fair English essay assessment systems.

References

Applebee, A. N. (1984). *Contexts for Learning to Write: Studies of Secondary School Instruction.* Norwood, N. J.: Ablex.

Bacha, N. (2001). Writing evaluation: What can analytic versus holistic essay scoring tell us? *System, 29*(3), 371-383.

Bereiter, C. (2003). Foreword. In M. D. Shermis, & J. C. Burstein (Eds.), *Automated Essay Scoring: A cross disciplinary approach*. Mahwah, N.J.: Lawrence Erlbaum Associates.

Bond, T. G. (2004). Validity and assessment: A Rasch measurement perspective. *Metodologia de las Ciencias del Comportamiento*, *5*(2), 179-194.

Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences* (2nd ed.). Mahwah, N.J.: Lawrence Erlbaum Associates.

Borman, W. C. (1978). Exploring upper limits of reliability and validity in job performance ratings. *Journal of Applied Psychology 63*(2), 135-144.

Breland, H. M. (1996). Computer-assisted writing assessment: The politics of science versus the humanities. In E. M. White, W. D. Lutz, & S. Kamusikiri (Eds.), *Assessment of writing: Politics, policies, practices* (pp. 249-256). New York: Modern Language Association of America.

Brookhart, S. M. (1999). *The art and science of classroom assessment: The missing part of pedagogy.* Washington: The George Washington University Press.

Burkhalter, N. (1995). A Vygotsky-based curriculum for teaching persuasive writing in the elementary grades. *Language Arts, 72*(3), 192-199.

Burstein, J. (2003). The e-rater scoring engine: Automated Essay Scoring with natural language processing. In M. D. Shermis, & J. C. Burstein (Eds.), *Automated Essay Scoring: A cross disciplinary approach* (pp. 113-121). Mahwah, N.J.: Lawrence Erlbaum Associates.

Burstein, J., Braden-Harder, L., Chodorow, M., Hua, S., Kaplan, B., Kukich, K., Lu, C., Nolan, J., Rock, D., & Wolff, S. (1998). *Computer analysis of essay content for automated score prediction.* Educational Testing Service. Princeton, NJ.

Burstein, J., & Chodorow, M. (2002). Directions in automated essay analysis. In R. B. Kaplan (Ed.), *The Oxford handbook of applied linguistics* (pp. 487-497). Oxford: University of Oxford Press.

Calfee, R. (2000). To grade or not to grade. In M. A. Hearst (Ed.), The debate on automated essay grading. *IEEE Intelligent Systems, 15*(5), 35-37.

Carlisle, R., & McKenna, E. (1991). Placement of ESL/EFL undergraduate writers in college-level writing programs. In L. Hamp-Lyons (Ed.), *Assessing second language writing in academic contexts* (pp. 197-211). Norwood, N.J.: Ablex.

Carlson, S. B. (1991). Program evaluation procedures: Reporting the program publicly within the political context. In L. Hamp-Lyons (Ed.), *Assessing second language writing in academic contexts* (pp. 293-320). Norwood, N.J.: Ablex.

Carroll, J. B., Davies, P., & Richman, B. (1971). *Word frequency book.* Boston: Houghton Mifflin.

Chall, J. S. (1983). *Stages of reading development*. New York: McGraw-Hill.

Chall, J. S. (1988). The beginning years. In B. L. Zakaluk, & S. J. Samuels (Eds.), *Readability: Its past, present, and future* (pp. 2-13). Newark, DE: International Reading Association.

Charney, D. (1984). The validity of using holistic scoring to evaluate writing: A critical overview. *Research in the Teaching of English, 18*(1), 65-81.

Chavez, R. (1990). The development of story writing within an IBM writing to read program lab among language minority students: Preliminary findings of a naturalistic study. In C. J. Faltis, & R. A. DeVillar (Eds.), *Language minority students and computer* (pp. 121-144). New York: The Haworth Press.

Chung, K. W. K., & Baker, E. L. (2003). Issues in the reliability and validity of automated scoring of constructed responses. In M. D. Shermis, & J. C. Burstein (Eds.), *Automated Essay Scoring: A cross disciplinary approach* (pp. 23-40). Mahwah, N.J.: Lawrence Erlbaum Associates.

Chung, K. W. K., & O'Neil, H. F. (1997). *Methodological approaches to online scoring of essays.* Retrieved 3 March 2009, from http://research.cse.ucla.edu/Reports/TECH461.pdf

Coniam, D. (2009a). A comparison of onscreen and paper-based marking in the Hong Kong public examination system. *Educational Research and Evaluation, 15*(3), 243-263.

Coniam, D. (2009b). Experimenting with a computer essay-scoring program based on ESL student writing scripts. *ReCALL Journal, 21*(2), 259-279.

Cooper, C. R. (1977). Holistic evaluation of writing. In C. R. Cooper, & L. Odell (Eds.), *Evaluating writing: Describing, measuring, judging.* Buffalo: State University of New York, Buffalo.

Crowhurst, M., & Piche, G. L. (1979). Audience and mode of discourse effects on syntactic complexity in writing at two grade levels. *Research in the Teaching of English, 13*(2), 101-109.

Cumming, A. (1990). Expertise in evaluating second language compositions. *Language Testing, 7*(1), 31-51.

Diederich, P. B. (1974). *Measuring growth in English.* Urbana, ILL.: National Council of Teachers of English.

Diederich, P. B., French, J. W., & Carlton, S. T. (1961). Factors in judgments of writing ability. *Research Bulletin, 61*(15). Princeton, N.J.: Educational Testing Services.

Dikli, S. (2006). An overview of automated scoring of essays. *Journal of Technology, Learning and Assessment, 5*(1), 49-62.

Dowling, M. (2008). *Examining the Exams.* The Hong Kong Examinations and Assessment Authority. Retrieved 3 May 2009, from http://www.hkeaa.edu.hk/DocLibrary/Media/articles/markdowlinge.pdf

Elliot, S. (2003). IntelliMetric^{TM}: From here to validity. In M. D. Shermis, & J. C. Burstein (Eds.), *Automated Essay Scoring: A cross disciplinary perspective* (pp. 71-86). Mahwah, N.J.: Lawrence Erlbaum Associates.

Engelhard, G. (1992). The measurement of writing ability with a many-faceted Rasch model. *Applied Measurement in Education, 5*(3), 171-191.

Engelhard, G. (1994). Examining rater error in the assessment of written composition with a many-faceted Rasch model. *Journal of Educational Measurement, 31*(2), 93-112.

Engelhard, G. (1996). Evaluating rater accuracy in performance assessments. *Journal of Educational Measurement, 33*(1), 56-70.

Engelhard, G., Walker, E. S., Gordon, B., & Gabrielson, S. (1994). Writing tasks and gender: Influences on writing quality of black and white students. *Journal of Educational Research, 87*(4), 197-209.

Follman, J., & Anderson, J. (1967). An investigation of the reliability of five procedures for grading English themes. *Research in the Teaching of English, 1*(2), 190-200.

Foltz, P. W., Laham, D., & Landauer, T. K. (1999). The Intelligent Essay Assessor: Applications to Educational Technology. *Interactive Multimedia Electronic Journal of Computer-Enhanced Learning, 1* (2). http://imej.wfu.edu/articles/1999/2/04/index.asp

Freedman, S. (1979). How characteristics of student essays influence teachers' evaluations. *Journal of Educational Psychology, 71*(3), 328-338.

Godshalk, F., Swineford, F., & Coffman, W. (1966). *The measurement of writing abililty.* Princeton: English Testing Service.

Gottlieb, M. (2006). *Assessing English language learners: Bridges from language proficiency to academic achievement.* Thousand Oaks, CA: Corwin Press.

Graesser, A. C., McNamara, D. S., & Louwerse, M. M. (2003). What do readers need to learn in order to process coherence relations in narrative and expository text? In A. P. Sweet, & C. E. Snow (Eds.), *Rethinking Reading Comprehension* (pp.82-98). New York: Guilford.

Grobe, C. (1981). Syntactic maturity, mechanics and vocabulary as predictors of quality ratings. *Research in the Teaching of English, 15*(1), 75-85.

Hamp-Lyons, L. (1991a). Reconstructing ‘‘Academic writing proficiency”. In L. Hamp-Lyons (Ed.), *Assessing second language writing in academic contexts* (pp. 127-154). Norwood, N.J.: Ablex.

Hamp-Lyons, L. (1991b). Scoring procedures for ESL context. In L. Hamp-Lyons (Ed.), *Assessing second language writing in academic contexts* (pp. 241-276). Norwood, N.J.: Ablex.

Hamp-Lyons, L. (2000). Fairness in language testing. In A. J. Kunnan (Ed.), *Fairness and validation in language assessment: Selected papers from the 19th Language Testing Research Colloquium, Orlando, Florida* (pp. 30-34). Cambridge, England: Cambridge University Press.

Hamp-Lyons, L. (2001). Fourth generation writing assessment. In T. Silva, & P. K. Matsuda (Eds.), *On second language writing* (pp. 117-125). Mahwah, N.J.: Lawrence Erlbaum Associates.

Hamp-Lyons, L. (2007). The impact of testing practices on teaching: Ideologies and alternatives. In J. Cummins, & C. Davison (Eds.), *International handbook of English language teaching* (pp.487-504). Norwell: Springer.

Hamp-Lyons, L., & Prochnow, S. (1990). Prompt difficulty, task type, and performance. In S. Anivan (Ed.), *Current developments in language testing* (pp. 58-76). Singapore: SEA MEO/RELC.

Hedge, J. W., & Kavanagh, M. J. (1988). Improving the accuracy of performance evaluations: Comparison of three methods of performance appraiser training. *Journal of Applied Psychology, 73*(1), 68-73.

Hillegas, M. B. (1912). Scale for the measurement of quality in English composition by young people: The scale. *Teachers College Record*, *13*(4), 1-54.

Hong Kong Census and Statistics Department. (2010). Student Enrolment by Level of Education. Retrieved 22 September 2011, from http://www.censtatd.gov.hk/FileManager/EN/Content809/education.pdf

Hong Kong Education Department. (2000). *Circular Memorandum No. 501/2000.* Retrieved 3 June 2009, from http://www.edb.gov.hk/FileManager/EN/Content1376/cm501eng.pdf

Hong Kong Education Department. (2002). *English Language Education Key Learning Area Curriculum Guide (Primary 1 - Secondary 3)*. Curriculum Development Council.

Hong Kong Education and Manpower Bureau. (2004a). *English Language Curriculum Guide (Primary 1-6)*. Curriculum Development Council.

Hong Kong Education and Manpower Bureau. (2004b). *Promoting Assessment for Learning in English Language Education at Primary Level*. Curriculum Development Institute.

Hong Kong Education Bureau. (2007a). *English Language Curriculum and Assessment Guide (Secondary 4-6)*. Curriculum Development Council and Hong Kong Examinations and Assessment Authority.

Hong Kong Education Bureau. (2007b). *Policy and Initiatives of Special Education.* Retrieved 5 February 2010, from http://www.edb.gov.hk/FileManager/EN/Content7426/2009%20sped.pdf

Hong Kong Examinations and Assessment Authority. (2007a). *HKCEE English language examination report and question papers.* Hong Kong: Hong Kong Examinations and Assessment Authority.

Hong Kong Examinations and Assessment Authority. (2007b). *Onscreen Marking System.* Retrieved 3 May 2009, from http://www.hkeaa.edu.hk/DocLibrary/Media/Leaflets/ea-osm-eng.pdf

Hong Kong Examinations and Assessment Authority. (2008a). *Language Proficiency Assessment for Teachers (English Language) Handbook.* Retrieved 3 June 2009, from http://www.hkeaa.edu.hk/DocLibrary/Local/LanguageProficiencyAssessmentforTeachers/lpathandbookeng.pdf

Hong Kong Examinations and Assessment Authority. (2008b). *Territory-wide System Assessment Primary 6 English language question papers.* Retrieved 5 May 2009, from http://www.systemassessment.edu.hk/pri/2008pripaper/P6Eng/2008TSA6ERW1.pdf &

http://www.systemassessment.edu.hk/pri/2008pripaper/P6Eng/2008TSA6ERW4.pdf

Hong Kong Examinations and Assessment Authority. (2010). *Territory-wide System Assessment Report on the Basic Competencies of students in Chinese Language, English Language and Mathematics Key Stages 1-3.* Retrieved 22 June 2011, from http://www.bca.hkeaa.edu.hk/web/TSA/en/2010tsaReport/priSubjectreporteng.html

Huot, B. (1993). The influence of holistic scoring procedures on reading and rating student essays. In M. Williamson, & B. Hout (Eds.), *Validating holistic scoring for writing assessment: Theoretical and empirical foundations* (pp. 206-236). Cresskill, N.J.: Hampton Press.

Huot, B., & O'Neill, P. (Eds.). (2009). *Assessing writing: A critical sourcebook.* Urbana, IL: National Council of Teachers or English; New York: Bedford/St. Martin's.

James Cook University Research Committee. (2001). *James Cook University Statement and Guidelines on Research Practice.* Townsville, Queensland.

Janopoulos, M. (1993). Comprehension, communicative competence, and construct validity: Holistic scoring in an ESL perspective. In M. Williamson, & B. Hout (Eds.), *Validating holistic scoring for writing assessment: Theoretical and empirical foundations* (pp. 303-326). Cresskill, N.J.: Hampton Press.

Johnson, R. L., Penny, J., & Gordon, B. (2000). The relation between score resolution methods and interrater reliability: An empirical study of an analytic scoring rubric. *Applied Measurement in Education, 13*(2), 121-138.

Kukich, K. (2000). Beyond automated essay scoring. In M. A. Hearst (Ed.), The debate on automated essay grading. *IEEE Intelligent Systems, 15*(5), 22-27.

Landauer, T. K., Laham, D., & Foltz, P. W. (2003). Automated essay scoring and annotation of essays with the Intelligent Essay Assessor. In M. D. Shermis, & J. C. Burstein (Eds.), *Automated Essay Scoring: A cross disciplinary perspective* (pp. 87-112). Mahwah, N.J.: Lawrence Erlbaum Associates.

Lee, I. (2004). Error correction in L2 secondary writing classrooms: The case of Hong

Kong. *Journal of Second Language Writing, 13*(4), 285-312.

Legislative Council Panel on Education. (2005). *LC Paper No. CB(2)323/05-06(01): Grant to support the modernization and development of the Hong Kong Examinations and Assessment Authority's examination systems.* Retrieved 10 August 2010, from http://www.legco.gov.hk/yr05-06/english/panels/ed/papers/ed1114cb2-323-1e.pdf

Linacre, J. M. (1989). *Many-Facet Rasch Measurement.* Chicago: MESA Press.

Linacre, J. M. (1995). Categorical misfit statistics. *Rasch Measurement Transactions, 9*(3), 450-451. Retrieved 15 December 2010, from http://www.rasch.org/rmt/rmt93j.htm

Linacre, J. M. (1997). *Judging plans and facets (Research Note No. 3)*. Chicago: University of Chicage. MESA Psychometric Laboratory. Retrieved 9 May 2009, from http://www.rasch.org/rn3.htm

Linacre, J. M. (1999). Investigating rating scale category utility. *Journal of Outcome Measurement, 3*(2), 103-122.

Linacre, J. M. (2002). Optimizing rating scale category effectiveness. *Journal of Applied Measurement, 3*(1), 85-106.

Linacre, J. M. (2010a). *A user's guide to FACETS: Rasch-model computer programs*. Chicago, IL.

Linacre, J. M. (2010b). *A user's guide to WINSTEPS® /MINISTEP: Rasch-model computer programs.* Chicago, IL.

Linacre, J. M. (2010c). *FACETS* Version 3.67.0 [Computer software]. Beaverton, Oregon: Winsteps.com.

Linacre, J. M. (2010d). *WINSTEPS®* Version 3.70.0.4 [Computer software]. Beaverton, Oregon: Winsteps.com.

Linacre, J. M., Engelhard, G., Tatum, D. S., & Myford, C. M. (1994). Measurement with judges: Many-faceted conjoint measurement. *International Journal of Educational Research, 21*(6), 569-577.

Linacre, J. M., & Wright, B. D. (1989). The length of a logit. *Rasch Measurement Transactions, 3*(2), 54-55. Retrieved 25 July 2011, from http://www.rasch.org/rmt/rmt32b.htm

Linacre, J. M., & Wright, B. D. (2002). Construction of measures from Many-facet data. *Journal of Applied Measurement, 3*(4), 484-509.

Lloyd-Jones, R. (1977). Primary trait scoring. In C. Cooper, & L. Odell (Eds.), *Evaluating writing: Describing, measuring, judging* (pp. 33-66). Urbana, IL: National Council of Teachers of English.

Lunz, M. E., & Linacre, J. M. (1998). Measurement designs using multifacet Rasch modeling. In G. A. Marcoulides (Ed.), *Modern methods for business research* (pp. 44-77). Mahwah, N.J.: Lawrence Erlbaum Associates.

Lunz, M. E., Wright, B. D., & Linacre, J. M. (1990). Measuring the impact of judge severity on examination scores. *Applied Measurement in Education, 3*(4), 331-345.

Mara, M., Maskova, J., Fucikova, Z., Kuzel, D., Belsan, T., & Sosna, O. (2008). Midterm clinical and first reproductive results of a randomized controlled trial comparing uterine fibroid embolization and myomectomy. *Cardiovasc Intervent Radiol, 31*(1), 73-85.

McColly, W. (1970). What does educational research say about the judging of writing ability? *Journal of Educational Research, 64*(4), 147-156.

McIntyre, R. M., Smith, D. E., & Hassett, C. E. (1984). Accuracy of performance ratings as affected by rater training and perceived purpose of rater training and perceived purpose of rating. *Journal of Applied Psychology*, *69*(1), 147-156.

McNamara, T. F. (1996). *Measuring second language performance.* New York: Longman.

MetaMetrics, Inc. (2008a). *Texas higher education coordinating board: Text measurement and analysis (MetaMetrics technical report update).* Retrieved 25 February 2010, from http://www.lexile.com/m/resources/materials/MetaMetricsTechnicalReportUpdate.pdf

MetaMetrics, Inc. (2008b). *The Lexile Framework for Writing overview.* Retrieved 23 March 2009, from http://www.metametricsinc.com/TLFforWritngandMyWritingWeb.pdf

Moe, A. J. (1980). Analyzing text with computers. *Educational Technology, 20*(7), 29-31.

Moskal, B. M. (2000). Scoring rubrics: What, when and how? *Practical Assessment, Research & Evaluation, 7*(3). Retrieved 23 May 2010, from http://pareonline.net/getvn.asp?v=7&n=3

Moskal, B. M., & Leydens, J. A. (2000). Scoring rubric development: Validity and reliability. *Practical Assessment, Research & Evaluation*, *7*(10). Retrieved 23 September 2010, from http://pareonline.net/getvn.asp?v=7&n=10

Myers, M. (1980). *A procedure for writing assessment and holistic scoring.* Urbana, IL: National Council of Teachers of English.

Myers, M. (2003). What can computers and AES contribute to a K-12 writing program? In M. D. Shermis, & J. Burstein (Eds.), *Automated Essay Scoring: A cross-disciplinary perspective* (pp. 3-20). Mahwah, N.J.: Lawrence Erlbaum Associates.

National Health and Medical Research Council, Australian Research Council, & Australian Vice-Chancellors' Committee. (2007). *National statement on ethical conduct in human research.* ACT: Australian Government. Retrieved 24 September 2010, from http://www.nhmrc.gov.au/filesnhmrc/file/publications/synopses/e72-jul09.pdf

Nippold, M. A., Ward-Lonergan, J. M., & Fanning, J. L. (2005). Persuasive writing in children, adolescents, and adults: A study of syntactic, semantic, and pragmatic development. *Language, Speech, and Hearing Services in Schools, 36*(2), 125-138.

Page, E. B. (1966). The imminence of grading essays by computer. *Phi Delta Kappan, 47*(5), 238-243.

Page, E. B. (2003). Project Essay Grade: PEG. In M. D. Shermis, & J. Burstein (Eds.), *Automated Essay Scoring: A cross-disciplinary perspective* (pp. 43-54). Mahwah, NJ: Lawrence Erlbaum Associates.

Page, E. B., & Petersen, N. S. (1995). The computer moves into essay grading. *Phi Delta Kappan, 76*(7), 561-565.

Page, E. B., Poggio, J. P., & Keith, T. Z. (1997, March). *Computer analysis of student essays: Finding trait differences in the student profile.* Symposium given at American Educational Research Association meetings, Chicago.

Pearson. (2009a). *General overview of WriteToLearn™ and its components.* Retrieved 2 March 2010, from http://www.pearsonkt.com/papers/WTLOverview-040609.pdf

Pearson. (2009b). *PTE Academic Automated Scoring*. Retrieved 2 May 2009, from http://www.pearsonpte.com/SiteCollectionDocuments/AutomatedScoringUS.pdf

Pearson. (2010). *Intelligent Essay Assessor (IEA) fact sheet.* Retrieved 10 Dec 2011, from http://kt.pearsonassessments.com/download/IEA-FactSheet-20100401.pdf

Peregoy, S., & Boyle, O. (2005). *Reading, writing, and learning in ESL: A resource book for K-12 teachers.* White Plains, N.Y.: Longman.

Popham, W. J. (1990). *Modern Educational Measurement: A Practitioner's Perspective.* Englewood Cliffs, NJ: Prentice Hall.

Popham, W, J. (1997). What's wrong--and what's right--with rubrics. *Educational Leadership, 55*(2), 72-75.

Powers, D. E., Burstein, J. C., Chodorow, M. S., Fowles, M. E., & Kukich, K. (2002). Stumping e-rater: Challenging the validity of automated essay scoring. *Computers in Human Behavior, 18*(2), 103-134.

Qian, D. D. (2008). English language assessment in Hong Kong: A survey of practices, developments and issues. *Language Testing 25*(1), 85-110.

Quellmalz, E., & Burry, J. (1983). *Analytic scales for assessing students' expository and narrative writing skills* (CSE Resource Paper No. 5). Los Angeles: University of California, Center for the Study of Evaluation.

Quellmalz, E., Capell, F., & Chou, C. (1982). Effects of discourse and response mode on the measurement of writing competence. *Journal of Educational Measurement, 19*(4), 241-258.

Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests.* Chicago: University of Chicago Press.

Rudner, L., & Gagne, P. (2001). *An overview of three approaches to scoring written essays by computer.* Retrieved 10 March 2009, from http://pareonline.net/htm/v7n26.htm

Rudner, L., Garcia, V., & Welch, C. (2006). An evaluation of the IntelliMetric™ essay scoring system. *Journal of Technology, Learning, and Assessment, 4*(4), 86-107.

Rudner, L., & Liang, T. (2002). Automated essay scoring using Bayes' theorem. *Journal of Technology, Learning, and Assessment, 1*(2), 25-47.

Sakyi, A. A. (2000). Validation of holistic scoring for ESL writing assessment: how raters evaluate compositions. In A. J. Kunnan (Ed.), *Fairness and validation in language assessment: Selected papers from the 19th Language Testing Research Colloquium, Orlando, Florida* (pp. 129-152). Cambridge, England: Cambridge University Press.

Santos, T. (1988). Professors' reactions to the academic writing of non-native speaking students. *TESOL Quarterly, 22*(1), 69-90.

Schaefer, E. J. (2004). *Multi-faceted Rasch analysis and native-English-speaker ratings of Japanese EFL essays.* Japan: Temple University. Published doctoral thesis.

Shermis, M. D., & Barrera, F. D. (2002, April). *Exit assessments: Evaluating writing ability through Automated Essay Scoring.* Paper presented at the Annual Meeting of the American Educational Research Association. New Orleans, LA.

Shermis, M. D., & Burstein, J. C. (2003). *Automated Essay Scoring: A Cross-disciplinary perspective*. Mahwah, N.J.: Lawrence Erlbaum Associates.

Shermis, M. D., Koch, C. M., Page, E. B., Keith, T. Z., & Harrington, S. (2002). Trait ratings for automated essay grading. *Educational and Psychological Measurement*, *62*(1), 5-18.

Shermis, M. D., Mzumara, H. R., Olson, J., & Harrington, S. (2001). On-line grading of student essays: PEG goes on the World Wide Web. *Assessment & Evaluation in Higher Education, 26*(3), 248-259.

Smith, M. (2009). *The reading-writing connection.* Durham, NC: MetaMetrics Inc.

Starch, D., & Elliott, E. C. (1912). Reliability of the grading of high school work in English. *School Review, 20*, 442-457.

Stenner, A. J. (1999). *Instructional uses of the Lexile Framework.* Durham, NC: MetaMetrics Inc.

Stewart, M., & Grobe, C. (1979). Syntactic maturity, mechanics of writing and teachers' quality ratings. *Research in the Teaching of English, 13*(3), 207-215.

Swartz, C. W., Kroening, J. M., & Sanford-Moore, E. E. (2008). Using a developmental scale to enhance reporting of state-wide, high-stakes writing tests. *Innovation Quarterly, 3*(3), 7-8.

Thorndike, E. L. (1904). *An introduction to the theory of mental and social measurements.* New York: Teachers College, Columbia University.

U.S. Department of Education, Institute of Education Sciences. (1998a). *Genres of Prompts 1998.* National Center for Education Statistics. Retrieved 11 February 2009, from http://nces.ed.gov/nationsreportcard/

U.S. Department of Education, Institute of Education Sciences. (1998b). *Sample Questions Writing 1998.* National Center for Education Statistics. Retrieved 10 February 2009, from http://nces.ed.gov/nationsreportcard/

U.S. Department of Education, Institute of Education Sciences. (2002). *Sample Questions: Writing 2002.* National Center for Education Statistics. Retrieved 10 February 2009, from http://nces.ed.gov/nationsreportcard/

U.S. Department of Education, Institute of Education Sciences. (2009). *National*

Assessment of Educational Progress (NAEP) Overview. National Center for Education Statistics. Retrieved 10 February 2010, from http://nces.ed.gov/nationsreportcard/aboutnaep.asp

Valenti, S., Neri, F., & Cucchiarelli, A. (2003). An overview of current research on automated essay grading. *Journal of Information Technology Education, 2*, 319-330.

Vaughan, C. (1991). Holistic Assessment: What goes on in the raters' mind? In L. Hamp-Lyons (Ed.), *Assessing second language writing in academic contexts* (pp. 111-125). Norwood, NJ: Ablex.

Warschauer, M., & Grimes, D. (2008). Automated writing assessment in the classroom. *Pedagogies: An International Journal, 3*(1), 22-26.

Warschauer, M., & Ware, P. (2006). Automated writing evaluation: Defining the classroom research agenda. *Language Teaching Research 10*(2), 1-24.

Weigle, S. C. (1994). Effects of training on raters of ESL compositions. *Language Testing, 11*(2), 197-223.

White, E. (1984). Holisticism. *College Composition and Communication, 35*(4), 400-409.

Wohlpart, A. J., Lindsey, C., & Rademacher, C. (2008). The reliability of computer software to score essays: Innovations in a humanities course. *Computers and Composition, 25* (2), 203-223.

Wright, B. D. (1984). Despair and hope for educational measurement. *Contemporary Education Review, 3*(1), 281-288.

Wright, B. D. (1996). Reliability and separation. *Rasch Measurement Transactions, 9*(4), 472.

Wright, B. D., & Linacre, J. M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions, 8*(3), 370.

Wright, B. D., & Masters, G. N. (1981). *The measurement of knowledge and attitude*

(Research Memorandum No. 30). Chicago: University of Chicago, MESA Psychometric Laboratory.

Wright, B. D., & Masters, G. N. (1982). *Rating Scale Analysis: Rasch Measurement.* Chicago: MESA Press.

Wright, B. D., & Stone, M. H. (1979). *Best Test Design: Rasch Measurement.* Chicago: MESA Press.

Yan, Z. (2009). *The development of Rasch measurement physical fitness scale for Hong Kong primary school-aged students*. Australia: James Cook University. Published doctoral thesis.

Appendix 1

The prompts of genres were used in the study from the National Assessment of Educational Progress (NAEP) of the United States Department of Education (1998):

Appendix 2

Writing Prompts from the National Assessment of Educational Progress of the United States Department of Education

Prompt	Title	Year/Block
4AI	Describe a favorite object	1998-4W17
4AN	Write a story about a castle	1998-4W6
4AP	Convince a friend to become visible	1998-4W18
4BI	Describe lunchtime	2002-4W17
4BN	Describe a very unusual day*	2002-4W5
4BP	Convince the school librarian	2002-4W21
8AI	Design an educational TV show	1998-8W15
8AN	Write a story about a visitor from space	1998-8W9
8AP	Debate lengthening the school year	1998-8W19
8BI	Which book would you save?	2002-8W16
8BN	Your first day as president	2002-8W4
8BP	Support a school schedule	2002-8W20
12AI	Give writing advice to a younger student.	1998-12W10
12AN	Write a story about a special object	1998-12W5
12AP	Debate the importance of voting	1998-12W21
12BI	Which book would you save?	2002-12W14
12BN	Create a tall tale*	2002-12W3
12BP	Who are the heroes of today?	2002-12W20

* The wordings of the prompt are altered slightly from the NAEP wordings.

Appendix 3

Writing Prompts for the Students in the United States

Prompt	Title/Text of the Prompt	Year/Block of NAEP
4AI	Describe a favorite object We all have favorite objects that we care about and would not want to give up. Think of one object that is important to you. For example, it could be a book, a piece of clothing, a game, or any object you care about. Write about your favorite object. Be sure to describe the object and explain why it is valuable or important to you.	1998-4W17
4AN	Write a story about a castle One morning a child looks out the window and discovers that a huge castle has appeared overnight. The child rushes outside to the castle and hears strange sounds coming from it. Someone is living in the castle! The castle door creaks open. The child goes in. Write a story about who the child meets and what happens inside the castle.	1998-4W6
4AP	Convince a friend to become visible Pretend you have a friend who is invisible and you would like other people to meet him or her. Write a letter to this invisible person. Convince your friend to become visible so that others may meet him or her. In your letter, use details and examples.	1998-4W18

4BI	<u>Describe lunchtime</u> Describe what lunchtime is like for you on a school day. Be sure to tell about your lunchtime so that someone who has never had lunch with you on a school day can understand where you have lunch and what lunchtime is like.	2002-4W17
4BN*	<u>Describe very unusual day</u> One day you wake up and go down to breakfast. You eat what you normally eat. Your breakfast is the last normal thing that happens to you all day. Write a story called 'The Very Unusual Day' about what happens that day, from right after breakfast until you go to bed again.	2002-4W5
4BP	<u>Convince the school librarian</u> Imagine this situation: Your favorite book is missing from your school library. It might be a book that you like to read over and over again. Or it might be a book that your teacher or parent has read to you. Some of your friends also like to read this book. The school librarian is not sure she wants to buy the book again. Write a letter to convince your school librarian to buy the book again. In your letter, give lots of reasons why the book should be in your school library.	2002-4W21
8AI	<u>Design an educational TV show</u> A public television network is seeking ideas for a new series of shows that would be educational for teenagers. The series will include ten one-hour episodes and will be shown once a week. Some of the titles under consideration are: 'Great Cities of the World' 'Women in History'	1998-8W15

	‘Nature Walks’ ‘American Legends’ Choose one of these titles. Write a letter to the network describing your ideas for a new educational series. In your letter, describe what one episode might be like. Use specific examples of what information you would include in the episode so the network president will be able to imagine what the series would be like.	
8AN	<u>Write a story about a visitor from space</u> Imagine this situation! A noise outside awakens you one night. You look out the window and see a spaceship. The door of the spaceship opens, and out walks a space creature. What does the creature look like? What do you do? Write a story about what happens next.	1998-8W9
8AP	<u>Debate lengthening the school year</u> Many people think that students are not learning enough in school. They want to shorten most school vacations and make students spend more of the year in school. Other people think that lengthening the school year and shortening vacations is a bad idea because students use their vacations to learn important things outside of school. What is your opinion? Write a letter to your school board either in favor or against lengthening the school year. Give specific reasons to support your opinion that will convince the school board to agree with you.	1998-8W19
8BI	<u>Which book would you save?</u> A novel written in the 1950’s describes a world where people are not allowed to read books. A small group of people who want to save books memorize them so that the books won't be forgotten. For example, an old man who has memorized the	2002-8W16

	novel *The Call of the Wild* helps a young boy memorize it by reciting the story to him. In this way, the book is saved for the future. If you were told that you could save just one book for future generations, which book would you choose? Write an essay in which you discuss which book you would choose to save for future generations and what it is about the book that makes it important to save. Be sure to discuss in detail why the book is important to you and why it would be important to future generations.	
8BN	<u>Your first day as president</u> Imagine that you wake up one morning to discover that you have become the President of the United States. Write a story about your first day as President.	2002-8W4
8BP*	<u>Support a school schedule</u> Suppose a research study showed that teenagers have low energy levels in the morning and that adults have low energy levels at night. The study recommends that teenagers should stay up later at night and sleep later in the morning. The study also recommends that adults go to bed earlier and get up earlier. Write a letter to your principal arguing for or against the proposition that classes at your school should begin much later in the day. Be sure to give detailed reasons to support your argument and make it convincing.	2002-8W20

12AI	<u>Give writing advice to a younger student.</u> Your school has a program in which a twelfth grader acts as a mentor for a tenth grader at the beginning of each school year. The mentor’s job is to help the tenth grader have a successful experience at your school. The tenth grader you are working with is worried about being able to write well enough for high school classes. Write a letter to your tenth grader explaining what kind of writing is expected in high school classes and what the student can do to be a successful writer in high school. As you plan your responses, think about your own writing experiences. How would you describe ‘goo’ writing? What advice about writing has been helpful to you? What writing techniques do you use?	1998-12W10
12AN	<u>Write a story about a special object</u> The following excerpt is from a poem by Walt Whitman. There was a child who went forth every day, And the first object he look'd upon, that object he became, And that object became part of him for the day or a certain part of the day, Or for many years or stretching cycles of years. Whitman's poem suggests that certain objects become important to us and remain important to us even if we no longer have them. Write a story in which you tell about an object that remains important to the main character over a period of years. The main character could be you or someone you know. In your story, describe the main character's first encounter with the object, why the object is so important to the character, and how, over the years, it remains a part of the character's life.	1998-12W5

12AP	<u>Debate the importance of voting</u> Your school is sponsoring a voter registration drive for 18-year-old high school students. You and three of your friends are talking about the project. Your friends say the following, Friend 1: 'I'm working on the young voters' registration drive. Are you going to come to it and register? You're all 18, so you can do it. We're trying to help increase the number of young people who vote and it shouldn't be too hard - I read that the percentage of 18- to 30-year-olds who vote increased in recent years. We want that percentage to keep going up.' Friend 2: 'I'll be there. People should vote as soon as they turn 18. It's one of the responsibilities of living in a democracy.' Friend 3: 'I don't know if people should even bother to register. One vote in an election isn't going to change anything.' Do you agree with friend 2 or 3? Write a response to your friends in which you explain whether you will register to vote. Be sure to explain why and support your position with examples from your reading or experience. Try to convince the friend with whom you disagree that your position is the right one.	1998-12W21
12BI	<u>Which book would you save?</u> A novel written in the 1950's describes a world where people are not allowed to read books. A small group of people who want to save books memorize them so that the books won't be forgotten. For example, an old man who has memorized the novel *The Call of the Wild* helps a young boy memorize it by reciting the story to him. In this way, the book is saved for the future. If you were told that you could save just one book for future generations, which book would you choose?	2002-12W14

	Write an essay in which you discuss which book you would choose to save for future generations and what it is about the book that makes it important to save. Be sure to discuss in detail why the book is important to you and why it would be important to future generations.	
12BN*	<u>Create a tall tale</u> A tall tale is a type of story that uses exaggeration to solve a real-life problem. As the story progresses, the main character demonstrates superhuman abilities to overcome ordinary obstacles. Imagine that you will participate in a 'tall-tale writing contest' at your school. Write you own tall tale. You can write about yourself, someone you know, or someone you imagine. Be sure to give your main character whatever superhuman abilities are necessary to save the day.	2002-12W3
12BP	<u>Who are the heroes of today?</u> Who are our heroes? The media attention given to celebrities suggests that these people are today's heroes. Yet ordinary people perform extraordinary acts of courage every day that go virtually unnoticed. Are these people the real heroes? Write an essay in which you define heroism and argue who you think our heroes really are - ordinary people or mass-media stars, or maybe both. Be sure to use examples of specific celebrities, other people you have heard or read about, or people from your own community to support your position.	2002-12W20

* The wordings of the prompt are altered slightly from the NAEP wordings.

Appendix 4

Writing Prompts for Hong Kong Students

Prompt	Title/Text of the Prompt	Year/Block of NAEP
4BI#	My Lunch Time Describe what lunch time is like for you on a school day. Be sure to tell about your lunch time so that someone who has never had lunch with you on a school day can understand where you have lunch and what lunch time is like.	2002-4W17
4BN	The Very Unusual Day One day you wake up and go down to breakfast. You eat what you normally eat. Your breakfast is the last normal thing that happens to you all day. Write a story called 'The Very Unusual Day' about what happens that day, from right after breakfast until you go to bed again.	2002-4W5
4BP#	My Favourite Book Your favourite book is missing from your school library. It might be a book that you like to read over and over again. Or it might be a book that your teacher or parent has read to you. Some of your friends also like to read this book. The school librarian is not sure she wants to buy the book again. Write a letter to convince your school librarian to buy the book again. In your letter, give lots of reasons why the book should be in your school library.	2002-4W21

The wordings of the prompt are altered slightly from the wordings of prompts for the students in the United States.

Appendix 5

Holistic Scoring Rubrics

Grade 4 Informative Prompt

Score	Description
(6) Excellent	• Develops ideas well and uses specific, relevant details across the response. • Well organized with clear transitions. • Sustains varied sentence structure and exhibits specific word choices. • Exhibits control over sentence boundaries; errors in grammar, spelling, and mechanics do not interfere with understanding.
(5) Skillful	• Develops ideas with some specific, relevant details. • Clearly organized; information is presented in an orderly way, but response may lack transitions. • Exhibits some variety in sentence structure and exhibits some specific word choices. • Generally exhibits control over sentence boundaries; errors in grammar, spelling, and mechanics do not interfere with understanding.
(4) Sufficient	• Clear but sparsely developed; may have a few details. • Provides a clear sequence of information; provides pieces of information that are generally related to each other. • Generally has simple sentences and simple word choice; may exhibit uneven control over sentence boundaries. • Has sentences that consist mostly of complete, clear, distinct thoughts; errors in grammar, spelling, and mechanics generally do not interfere with understanding.
(3) Uneven	May be characterized by one or more of the following: • Provides limited or incomplete information; may be list-like or have the quality of an outline. • Disorganized or provides a disjointed sequence of information. • Exhibits uneven control over sentence boundaries and may have some inaccurate word choices. • Errors in grammar, spelling, and mechanics sometimes interfere with understanding.
(2) Insufficient	May be characterized by one or more of the following: • Provides little information and makes little attempt at development. • Very disorganized OR too brief for reader to detect organization. • Exhibits little control over sentence boundaries and sentence formation; word choice is inaccurate in much of the response. • Characterized by misspellings, missing words, incorrect word order; errors in grammar, spelling, and mechanics are severe enough to make understanding very difficult in much of the response.
(1) Unsatisfactory	May be characterized by one or more of the following: • Attempts a response, but may only paraphrase the prompt or be extremely brief. • Exhibits no control over organization. • Exhibits no control over sentence formation; word choice is inaccurate across the response. • Characterized by misspellings, missing words, incorrect word order; errors in grammar, spelling, and mechanics severely impede understanding across the response.

Grade 8 Informative Prompt

Score	Description
(6) Excellent	● Develops and shapes information with well-chosen details across the response. ● Well organized with strong transitions. ● Sustains variety in sentence structure and exhibits good word choice. ● Errors in grammar, spelling, and punctuation are few and do not interfere with understanding.
(5) Skillful	● Develops and shapes information with details in parts of the response. ● Clearly organized, but may lack some transitions and/or have occasional lapses in continuity. ● Exhibits some variety in sentence structure and some good word choices. ● Errors in grammar, spelling, and punctuation do not interfere with understanding.
(4) Sufficient	● Develops information with some details. ● Organized with ideas that are generally related, but has few or no transitions. ● Exhibits control over sentence boundaries and sentence structure, but sentences and word choice may be simple and unvaried. ● Errors in grammar, spelling, and punctuation do not interfere with understanding.
(3) Uneven	May be characterized by one or more of the following: ● Presents some clear information, but is list-like, undeveloped, or repetitive OR offers no more than a well-written beginning. ● Unevenly organized; the response may be disjointed. ● Exhibits uneven control over sentence boundaries and sentence structure; may have some inaccurate word choices. ● Errors in grammar, spelling, and punctuation sometimes interfere with the understanding.
(2) Insufficient	May be characterized by one or more of the following: ● Presents fragmented information OR may be very repetitive OR may be very undeveloped. ● Very disorganized; thoughts are tenuously connected OR the response is too brief to detect organization. ● Minimal control over sentence boundaries and sentence structure; word choice may often be inaccurate. ● Errors in grammar or usage (such as missing words or incorrect word use or word order), spelling, and punctuation interfere with understanding in much of the response.
(1) Unsatisfactory	May be characterized by one or more of the following: ● Attempts to respond to prompt, but provides little or no coherent information; may only paraphrase the prompt. ● Has no apparent organization OR consists of a single statement. ● Minimal or no control over sentence boundaries and sentence structure; word choice may be inaccurate in much or all of the response. ● A multiplicity of errors in grammar or usage (such as missing words or incorrect word use or word order), spelling, and punctuation severely impedes understanding across the response.

Grade 12 Informative Prompt

Score	Description
(6) Excellent	● Information is presented effectively and consistently supported with well-chosen details. ● Information is focused and well organized, with a sustained controlling idea and effective use of transitions. ● Response consistently exhibits variety in sentence structure and precision in word choice. ● Errors in grammar, spelling, and punctuation are few and do not interfere with understanding.
(5) Skillful	● Information is presented clearly and supported with pertinent details in much of the response. ● Response is well organized, but may lack some transitions. ● Response exhibits some variety in sentence structure and uses good word choice; occasionally, words may be used inaccurately. ● Errors in grammar, spelling, and punctuation do not interfere with understanding.
(4) Sufficient	● Information is presented clearly and supported with some pertinent details. ● Information is generally organized, but has few or no transitions among parts. ● Sentence structure may be simple and unvaried; word choice is mostly accurate. ● Errors in grammar, spelling, and punctuation do not interfere with understanding.
(3) Uneven	May be characterized by one or more of the following: ● Information is presented clearly in parts; other parts are undeveloped or repetitive OR response is no more than a well-written beginning. ● Information is organized in parts of the response; other parts are disjointed and/or lack transitions. ● Exhibits uneven control over sentence boundaries and sentence structure; may exhibit some inaccurate word choices. ● Errors in grammar, spelling, and punctuation sometimes interfere with understanding.
(2) Insufficient	May be characterized by one or more of the following: ● Provides information that is very undeveloped or list-like. ● Much of the response is disorganized or unfocused, OR the response is too brief to detect organization. ● Author has minimal control over sentence boundaries and sentence structure; word choice may often be inaccurate. ● Errors in grammar, spelling, and punctuation interfere with understanding in much of the response.
(1) Unsatisfactory	May be characterized by one or more of the following: ● Responds to prompt, but may be incoherent OR provides very minimal information OR merely paraphrases the prompt. ● Exhibits little or no apparent organization. ● Minimal or no control over sentence boundaries and sentence structure; word choice may be inaccurate in much or all of the response. ● Errors in grammar, spelling, and punctuation severely impede understanding across the response.

Grade 4 Narrative Prompt

Score	Description
(6) Excellent	● Tells a well-developed story with relevant descriptive details across the response. ● Events are well connected and tie the story together with transitions across the response. ● Sustains varied sentence structure and exhibits specific word choices. ● Exhibits control over sentence boundaries; errors in grammar, spelling, and mechanics do not interfere with understanding.
(5) Skillful	● Tells a clear story with some development, including some relevant descriptive details. ● Events are connected in much of the response; may lack some transitions. ● Exhibits some variety in sentence structure and exhibits some specific word choices. ● Generally exhibits control over sentence boundaries; errors in grammar, spelling, and mechanics do not interfere with understanding.
(4) Sufficient	● Tells a clear story with little development; has few details. ● Events are generally related; may contain brief digressions or inconsistencies. ● Generally has simple sentences and simple word choice; may exhibit uneven control over sentence boundaries. ● Has sentences that consist mostly of complete, clear, distinct thoughts; errors in grammar, spelling, and mechanics generally do not interfere with understanding.
(3) Uneven	May be characterized by one or more of the following: ● Attempts to tell a story, but tells only part of a story, gives a plan for a story, or is list-like. ● Lacks a clear progression of events; elements may not fit together or be in sequence. ● Exhibits uneven control over sentence boundaries and may have some inaccurate word choices. ● Errors in grammar, spelling, and mechanics sometimes interfere with understanding.
(2) Insufficient	May be characterized by one or more of the following: ● Attempts a response, but is no more than a fragment or the beginning of a story OR is very repetitive. ● Very disorganized OR too brief to detect organization. ● Exhibits little control over sentence boundaries and sentence formation; word choice is inaccurate in much of the response. ● Characterized by misspellings, missing words, incorrect word order, errors in grammar, spelling, and mechanics are severe enough to make understanding very difficult in much of the response.
(1) Unsatisfactory	May be characterized by one or more of the following: ● Attempts a response but may only paraphrase the prompt or be extremely brief. ● Exhibits no control over organization. ● Exhibits no control over sentence formation; word choice is inaccurate across the response. ● Characterized by misspellings, missing words, incorrect word order; errors in grammar, spelling, and mechanics severely impede understanding across the response.

Grade 8 Narrative Prompt

Score	Description
(6) Excellent	• Tells a clear story that is well developed and shaped with well-chosen details across the response. • The story is well organized with strong transitions. • Sustains variety in sentence structure and exhibits good word choice. • Errors in grammar, spelling, and punctuation are few and do not interfere with understanding.
(5) Skillful	• Tells a clear story that is developed and shaped with details in parts of the response. • The story is clearly organized, but may lack some transitions and/or have occasional lapses in continuity. • Exhibits some variety in sentence structure and some good word choices. • Errors in grammar, spelling, and punctuation do not interfere with understanding.
(4) Sufficient	• Tells a clear story that is developed with some details. • The parts of the story are generally related, but there are few or no transitions. • Exhibits control over sentence boundaries and sentence structure, but sentences and word choice may be simple and unvaried. • Errors in grammar, spelling, and punctuation do not interfere with understanding.
(3) Uneven	May be characterized by one or more of the following: • Attempts to tell a story, but parts of the story are unclear, undeveloped, list-like, or repetitive OR offers no more than a well-written beginning. • Unevenly organized; parts of the story may be unrelated to one another. • Exhibits uneven control over sentence boundaries and sentence structure; may have some inaccurate word choices. • Errors in grammar, spelling, and punctuation sometimes interfere with understanding.
(2) Insufficient	May be characterized by one or more of the following: • Attempts to tell a story, but the attempt may be a fragment and/or very undeveloped. • Very disorganized throughout the response OR too brief to detect organization. • Minimal control over sentence boundaries and sentence structure; word choice may often be inaccurate. • Errors in grammar or usage (such as missing words or incorrect word use or word order), spelling, and punctuation interfere with understanding in much of the response.
(1) Unsatisfactory	May be characterized by one or more of the following: • Responds to prompt, but provides little or no coherent content OR merely paraphrases the prompt. • Has no apparent organization OR consists of a single statement. • Minimal or no control over sentence boundaries and sentence structure; word choice may be inaccurate in much or all of the response. • A multiplicity of errors in grammar or usage (such as missing words or incorrect word use or word order), spelling, and punctuation severely impedes understanding across the response.

Grade 12 Narrative Prompt

Score	Description
(6) Excellent	● Tells a clear story that is consistently well developed and detailed; details enhance story being told. ● Well organized; integrates narrative events into a smooth telling; effective transitions move the story forward. ● Consistently exhibits variety in sentence structure and precision in word choice. ● Errors in grammar, spelling, and punctuation are few and do not interfere with understanding.
(5) Skillful	● Tells a clear story that is well developed and elaborated with details in much of the response. ● Well organized with story elements that are connected across most of the response; may have occasional lapse in transitions. ● Exhibits some variety in sentence structure and uses good word choice; occasionally, words may be used inaccurately. ● Errors in grammar, spelling, and punctuation do not interfere with understanding
(4) Sufficient	● Tells a clear story that is developed with some pertinent details. ● Generally organized, but transitions among parts of the story may be lacking. ● Sentence structure may be simple and unvaried; word choice is mostly accurate. ● Errors in grammar, spelling, and punctuation do not interfere with understanding.
(3) Uneven	May be characterized by one or more of the following: ● Tells a story that may be clear and developed in parts; other parts are unfocused, repetitive, or minimally developed OR response is no more that a well-written beginning. ● Organized in parts of the response; other parts are disjointed and/or lack transitions. ● Exhibits uneven control over sentence boundaries and sentence structure; may exhibit some inaccurate word choices. ● Errors in grammar, spelling, and punctuation sometimes interfere with understanding.
(2) Insufficient	May be characterized by one or more of the following: ● Attempts to tell a story, but is very undeveloped, list-like, or fragmentary. ● Disorganized or unfocused in much of the response OR the response is too brief to detect organization. ● Minimal control over sentence boundaries and sentence structure; word choice may often be inaccurate. ● Errors in grammar, spelling, and punctuation interfere with understanding in much of the response.
(1) Unsatisfactory	May be characterized by one or more of the following: ● Responds to prompt but provides little or no coherent content OR merely paraphrases the prompt. ● Has little or no apparent organization. ● Minimal or no control over sentence boundaries and sentence structure; word choice may be inaccurate in much or all of the response. ● Errors in grammar, spelling, and punctuation severely impede understanding across the response.

Grade 4 Persuasive Prompt

Score	Description
(6) Excellent	• Takes a clear position and develops support with well-chosen details, reasons, or examples across the response. • Well organized; maintains focus. • Sustains varied sentence structure and exhibits specific word choices. • Exhibits control over sentence boundaries; errors in grammar, spelling, and mechanics do not interfere with understanding.
(5) Skillful	• Takes a clear position and develops support with some specific details, reasons, or examples. • Provides some organization of ideas by, for example, using contrast or building to a point. • Exhibits some variety in sentence structure and exhibits some specific word choices. • Generally exhibits control over sentence boundaries; errors in grammar, spelling, and mechanics do not interfere with understanding.
(4) Sufficient	• Takes a clear position with support that is clear and generally related to the issue. • Generally organized. • Generally has simple sentences and simple word choice; may exhibit uneven control over sentence boundaries. • Has sentences that consist mostly of complete, clear, distinct thoughts; errors in grammar, spelling, and mechanics generally do not interfere with understanding.
(3) Uneven	May be characterized by one or more of the following: • Takes a position and offers limited or incomplete support; some reasons may not be clear or related to the issue. • Disorganized OR provides a disjointed sequence of information. • Exhibits uneven control over sentence boundaries and may have some inaccurate word choices. • Errors in grammar, spelling, and mechanics sometimes interfere with understanding.
(2) Insufficient	May be characterized by one or more of the following: • Takes a position, but provides only minimal support (generalizations or a specific reason or example); OR attempts to take a position but the position is unclear. • Very disorganized or too brief to detect organization. • May exhibit little control over sentence boundaries and sentence formation; word choice is inaccurate in much of the response. • Characterized by misspellings, missing words, incorrect word order; errors in grammar, spelling, and mechanics may be severe enough to make understanding very difficult in much of the response.
(1) Unsatisfactory	May be characterized by one or more of the following: • Takes a position but provides no support, OR attempts to take a position (is on topic) but the position is very unclear; may only paraphrase the prompt. • Exhibits no control over organization. • Exhibits no control over sentence formation; word choice is inaccurate across the response. • Characterized by misspellings, missing words, incorrect word order; errors in grammar, spelling, and mechanics severely impede understanding across the response.

Grade 8 Persuasive Prompt

Score	Description
(6) Excellent	● Takes a clear position and develops it consistently with well-chosen reasons and/or examples across the response. ● Well organized with strong transitions. ● Sustains variety in sentence structure and exhibits good word choice. ● Errors in grammar, spelling, and punctuation are few and do not interfere with understanding.
(5) Skillful	● Takes a clear position and develops it with reasons and/or examples in parts of the response. ● Clearly organized, but may lack some transitions and/or have occasional lapses in continuity. ● Exhibits some variety in sentence structure and some good word choices. ● Errors in grammar, spelling, and punctuation do not interfere with understanding.
(4) Sufficient	● Takes a clear position and supports it with some reasons and/or examples. ● Organized with ideas that are generally related, but there are few or no transitions. ● Exhibits control over sentence boundaries and sentence structure, but sentences and word choice may be simple and unvaried. ● Errors in grammar, spelling, and punctuation do not interfere with understanding.
(3) Uneven	May be characterized by one or more of the following: ● Takes a position and offers support, but may be unclear, repetitive, list-like, or undeveloped. ● Unevenly organized; the response may be disjointed. ● Exhibits uneven control over sentence boundaries and sentence structure; may have some inaccurate word choices. ● Errors in grammar, spelling, and punctuation sometimes interfere with understanding.
(2) Insufficient	May be characterized by one or more of the following: ● Takes a position, but may be very unclear, very undeveloped, or very repetitive. ● Very disorganized; thoughts are tenuously connected OR the response is too brief to detect organization. ● Minimal control over sentence boundaries and sentence structure; word choice may often be inaccurate. ● Errors in grammar or usage (such as missing words or incorrect word use or word order), spelling, and punctuation interfere with understanding in much of the response.
(1) Unsatisfactory	May be characterized by one or more of the following: ● Attempts to take a position (addresses topic) but response is incoherent OR takes a position but provides no support; may only paraphrase the prompt. ● Has no apparent organization OR consists of a single statement. ● Minimal or no control over sentence boundaries and sentence structure; word choice may be inaccurate in much or all of the response. ● A multiplicity of errors in grammar or usage (such as missing words or incorrect word use or word order), spelling, and punctuation severely impedes understanding across the response.

Grade 12 Persuasive Prompt

Score	Description
(6) Excellent	● Takes a clear position and supports it consistently with well-chosen reasons and/or examples; may use persuasive strategy to convey an argument. ● Focused and well organized, with effective use of transitions. ● Consistently exhibits variety in sentence structure and precision in word choice. ● Errors in grammar, spelling, and punctuation are few and do not interfere with understanding.
(5) Skillful	● Takes a clear position and supports it with pertinent reasons and/or examples through much of the response. ● Well organized, but may lack some transitions. ● Exhibits some variety in sentence structure and uses good word choice; occasionally, words may be used inaccurately. ● Errors in grammar, spelling, and punctuation do not interfere with understanding.
(4) Sufficient	● Takes a clear position and supports it with some pertinent reasons and/or examples; there is some development. ● Generally organized, but has few or no transitions among parts. ● Sentence structure may be simple and unvaried; word choice is mostly accurate. ● Errors in grammar, spelling, and punctuation do not interfere with understanding.
(3) Uneven	May be characterized by one or more of the following: ● Takes a position and provides uneven support; may lack development in parts or be repetitive OR is no more than a well-written beginning. ● Organized in parts of the response; other parts are disjointed and/or lack transitions. ● Exhibits uneven control over sentence boundaries and sentence structure; may exhibit some inaccurate word choices. ● Errors in grammar, spelling, and punctuation sometimes interfere with understanding.
(2) Insufficient	May be characterized by one or more of the following: ● Takes a position but is very undeveloped. ● Disorganized or unfocused in much of the response OR clear but very brief. ● Minimal control over sentence boundaries and sentence structure; word choice may often be inaccurate. ● Errors in grammar, spelling, and punctuation interfere with understanding in much of the response.
(1) Unsatisfactory	May be characterized by one or more of the following: ● Attempts to take a position (addresses topic) but position is very unclear OR takes a position but provides minimal or no support; may only paraphrase the prompt. ● Little or no apparent organization. ● Minimal or no control over sentence boundaries and sentence structure; word choice may be inaccurate in much or all of the response. ● Errors in grammar, spelling, and punctuation severely impede understanding across the response.

Appendix 6

Administrative documentation
has been removed

Appendix 7

Data Use Agreement

Administrative documentation
has been removed

Appendix 8

Administrative documentation has been removed

Appendix 9

Category Structure for the scores of the Lexile Analyzer of US student essays on the 4AI prompt (50 categories)

SUMMARY OF CATEGORY STRUCTURE. Model="R"

CATEGORY LABEL	SCORE	OBSERVED COUNT	%	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	STRUCTURE CALIBRATN	CATEGORY MEASURE	
1	1	1	0	-.69	-.98	3.43	9.90	NONE	(-1.29)	01
2	2	0	0			.00	.00	NULL	-1.20	
3	3	0	0			.00	.00	NULL	-1.15	
4	4	0	0			.00	.00	NULL	-1.12	
5	5	0	0			.00	.00	NULL	-1.10	
6	6	0	0			.00	.00	NULL	-1.08	
7	7	0	0			.00	.00	NULL	-1.07	
8	8	0	0			.00	.00	NULL	-1.05	
9	9	0	0			.00	.00	NULL	-1.04	
10	10	0	0			.00	.00	NULL	-1.03	
11	11	0	0			.00	.00	NULL	-1.02	
12	12	0	0			.00	.00	NULL	-1.01	
13	13	1	0	-1.02*	-.90	.02	.01	-11.05	-1.00	13
14	14	0	0			.00	.00	NULL	-.99	
15	15	0	0			.00	.00	NULL	-.98	
16	16	0	0			.00	.00	NULL	-.98	
17	17	2	0	-.84	-.86	1.00	.59	-4.15	-.97	17
18	18	2	0	-.81	-.85	1.17	1.11	-.84	-.96	18
19	19	0	0			.00	.00	NULL	-.95	
20	20	0	0			.00	.00	NULL	-.94	
21	21	0	0			.00	.00	NULL	-.93	
22	22	0	0			.00	.00	NULL	-.92	
23	23	1	0	-.55	-.80	2.46	5.01	-3.37	-.91	23
24	24	0	0			.00	.00	NULL	-.90	
25	25	1	0	-.72*	-.77	1.29	.84	-1.54	-.89	25
26	26	2	0	-.58	-.75	1.59	3.31	-1.44	-.88	26
27	27	4	0	-.68*	-.73	1.75	1.86	-1.42	-.86	27
28	28	1	0	-.56	-.70	1.74	2.01	.68	-.85	28
29	29	4	0	-.49	-.68	1.71	2.51	-2.07	-.83	29
30	30	1	0	-.04	-.64	3.13	6.78	.73	-.81	30
31	31	1	0	-.78*	-.61	.00	.00	-.62	-.79	31
32	32	8	0	-.59	-.56	.94	.91	-2.66	-.76	32
33	33	7	0	-.32	-.51	1.35	2.22	-.40	-.73	33
34	34	16	0	-.51*	-.44	.60	.77	-1.30	-.69	34
35	35	24	1	-.36	-.35	1.21	.98	-.80	-.63	35
36	36	47	1	-.23	-.24	.99	1.03	-.97	-.55	36
37	37	52	2	-.04	-.11	1.06	1.14	-.28	-.43	37
38	38	87	3	.05	.05	1.54	1.09	-.55	-.27	38
39	39	110	3	.22	.24	1.83	1.08	-.09	-.07	39
40	40	199	6	.43	.46	.96	.99	-.24	.18	40
41	41	303	9	.72	.72	.94	.96	.17	.49	41
42	42	387	11	.99	1.03	.99	.94	.62	.89	42
43	43	441	13	1.42	1.41	.92	.94	1.08	1.39	43
44	44	510	15	1.91	1.90	.92	.93	1.50	1.98	44
45	45	465	13	2.48	2.51	1.10	.98	2.29	2.69	45
46	46	386	11	3.18	3.17	1.06	.92	3.02	3.53	46
47	47	259	8	3.80	3.81	.95	.96	3.89	4.53	47
48	48	105	3	4.56	4.48	.86	.87	5.03	5.80	48
49	49	23	1	5.77	5.50	.82	.85	6.44	7.48	49
50	50	3	0	6.21	7.16	1.99	1.70	8.35	(9.55)	50
MISSING		7149	67	1.83						

Category Structure for the scores of the Lexile Analyzer of US student essays on the 4AI prompt (23 categories)

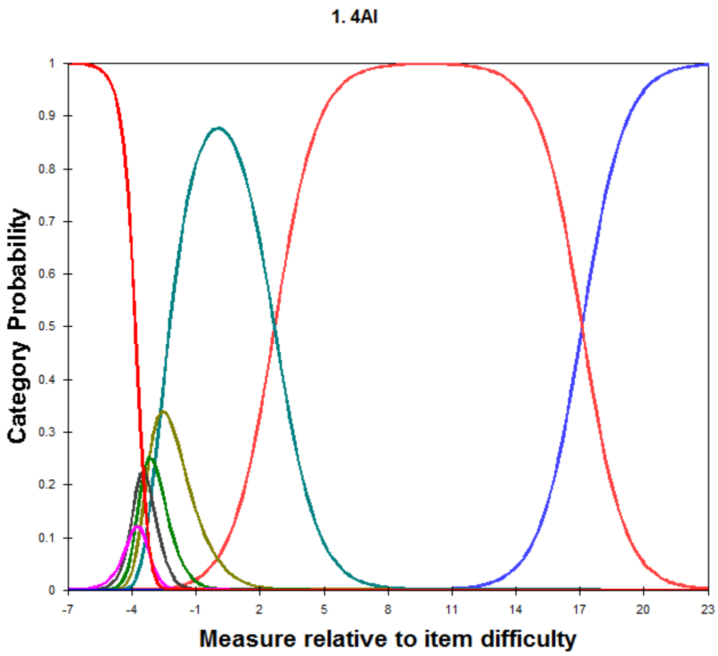
SUMMARY OF CATEGORY STRUCTURE. Model="R"

CATEGORY LABEL	SCORE	OBSERVED COUNT	%	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	STRUCTURE CALIBRATN	CATEGORY MEASURE	
1	1	1	0	-1.20	-1.85	3.51	9.90	NONE	(-2.43)	01
2	2	0	0			.00	.00	NULL	-2.20	
3	3	0	0			.00	.00	NULL	-2.08	
4	4	0	0			.00	.00	NULL	-2.01	
5	5	0	0			.00	.00	NULL	-1.95	
6	6	1	0	-1.94*	-1.67	.02	.01	-8.59	-1.91	06
7	7	0	0			.00	.00	NULL	-1.86	
8	8	2	0	-1.56	-1.58	.92	.54	-3.88	-1.82	08
9	9	2	0	-1.46	-1.53	1.12	1.07	-1.53	-1.78	09
10	10	0	0			.00	.00	NULL	-1.73	
11	11	1	0	-.95	-1.42	2.29	3.88	-2.21	-1.68	11
12	12	3	0	-1.10*	-1.34	1.44	2.21	-2.46	-1.62	12
13	13	8	0	-1.07	-1.23	1.67	1.80	-2.25	-1.56	13
14	14	3	0	-.66	-1.09	1.61	2.78	-.17	-1.47	14
15	15	14	0	-.77*	-.89	1.13	1.43	-2.53	-1.36	15
16	16	43	1	-.61	-.58	.94	.95	-1.86	-1.16	16
17	17	110	3	-.03	-.07	.98	.99	-1.28	-.77	17
18	18	261	8	.64	.69	1.49	1.02	-.58	-.01	18
19	19	678	20	1.72	1.76	.95	.91	.24	1.27	19
20	20	1061	31	3.41	3.39	.92	.95	2.07	3.40	20
21	21	934	27	5.76	5.77	1.08	.97	4.66	6.35	21
22	22	317	9	8.07	8.03	.96	.96	8.02	10.20	22
23	23	14	0	10.89	10.83	.94	.95	12.36	(13.47)	23
MISSING		7149	67	3.82						

Category Structure for the scores of the Lexile Analyzer of US student essays on the 4AI prompt (9 categories)

SUMMARY OF CATEGORY STRUCTURE. Model="R"

CATEGORY LABEL	SCORE	OBSERVED COUNT	%	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	STRUCTURE CALIBRATN	CATEGORY MEASURE	
1	1	1	0	-2.19	-3.31	2.19	5.98	NONE	(-4.60)	1
2	2	0	0			.00	.00	NULL	-4.07	
3	3	1	0	-3.63*	-2.98	.03	.01	-6.26	-3.75	3
4	4	4	0	-2.79	-2.72	.90	.55	-4.22	-3.48	4
5	5	10	0	-2.13	-2.30	1.20	1.30	-3.44	-3.15	5
6	6	38	1	-1.29	-1.20	.92	.86	-3.18	-2.55	6
7	7	606	18	1.63	1.63	1.02	.55	-2.67	.04	7
8	8	2547	74	9.22	9.22	1.00	1.03	2.68	9.88	8
9	9	246	7	17.24	17.18	.91	.40	17.09	(18.19)	9
MISSING		7113	67	8.27						



Category Probability Curves for the scores of the Lexile Analyzer of US student essays on the 4AI prompt (9 categories)

Appendix 10

Prompt polarity for the scores of Lexile Analyzer of US student essays

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Prompt
17	722	98	-.88	.11	1.57	3.4	1.55	3.1	.66	.75	33.7	43.0	12BN
1	457	139	.20	.09	1.11	.9	1.09	.8	.69	.75	43.1	39.7	4AI
18	657	93	-.49	.11	1.36	2.3	1.38	2.3	.70	.75	38.7	40.9	12BP
3	410	133	.42	.09	.95	-.4	.97	-.2	.74	.74	32.1	40.1	4AP
2	512	134	-.37	.09	1.08	.6	1.08	.7	.74	.77	37.9	39.2	4AN
4	948	264	.40	.06	1.16	1.8	1.17	1.9	.76	.78	33.0	39.5	4BI
14	1190	188	.21	.07	1.02	.3	1.04	.4	.76	.76	38.0	38.7	12AN
6	829	264	.89	.07	.97	-.3	.95	-.5	.77	.77	45.6	41.0	4BP
13	1286	188	-.32	.08	.88	-1.2	.87	-1.3	.78	.74	39.0	39.5	12AI
15	1161	189	.41	.07	.78	-2.3	.79	-2.2	.80	.77	45.7	38.9	12AP
9	1169	250	.18	.06	.85	-1.7	.85	-1.7	.81	.81	41.6	38.7	8AP
5	1116	266	-.24	.06	1.11	1.3	1.11	1.2	.82	.80	35.6	38.9	4BN
7	1174	246	.03	.06	.88	-1.3	.88	-1.4	.82	.80	38.2	38.7	8AI
16	612	97	.30	.10	.61	-3.2	.60	-3.2	.83	.78	39.2	39.5	12BI
8	1410	254	-.75	.06	1.05	.6	1.04	.4	.84	.81	36.6	39.0	8AN
12	1183	216	.24	.07	.92	-.8	.91	-.9	.84	.83	39.5	38.7	8BP
10	1174	218	.35	.07	.85	-1.6	.87	-1.4	.84	.82	44.2	38.6	8BI
11	1357	216	-.60	.07	.90	-1.1	.89	-1.2	.85	.82	43.7	39.8	8BN
MEAN	964.8	191.8	.00	.08	1.00	-.2	1.00	-.2			39.2	39.6	
S.D.	317.3	60.2	.47	.02	.21	1.7	.21	1.6			4.0	1.1	

Prompt polarity for the scores of IEA of US student essays

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Prompt
17	669	98	-.55	.10	1.61	3.7	1.54	3.3	.61	.76	30.6	37.7	12BN
16	579	97	.24	.10	.95	-.3	.95	-.3	.70	.78	33.0	36.0	12BI
18	602	93	-.22	.10	1.28	1.8	1.25	1.7	.71	.76	38.7	36.3	12BP
1	427	139	.40	.09	1.06	.5	1.09	.8	.73	.74	45.6	41.6	4AI
14	1161	188	.26	.07	1.16	1.6	1.15	1.4	.75	.76	37.4	35.9	12AN
13	1350	188	-.70	.08	1.33	2.9	1.22	2.0	.75	.72	30.5	39.3	12AI
3	415	133	.35	.09	.82	-1.5	.94	-.4	.76	.75	45.0	40.9	4AP
6	940	264	.38	.06	.79	-2.5	.81	-2.2	.79	.76	44.2	39.4	4BP
4	966	264	.28	.06	.94	-.7	.92	-.9	.79	.77	41.9	39.1	4BI
11	1289	216	-.07	.07	1.01	.1	1.02	.3	.79	.82	37.9	36.7	8BN
15	1140	189	.39	.07	1.07	.7	1.04	.4	.80	.76	35.1	35.8	12AP
8	1381	254	-.49	.06	.94	-.6	.96	-.4	.80	.80	39.5	36.7	8AN
2	542	134	-.65	.09	.96	-.3	.98	-.1	.81	.80	37.4	36.9	4AN
7	1171	246	.09	.06	.90	-1.1	.91	-1.0	.81	.79	42.0	36.7	8AI
9	1188	250	.13	.06	.89	-1.3	.89	-1.3	.81	.80	41.4	36.7	8AP
5	1184	266	-.51	.06	.93	-.8	.94	-.7	.82	.80	38.2	37.2	4BN
10	1169	218	.49	.06	.96	-.4	.95	-.5	.83	.82	40.3	36.2	8BI
12	1230	216	.19	.07	.78	-2.5	.79	-2.4	.86	.81	41.1	36.4	8BP
MEAN	966.8	191.8	.00	.07	1.02	.0	1.02	.0			38.9	37.5	
S.D.	323.8	60.2	.40	.01	.20	1.6	.18	1.4			4.3	1.7	

Category Structure for the scores of the Lexile Analyzer of US student essays on the 4AI prompt (new 9 categories)

SUMMARY OF CATEGORY STRUCTURE. Model="R"

CATEGORY LABEL	SCORE	OBSERVED COUNT	%	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	STRUCTURE CALIBRATN	CATEGORY MEASURE	
1	1	163	5	-3.32	-3.35	1.01	1.01	NONE	(-4.70)	1
2	2	255	7	-2.34	-2.39	1.13	1.15	-3.38	-3.09	2
3	3	404	12	-1.60	-1.57	.98	.97	-2.43	-1.99	3
4	4	570	17	-.83	-.80	.98	.98	-1.53	-1.00	4
5	5	643	19	.02	.00	.98	.98	-.53	.01	5
6	6	552	16	.81	.83	.97	.96	.57	1.01	6
7	7	423	12	1.64	1.64	.93	.91	1.50	1.99	7
8	8	273	8	2.42	2.41	.98	1.01	2.46	3.09	8
9	9	170	5	3.35	3.30	.89	.91	3.33	(4.67)	9
MISSING		7097	67	.08						

Category Structure for the scores of IEA of US student essays on the 4AI prompt (new 9 categories)

SUMMARY OF CATEGORY STRUCTURE. Model="R"

CATEGORY LABEL	SCORE	OBSERVED COUNT	%	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	STRUCTURE CALIBRATN	CATEGORY MEASURE	
1	1	162	5	-3.33	-3.40	1.09	1.07	NONE	(-4.65)	1
2	2	257	7	-2.21	-2.15	1.02	1.01	-3.36	-2.93	2
3	3	441	13	-1.23	-1.24	1.01	1.03	-2.21	-1.71	3
4	4	499	14	-.54	-.55	.88	.87	-1.01	-.78	4
5	5	670	19	.06	.09	.97	.98	-.52	.08	5
6	6	539	16	.81	.75	.89	.89	.64	.93	6
7	7	416	12	1.36	1.41	1.15	1.14	1.34	1.75	7
8	8	309	9	2.09	2.04	.93	.94	2.02	2.78	8
9	9	160	5	2.70	2.73	1.03	1.03	3.11	(4.41)	9
MISSING		7075	67	.07						

Coherence of categories for the scores of the Lexile Analyzer of US student essays on the 4AI prompt (new 9 categories)

CATEGORY LABEL	STRUCTURE MEASURE	S.E.	SCORE-TO-MEASURE AT CAT.	----ZONE----		50% CUM. PROBABLTY	COHERENCE M->C	C->M	RMSR	ESTIM DISCR	
1	NONE		(-4.70)	-INF	-3.94		69%	26%	1.1973		1
2	-3.38	.11	-3.09	-3.94	-2.50	-3.66	35%	35%	1.1319	.98	2
3	-2.43	.07	-1.99	-2.50	-1.50	-2.46	34%	36%	.9428	.91	3
4	-1.53	.06	-1.00	-1.50	-.50	-1.51	39%	43%	.8904	.98	4
5	-.53	.05	.01	-.50	.52	-.51	40%	45%	.8974	1.06	5
6	.57	.05	1.01	.52	1.50	.53	37%	38%	.9137	.99	6
7	1.50	.06	1.99	1.50	2.50	1.50	40%	43%	.9064	1.07	7
8	2.46	.07	3.09	2.50	3.92	2.47	41%	45%	1.0067	1.10	8
9	3.33	.10	(4.67)	3.92	+INF	3.63	72%	24%	1.1375	1.12	9

Coherence of categories for the scores of IEA of US student essays on the 4AI prompt (new 9 categories)

CATEGORY LABEL	STRUCTURE MEASURE	S.E.	SCORE-TO-MEASURE AT CAT.	----ZONE----		50% CUM. PROBABLTY	COHERENCE M->C	C->M	RMSR	ESTIM DISCR	
1	NONE		(-4.65)	-INF	-3.85		70%	34%	1.1565		1
2	-3.36	.11	-2.93	-3.85	-2.27	-3.60	46%	35%	1.1355	.93	2
3	-2.21	.08	-1.71	-2.27	-1.23	-2.23	41%	40%	1.0447	1.05	3
4	-1.01	.06	-.78	-1.23	-.35	-1.16	34%	42%	.8930	1.02	4
5	-.52	.05	.08	-.35	.51	-.39	40%	44%	.9166	.99	5
6	.64	.05	.93	.51	1.34	.54	38%	42%	.9231	1.12	6
7	1.34	.06	1.75	1.34	2.22	1.33	35%	39%	1.0855	.94	7
8	2.02	.07	2.78	2.22	3.64	2.14	42%	40%	1.0686	1.09	8
9	3.11	.10	(4.41)	3.64	+INF	3.37	65%	11%	1.3862	.97	9

Appendix 11

```
+--------------------------------------------------------------------------------------------------------------+
|Measr|-Prompt                    |-US Rater                                                 |+Students   |R6SCA|
|-----+---------------------------+----------------------------------------------------------+------------+-----|
|   6 +                           +                                                          + .          + (6) |
|     |                           |                                                          |            |     |
|     |                           |                                                          |            |     |
|   5 +                           +                                                          + *          +     |
|     |                           |                                                          | .          |     |
|     |                           |                                                          | *.         | --- |
|   4 +                           +                                                          + *.         +     |
|     |                           |                                                          | **.        |     |
|     |                           |                                                          | ***.       |  5  |
|   3 +                           +                                                          + ***.       +     |
|     |                           |                                                          | ****.      |     |
|     |                           |                                                          | *****.     | --- |
|   2 + 15-12AP                   +                                                          + *****.     +     |
|     | 16-12BI                   |                                                          | ****       |     |
|     | 9-8AP                     |                                                          | *****.     |     |
|   1 + 13-12AI  14-12AN  11-8BN  + US200                                                    + ******.    +  4  |
|     | 17-12BN  18-12BP  12-8BP  | US209  US211                                             | *******.   |     |
|     | 10-8BI                    | US103  US104  US206                                      | *********. |     |
*   0 * 7-8AI    8-8AN            * US102  US105  US201  US203  US204  US205  US208  US210   * *******    * --- *
|     |                           | US101  US202  US207                                      | ********.  |     |
|     |                           | US106                                                    | ******     |     |
|  -1 +                           +                                                          + ******.    +  3  |
|     | 5-4BN    6-4BP            | US108                                                    | *****      |     |
|     | 4-4BI                     |                                                          | *****.     |     |
|  -2 + 1-4AI    2-4AN    3-4AP   +                                                          + ****.      + --- |
|     |                           |                                                          | **.        |     |
|     |                           |                                                          | ***.       |     |
|  -3 +                           +                                                          + *.         +     |
|     |                           |                                                          | *.         |  2  |
|     |                           |                                                          | .          |     |
|  -4 +                           +                                                          + *.         +     |
|     |                           |                                                          | .          | --- |
|     |                           |                                                          | .          |     |
|  -5 +                           +                                                          +            +     |
|     |                           |                                                          | .          |     |
|     |                           |                                                          | .          |     |
|  -6 +                           +                                                          + .          +     |
|     |                           |                                                          |            |     |
|     |                           |                                                          |            |     |
|  -7 +                           +                                                          +            +     |
|     |                           |                                                          |            |     |
|     |                           |                                                          | .          |     |
|  -8 +                           +                                                          +            + (1) |
|-----+---------------------------+----------------------------------------------------------+------------+-----|
|Measr|-Prompt                    |-US Rater                                                 | * = 5      |R6SCA|
+--------------------------------------------------------------------------------------------------------------+
```

Vertical map for the US student essays with scoring of US raters

Calibration of the students for US student essays using US raters

	Student		Observed	Measure	S.E.	Infit		Outfit	
	Number	Group	score	(Logit)		MnSq	ZStd	MnSq	ZStd
1	US001	104	3.7	-1.13	0.30	0.66	-1.1	0.66	-1.2
2	US002	104	3.8	-0.81	0.33	1.31	0.9	1.32	0.9
3	US003	104	4.1	-0.31	0.30	0.92	-0.1	0.92	-0.1
4	US004	104	4.2	-0.14	0.30	0.66	-1.2	0.66	-1.1
5	US005	104	3.1	-2.40	0.30	0.99	0	0.99	0
6	US006	104	4.1	-0.31	0.30	0.32	-3.0	0.32	-3.0
8	US007	104	3.3	-2.13	0.30	0.70	-1.0	0.70	-1.0
9	US008	104	3.2	-2.22	0.30	0.40	-2.6	0.40	-2.6
10	US009	104	4.1	-0.31	0.30	1.25	0.8	1.25	0.8
11	US010	104	3.5	-1.50	0.30	0.77	-0.7	0.77	-0.7
12	US011	104	5.3	2.08	0.31	0.64	-1.5	0.62	-1.6

13	US012	104	1.7	-5.71	0.41	0.64	-1.1	0.68	-1.0
14	US013	104	4.3	0.21	0.29	0.99	0	0.99	0
16	US014	104	4.6	0.72	0.29	1.05	0.2	1.07	0.3
17	US015	104	4.1	-0.46	0.33	0.45	-1.9	0.45	-1.9
18	US016	104	3.4	-1.86	0.30	0.86	-0.4	0.86	-0.4
19	US017	104	1.6	-6.00	0.39	1.58	1.8	1.48	1.5
21	US018	104	3.8	-1.04	0.30	1.65	1.9	1.65	1.9
22	US019	104	3.7	-1.23	0.30	0.42	-2.4	0.42	-2.4
23	US020	104	4.2	-0.04	0.30	0.71	-0.9	0.72	-0.9
24	US021	104	3.3	-2.03	0.30	1.42	1.3	1.41	1.3
26	US022	106	2.6	-2.51	0.30	0.57	-1.7	0.57	-1.7
28	US023	106	4.0	0.48	0.30	1.07	0.3	1.07	0.3
29	US024	106	3.1	-1.41	0.30	0.45	-2.3	0.45	-2.3
30	US025	106	2.8	-1.96	0.30	2.06	2.9	2.05	2.9
31	US026	106	3.4	-0.69	0.30	0.65	-1.2	0.65	-1.2
33	US027	106	4.0	0.48	0.30	1.18	0.6	1.20	0.7
35	US028	106	3.3	-0.96	0.30	0.65	-1.2	0.65	-1.2
36	US029	106	3.3	-1.05	0.30	1.04	0.2	1.03	0.1
37	US030	106	3.2	-1.14	0.30	1.21	0.7	1.21	0.7
39	US031	106	3.5	-0.51	0.30	1.16	0.6	1.16	0.6
40	US032	106	3.1	-1.41	0.30	0.81	-0.6	0.81	-0.6
41	US033	106	3.1	-1.32	0.30	0.86	-0.4	0.86	-0.4
42	US034	106	3.0	-1.69	0.30	0.69	-1.1	0.69	-1.1
44	US035	106	4.3	1.09	0.29	1.66	2.0	1.72	2.2
45	US036	106	3.7	-0.06	0.30	1.30	1.0	1.30	1.0
46	US037	106	3.8	0.12	0.30	0.71	-1.0	0.72	-0.9
47	US038	106	3.3	-0.96	0.30	0.54	-1.7	0.54	-1.7
48	US039	106	2.5	-2.58	0.30	1.07	0.3	1.06	0.2
49	US040	106	3.8	0.23	0.30	1.23	0.8	1.23	0.8
50	US041	104	3.8	-0.91	0.30	1.03	0.2	1.03	0.2
51	US042	104	2.6	-3.52	0.30	0.40	-2.7	0.40	-2.7
52	US043	104	3.6	-1.37	0.30	0.89	-0.3	0.89	-0.3
53	US044	104	3.8	-0.91	0.30	1.26	0.9	1.26	0.9
54	US045	104	4.0	-0.36	0.30	0.53	-1.8	0.53	-1.7
55	US046	104	2.7	-3.26	0.30	0.42	-2.5	0.42	-2.5
57	US047	104	3.6	-1.37	0.30	0.86	-0.3	0.86	-0.3
58	US048	104	3.4	-1.88	0.37	0.70	-0.8	0.70	-0.8
59	US049	104	2.6	-3.43	0.30	1.95	2.7	1.95	2.7

61	US050	104	3.0	-2.65	0.33	0.95	0	0.95	0
62	US051	104	4.5	0.68	0.29	0.86	-0.4	0.86	-0.4
63	US052	104	3.2	-2.27	0.30	0.94	-0.1	0.94	-0.1
64	US053	104	3.8	-0.93	0.33	0.46	-2.0	0.45	-2.0
65	US054	104	3.6	-1.28	0.30	1.25	0.8	1.25	0.8
66	US055	104	3.0	-2.72	0.30	1.02	0.1	1.02	0.1
67	US056	104	2.9	-2.90	0.30	0.93	-0.1	0.93	-0.1
68	US057	104	3.8	-1.01	0.37	1.50	1.3	1.50	1.3
70	US058	104	2.5	-3.87	0.30	2.23	3.4	2.24	3.4
71	US059	104	3.3	-2.17	0.33	0.40	-2.3	0.40	-2.3
72	US060	104	2.4	-4.06	0.30	1.40	1.3	1.41	1.3
74	US061	104	2.4	-3.92	0.37	0.76	-0.6	0.76	-0.6
75	US062	106	3.7	-0.17	0.30	1.20	0.7	1.20	0.7
76	US063	106	3.9	0.69	0.37	0.56	-1.3	0.54	-1.4
77	US064	106	3.7	-0.08	0.30	0.54	-1.7	0.54	-1.7
79	US065	106	3.7	-0.17	0.30	0.77	-0.7	0.78	-0.7
80	US066	106	4.5	1.49	0.29	0.74	-0.9	0.74	-0.9
81	US067	106	3.9	0.37	0.30	0.69	-1.0	0.69	-1.0
82	US068	106	3.8	0.19	0.30	0.86	-0.4	0.86	-0.4
83	US069	106	3.5	-0.54	0.30	0.64	-1.3	0.64	-1.2
84	US070	106	2.4	-3.00	0.31	0.57	-1.7	0.58	-1.7
86	US071	106	3.8	0.01	0.30	2.56	3.8	2.53	3.7
87	US072	106	3.0	-1.62	0.30	1.11	0.4	1.11	0.4
88	US073	106	2.7	-2.35	0.30	1.91	2.6	1.90	2.6
89	US074	106	3.6	-0.26	0.30	1.03	0.2	1.04	0.2
90	US075	106	4.3	1.12	0.29	1.46	1.5	1.46	1.5
91	US076	106	4.0	0.41	0.30	0.40	-2.5	0.40	-2.6
92	US077	106	2.5	-2.83	0.33	0.82	-0.5	0.82	-0.5
93	US078	106	4.2	0.84	0.29	1.08	0.3	1.06	0.3
94	US079	106	4.1	0.76	0.29	0.58	-1.6	0.58	-1.6
95	US080	106	2.5	-2.67	0.30	0.91	-0.2	0.90	-0.2
96	US081	106	3.5	-0.59	0.30	1.78	2.2	1.77	2.2
97	US082	108	3.4	0.37	0.30	0.33	-3.1	0.33	-3.1
98	US083	108	3.1	-0.16	0.30	0.43	-2.4	0.43	-2.4
99	US084	108	1.5	-3.90	0.41	1.04	0.2	1.10	0.4
100	US085	108	3.0	-0.51	0.30	0.91	-0.2	0.91	-0.2
101	US086	108	2.8	-0.78	0.30	1.25	0.9	1.25	0.9
102	US087	108	3.0	-0.33	0.30	1.25	0.9	1.26	0.9

103	US088	108	3.0	-0.33	0.30	0.90	-0.2	0.90	-0.2
104	US089	108	3.0	-0.42	0.30	0.57	-1.7	0.57	-1.7
105	US090	108	3.3	0.21	0.30	0.81	-0.6	0.81	-0.6
106	US091	108	2.2	-2.13	0.30	1.14	0.5	1.16	0.6
107	US092	108	3.5	0.75	0.30	1.13	0.5	1.12	0.4
108	US093	108	2.8	-0.87	0.30	1.15	0.6	1.15	0.6
109	US094	108	3.7	1.12	0.30	0.61	-1.4	0.61	-1.4
110	US095	108	3.3	0.12	0.30	1.26	0.9	1.26	0.9
111	US096	108	3.1	-0.15	0.30	1.69	2.0	1.68	2.0
112	US097	108	3.6	0.93	0.30	0.92	-0.1	0.92	-0.1
113	US098	108	2.8	-0.96	0.30	0.65	-1.3	0.65	-1.3
114	US099	108	2.8	-0.96	0.30	0.65	-1.3	0.65	-1.3
115	US100	108	3.0	-0.33	0.30	0.49	-2.1	0.49	-2.1
116	US101	108	2.9	-0.57	0.30	1.00	0	1.00	0
117	US102	108	3.0	-0.45	0.30	1.15	0.6	1.15	0.5
118	US103	108	2.6	-1.25	0.30	0.91	-0.2	0.90	-0.2
119	US104	108	3.6	0.82	0.30	0.93	-0.1	0.93	-0.1
120	US105	108	2.9	-0.62	0.30	0.47	-2.2	0.47	-2.2
121	US106	108	3.6	0.91	0.30	0.77	-0.7	0.77	-0.7
122	US107	108	3.3	0.27	0.30	0.79	-0.6	0.79	-0.6
123	US108	108	3.7	1.00	0.30	0.73	-0.9	0.74	-0.8
124	US109	108	3.2	-0.09	0.30	0.79	-0.6	0.79	-0.6
125	US110	108	3.2	-0.28	0.33	1.19	0.6	1.19	0.6
126	US111	108	3.4	0.36	0.30	0.98	0	0.98	0
127	US112	108	3.3	0.27	0.30	1.02	0.1	1.02	0.1
128	US113	108	3.4	0.36	0.30	0.68	-1.1	0.68	-1.1
129	US114	108	3.1	-0.27	0.30	0.61	-1.5	0.61	-1.5
130	US115	108	3.2	-0.09	0.30	0.35	-3.0	0.35	-3.0
131	US116	108	3.1	-0.18	0.30	0.50	-2.0	0.50	-2.0
132	US117	108	2.3	-1.88	0.30	0.30	-3.4	0.30	-3.4
133	US118	108	2.3	-1.97	0.30	0.61	-1.5	0.61	-1.5
134	US119	108	3.0	-0.45	0.30	0.37	-2.8	0.37	-2.8
135	US120	108	2.8	-1.01	0.30	0.52	-2.0	0.51	-2.0
136	US121	108	2.1	-2.30	0.31	0.79	-0.7	0.78	-0.7
137	US122	108	2.6	-1.25	0.30	0.28	-3.5	0.28	-3.5
138	US123	108	2.5	-1.65	0.33	0.86	-0.3	0.86	-0.3
139	US124	108	3.2	0.01	0.30	0.34	-3.0	0.34	-3.0
140	US125	108	2.3	-1.98	0.30	1.05	0.2	1.05	0.2

141	US126	108	2.5	-1.43	0.30	0.53	-1.9	0.53	-1.9
142	US127	108	2.7	-1.16	0.30	0.47	-2.2	0.46	-2.3
143	US128	108	3.1	-0.26	0.30	0.57	-1.6	0.57	-1.6
144	US129	108	3.5	0.35	0.37	1.34	0.9	1.34	0.9
145	US130	108	2.4	-1.65	0.33	0.43	-2.3	0.43	-2.3
146	US131	108	4.6	2.98	0.36	0.71	-0.8	0.71	-0.8
147	US132	108	4.1	2.00	0.30	0.82	-0.5	0.81	-0.5
148	US133	108	4.8	3.27	0.29	1.58	1.9	1.61	2.0
149	US134	108	4.0	1.82	0.30	0.74	-0.8	0.74	-0.8
150	US135	108	5.3	4.44	0.32	1.59	1.9	1.77	2.4
151	US136	108	5.5	4.99	0.35	1.01	0.1	0.87	-0.3
152	US137	108	4.9	3.61	0.29	1.09	0.4	1.10	0.4
153	US138	108	3.8	1.28	0.30	0.54	-1.7	0.54	-1.7
154	US139	108	4.5	2.66	0.29	1.09	0.3	1.09	0.4
155	US140	108	4.3	2.41	0.29	0.52	-1.9	0.52	-1.9
156	US141	108	4.1	1.99	0.30	0.42	-2.4	0.41	-2.4
157	US142	108	4.5	2.85	0.29	0.54	-1.9	0.54	-1.9
158	US143	108	4.6	3.02	0.29	1.07	0.3	1.06	0.2
158	US144	108	5.4	4.76	0.34	1.26	0.9	1.43	1.3
160	US145	108	4.6	2.93	0.29	0.99	0	0.97	0
161	US146	108	4.5	2.85	0.29	0.39	-2.8	0.40	-2.8
162	US147	108	4.4	2.66	0.32	1.18	0.6	1.17	0.6
163	US148	108	5.0	3.70	0.30	0.59	-1.7	0.61	-1.6
164	US149	108	5.5	5.11	0.39	0.93	-0.1	0.82	-0.4
165	US150	108	5.4	4.76	0.34	0.69	-1.1	0.73	-0.9
166	US151	108	4.6	3.02	0.29	0.75	-0.9	0.75	-0.9
167	US152	108	5.0	3.78	0.32	0.85	-0.4	0.83	-0.5
168	US153	108	2.3	-1.91	0.38	0.47	-1.8	0.48	-1.8
169	US154	108	3.3	0.16	0.33	1.22	0.7	1.21	0.7
170	US155	108	3.0	-0.54	0.30	0.64	-1.3	0.64	-1.3
171	US156	108	3.3	0.27	0.30	0.56	-1.7	0.56	-1.7
172	US157	108	3.4	0.36	0.30	0.67	-1.2	0.67	-1.2
173	US158	108	3.5	0.62	0.33	1.24	0.8	1.22	0.7
174	US159	108	3.4	0.38	0.30	0.62	-1.4	0.62	-1.4
175	US160	108	2.4	-1.73	0.30	0.35	-3.0	0.35	-3.0
176	US161	108	3.4	0.35	0.33	3.15	4.4	3.13	4.4
177	US162	108	4.0	1.72	0.30	0.95	0	0.96	0
178	US163	108	4.1	2.12	0.33	0.73	-0.8	0.74	-0.7

179	US164	108	2.1	-2.35	0.31	0.46	-2.4	0.47	-2.3
180	US165	108	2.5	-1.53	0.30	1.25	0.9	1.26	0.9
181	US166	108	1.8	-3.20	0.39	0.59	-1.3	0.60	-1.3
182	US167	108	3.1	-0.18	0.30	0.98	0	0.99	0
183	US168	108	3.5	0.54	0.30	0.82	-0.5	0.82	-0.5
184	US169	108	2.4	-1.91	0.33	1.43	1.3	1.44	1.3
185	US170	108	2.5	-1.62	0.30	0.46	-2.3	0.46	-2.3
186	US171	108	3.1	-0.27	0.30	1.27	0.9	1.27	0.9
187	US172	108	3.2	0	0.30	0.88	-0.3	0.89	-0.3
188	US173	108	3.5	0.63	0.30	0.89	-0.3	0.88	-0.3
189	US174	108	3.2	-0.09	0.30	1.23	0.8	1.23	0.8
190	US175	108	4.0	1.55	0.33	0.43	-2.1	0.43	-2.1
191	US176	108	4.1	1.90	0.30	0.79	-0.6	0.80	-0.6
192	US177	108	3.8	1.27	0.30	1.29	0.9	1.29	0.9
193	US178	108	4.2	2.19	0.30	0.81	-0.5	0.81	-0.5
194	US179	108	4.0	1.72	0.33	1.23	0.7	1.25	0.8
195	US180	108	3.5	0.71	0.30	0.85	-0.4	0.85	-0.4
196	US181	108	3.3	0.26	0.30	1.01	0.1	1.01	0.1
197	US182	108	3.3	0.16	0.33	0.99	0	0.99	0
198	US183	108	5.0	3.77	0.29	1.40	1.4	1.36	1.3
199	US184	108	3.3	0.35	0.30	1.06	0.2	1.06	0.2
200	US185	108	3.6	0.79	0.33	1.10	0.4	1.10	0.4
201	US186	108	4.2	2.25	0.30	1.17	0.6	1.19	0.7
202	US187	108	3.4	0.44	0.30	0.55	-1.7	0.55	-1.7
203	US188	108	3.3	0.26	0.30	1.96	2.7	1.95	2.6
204	US189	108	3.9	1.62	0.30	1.45	1.4	1.45	1.4
205	US190	108	3.9	1.56	0.33	0.43	-2.1	0.42	-2.1
206	US191	108	3.6	0.98	0.30	0.75	-0.8	0.74	-0.8
207	US192	108	2.8	-0.82	0.30	0.67	-1.2	0.67	-1.2
208	US193	108	3.0	-0.42	0.30	0.43	-2.4	0.43	-2.4
209	US194	108	3.6	0.84	0.30	0.59	-1.5	0.59	-1.5
210	US195	108	3.6	0.70	0.33	1.05	0.2	1.05	0.2
211	US196	108	4.2	2.16	0.30	0.83	-0.5	0.83	-0.5
212	US197	108	3.6	0.79	0.30	0.77	-0.7	0.77	-0.7
213	US198	108	3.7	0.97	0.30	0.71	-0.9	0.71	-0.9
214	US199	108	3.5	0.70	0.30	0.67	-1.1	0.67	-1.1
215	US200	108	3.1	-0.21	0.30	0.75	-0.8	0.75	-0.8
216	US201	108	3.9	1.47	0.33	1.47	1.3	1.46	1.3

217	US202	108	3.4	0.42	0.30	0.85	-0.4	0.85	-0.4
218	US203	108	3.9	1.43	0.30	1.64	1.8	1.65	1.9
219	US204	108	3.4	0.33	0.30	1.01	0.1	1.01	0.1
220	US205	108	4.0	1.61	0.30	0.38	-2.6	0.38	-2.6
221	US206	108	4.3	2.32	0.29	2.14	3.0	2.13	3.0
222	US207	108	4.1	1.88	0.30	0.70	-1.0	0.71	-0.9
223	US208	108	4.8	3.33	0.29	1.43	1.5	1.44	1.5
224	US209	108	4.0	1.61	0.30	0.68	-1.0	0.69	-1.0
225	US210	108	3.8	1.42	0.33	0.75	-0.7	0.75	-0.7
226	US211	104	4.0	-0.77	0.30	0.46	-2.1	0.46	-2.1
227	US212	104	3.5	-1.68	0.30	0.41	-2.5	0.41	-2.5
229	US213	104	4.1	-0.41	0.37	0.84	-0.3	0.84	-0.3
230	US214	104	2.9	-3.12	0.30	5.60	8.2	5.60	8.2
231	US215	104	3.9	-0.94	0.30	0.31	-3.0	0.31	-3.0
233	US216	104	3.7	-1.32	0.30	0.43	-2.3	0.43	-2.3
234	US217	104	3.5	-1.86	0.30	0.85	-0.4	0.85	-0.4
235	US218	104	4.2	-0.23	0.30	0.97	0	0.97	0
237	US219	104	4.3	-0.14	0.30	0.93	-0.1	0.91	-0.2
238	US220	104	3.6	-1.50	0.30	1.27	0.9	1.28	0.9
239	US221	104	2.5	-3.95	0.30	1.52	1.7	1.51	1.6
240	US222	104	4.1	-0.32	0.30	0.76	-0.7	0.75	-0.7
241	US223	104	3.8	-1.09	0.30	0.90	-0.2	0.90	-0.2
242	US224	104	2.4	-3.88	0.37	1.69	1.8	1.69	1.7
243	US225	104	3.3	-2.18	0.30	0.99	0	0.99	0
244	US226	104	4.3	0.08	0.29	1.22	0.7	1.21	0.7
245	US227	104	3.3	-2.29	0.33	0.64	-1.2	0.64	-1.2
247	US228	104	3.0	-2.63	0.30	1.56	1.7	1.56	1.7
248	US229	104	3.5	-1.83	0.37	0.84	-0.3	0.84	-0.3
249	US230	106	3.0	-1.40	0.30	0.80	-0.6	0.80	-0.6
250	US231	106	3.1	-1.31	0.30	1.01	0.1	1.00	0.1
251	US232	106	3.3	-0.86	0.30	0.45	-2.3	0.44	-2.3
252	US233	106	2.7	-2.12	0.30	1.00	0	0.98	0
253	US234	106	3.0	-1.58	0.30	0.49	-2.1	0.48	-2.1
254	US235	106	2.7	-2.08	0.30	1.66	2.0	1.72	2.2
256	US236	106	2.4	-2.74	0.31	0.57	-1.7	0.57	-1.7
258	US237	106	3.0	-1.44	0.30	1.14	0.5	1.14	0.5
260	US238	106	3.3	-0.71	0.30	0.87	-0.3	0.86	-0.3
261	US239	106	3.5	-0.40	0.30	1.15	0.5	1.16	0.6

262	US240	106	3.1	-1.24	0.30	0.30	-3.3	0.30	-3.3
263	US241	106	3.5	-0.42	0.30	1.32	1.0	1.33	1.0
265	US242	106	3.7	0.03	0.30	1.38	1.2	1.39	1.2
266	US243	106	3.8	0.12	0.30	0.70	-1.0	0.70	-1.0
268	US244	106	3.8	0.30	0.30	0.68	-1.1	0.68	-1.1
269	US245	106	2.4	-2.81	0.31	0.60	-1.5	0.59	-1.6
270	US246	106	2.9	-1.79	0.30	0.78	-0.7	0.78	-0.7
272	US247	106	3.7	-0.06	0.30	0.80	-0.6	0.80	-0.6
273	US248	106	3.8	0.21	0.30	2.28	3.3	2.28	3.3
275	US249	106	3.7	-0.06	0.30	0.80	-0.6	0.80	-0.6
276	US250	104	3.4	-1.89	0.30	0.44	-2.3	0.44	-2.3
277	US251	104	3.2	-2.43	0.30	1.07	0.3	1.07	0.3
278	US252	104	3.4	-1.89	0.30	0.80	-0.6	0.80	-0.6
279	US253	104	3.1	-2.52	0.30	0.67	-1.2	0.67	-1.2
280	US254	104	3.9	-0.79	0.30	0.48	-2.0	0.48	-2.0
281	US255	104	2.7	-3.41	0.30	0.44	-2.4	0.44	-2.4
282	US256	104	1.9	-5.34	0.35	1.00	0	1.01	0.1
284	US257	104	3.4	-1.89	0.30	0.72	-1.0	0.71	-1.0
285	US258	104	3.5	-1.80	0.30	0.71	-1.0	0.70	-1.0
286	US259	104	4.4	0.25	0.29	0.72	-1.0	0.71	-1.0
287	US260	104	3.0	-2.71	0.30	1.31	1.0	1.31	1.0
288	US261	104	4.1	-0.32	0.30	0.70	-1.0	0.71	-0.9
289	US262	104	2.2	-4.57	0.33	0.49	-1.9	0.49	-1.9
290	US263	104	2.1	-4.66	0.31	0.53	-1.9	0.54	-1.9
291	US264	104	2.9	-2.95	0.30	0.91	-0.2	0.91	-0.2
292	US265	104	3.9	-0.78	0.30	1.61	1.8	1.62	1.8
293	US266	104	2.4	-4.02	0.30	1.13	0.5	1.13	0.5
295	US267	104	2.5	-3.75	0.30	1.10	0.4	1.10	0.4
296	US268	104	2.2	-4.57	0.31	1.19	0.7	1.19	0.7
297	US269	104	5.0	1.36	0.29	1.43	1.6	1.42	1.5
298	US270	104	3.9	-0.77	0.33	1.45	1.2	1.45	1.3
299	US271	104	3.2	-2.44	0.33	3.34	4.8	3.35	4.8
300	US272	104	4.5	0.37	0.29	1.21	0.8	1.22	0.8
301	US273	104	4.1	-0.41	0.30	0.53	-1.7	0.53	-1.7
303	US274	104	3.6	-1.47	0.30	0.57	-1.6	0.57	-1.6
304	US275	104	4.7	0.92	0.29	0.48	-2.3	0.48	-2.4
305	US276	104	3.9	-0.76	0.30	0.84	-0.4	0.82	-0.5
306	US277	104	4.4	0.30	0.29	0.63	-1.4	0.65	-1.3

307	US278	104	4.8	0.96	0.29	1.27	1.0	1.27	1.0
308	US279	104	3.5	-1.68	0.30	0.43	-2.4	0.43	-2.4
309	US280	104	4.4	0.39	0.36	0.92	-0.1	0.91	-0.1
310	US281	104	4.2	-0.22	0.30	1.18	0.6	1.19	0.6
311	US282	104	4.1	-0.40	0.30	0.35	-2.8	0.35	-2.8
312	US283	104	3.8	-0.94	0.30	1.38	1.2	1.38	1.2
313	US284	104	4.1	-0.31	0.30	0.89	-0.2	0.89	-0.2
314	US285	104	1.2	-7.55	0.54	0.88	-0.1	0.82	-0.2
316	US286	104	4.3	0.04	0.29	1.63	1.9	1.65	1.9
317	US287	104	4.3	0.13	0.29	0.85	-0.4	0.86	-0.4
318	US288	104	4.5	0.47	0.29	0.74	-0.9	0.72	-1.0
319	US289	104	3.0	-2.67	0.30	0.40	-2.6	0.40	-2.6
320	US290	104	3.4	-1.69	0.37	0.59	-1.2	0.59	-1.2
321	US291	106	3.3	-0.77	0.30	0.99	0	0.99	0
322	US292	106	5.2	3.31	0.33	1.25	0.8	1.08	0.3
323	US293	106	4.3	1.15	0.29	1.15	0.6	1.17	0.6
324	US294	106	4.4	1.31	0.32	1.44	1.4	1.43	1.3
325	US295	106	3.1	-1.35	0.30	0.67	-1.2	0.67	-1.2
326	US296	106	4.2	1.06	0.29	1.22	0.8	1.22	0.8
327	US297	106	3.9	0.45	0.30	1.04	0.2	1.03	0.1
328	US298	106	4.7	2.09	0.30	1.16	0.6	1.15	0.6
329	US299	106	3.8	0.17	0.33	2.77	3.8	2.72	3.7
330	US300	106	4.4	1.49	0.29	0.94	-0.1	0.95	-0.1
331	US301	106	3.5	-0.45	0.30	0.62	-1.3	0.62	-1.3
332	US302	106	5.2	3.12	0.32	0.93	-0.1	0.90	-0.2
333	US303	106	4.1	0.80	0.29	0.59	-1.6	0.58	-1.6
334	US304	106	5.0	2.63	0.31	1.21	0.8	1.18	0.7
335	US305	106	4.5	1.51	0.32	0.54	-1.8	0.54	-1.8
336	US306	106	3.0	-1.78	0.37	0.87	-0.2	0.87	-0.2
337	US307	110	4.8	3.84	0.29	0.45	-2.6	0.45	-2.6
338	US308	110	3.9	2.04	0.30	1.28	0.9	1.28	0.9
339	US309	110	4.1	2.49	0.30	0.73	-0.9	0.73	-0.9
340	US310	110	3.8	1.77	0.30	0.89	-0.2	0.89	-0.2
341	US311	110	4.9	4.10	0.29	0.73	-1.0	0.74	-1.0
342	US312	110	5.0	4.27	0.29	1.01	0.1	1.00	0
343	US313	110	4.7	3.59	0.29	0.70	-1.2	0.69	-1.2
344	US314	110	4.5	3.26	0.29	0.69	-1.1	0.69	-1.1
345	US315	110	5.0	4.19	0.29	1.03	0.1	1.01	0.1

346	US316	110	4.0	2.33	0.30	0.97	0	0.96	0
347	US317	110	4.4	3.03	0.29	1.53	1.7	1.53	1.7
348	US318	110	4.1	2.42	0.30	1.23	0.8	1.23	0.8
349	US319	110	4.8	3.86	0.29	0.77	-0.8	0.77	-0.8
350	US320	110	5.3	4.92	0.32	0.80	-0.7	0.78	-0.8
351	US321	110	3.1	0.27	0.30	0.63	-1.3	0.63	-1.3
352	US322	110	3.0	0.18	0.30	2.38	3.6	2.39	3.6
353	US323	110	2.8	-0.35	0.30	1.46	1.5	1.46	1.5
354	US324	110	4.3	2.80	0.29	0.63	-1.3	0.64	-1.3
355	US325	110	3.3	0.63	0.30	0.64	-1.3	0.64	-1.3
356	US326	110	3.1	0.39	0.33	2.46	3.4	2.45	3.4
357	US327	110	3.1	0.34	0.30	1.64	1.9	1.64	1.9
358	US328	110	3.9	1.98	0.30	0.60	-1.4	0.59	-1.4
359	US329	110	3.0	0.16	0.30	1.07	0.3	1.06	0.3
360	US330	110	3.2	0.59	0.33	0.46	-2.0	0.46	-2.0
361	US331	110	4.0	2.36	0.30	0.92	-0.1	0.91	-0.1
362	US332	110	2.6	-1.06	0.37	1.11	0.4	1.10	0.4
363	US333	110	3.3	0.78	0.30	0.88	-0.3	0.87	-0.3
364	US334	110	4.0	2.33	0.30	0.79	-0.6	0.79	-0.6
365	US335	110	4.9	4.29	0.29	0.94	-0.1	0.93	-0.1
366	US336	110	4.3	2.95	0.29	0.93	-0.1	0.93	-0.1
367	US337	110	3.4	1.05	0.30	0.60	-1.5	0.60	-1.5
368	US338	110	3.5	1.33	0.30	0.50	-2.0	0.50	-2.0
369	US339	110	4.5	3.39	0.29	1.02	0.1	1.02	0.1
370	US340	110	3.5	1.26	0.30	1.65	1.9	1.64	1.9
371	US341	110	3.3	0.89	0.30	1.85	2.4	1.84	2.4
372	US342	110	4.8	3.97	0.29	1.11	0.4	1.09	0.4
373	US343	110	4.5	3.31	0.29	1.74	2.3	1.79	2.4
374	US344	110	4.8	4.05	0.29	1.28	1.0	1.30	1.1
375	US345	110	4.7	3.80	0.29	0.46	-2.5	0.46	-2.4
376	US346	110	4.1	2.53	0.30	0.90	-0.2	0.90	-0.2
377	US347	110	4.1	2.62	0.30	0.44	-2.3	0.44	-2.3
378	US348	110	4.1	2.62	0.30	1.00	0	1.00	0.1
379	US349	110	3.8	1.90	0.30	1.30	1.0	1.29	0.9
380	US350	110	4.0	2.26	0.30	0.97	0	0.97	0
381	US351	110	4.6	3.56	0.29	1.47	1.6	1.47	1.6
382	US352	110	3.4	0.99	0.30	1.03	0.2	1.03	0.2
383	US353	110	4.8	3.83	0.29	1.18	0.7	1.18	0.7

384	US354	110	4.1	2.39	0.30	1.10	0.4	1.11	0.4
385	US355	110	3.9	2.03	0.30	1.47	1.4	1.47	1.4
386	US356	110	4.5	3.25	0.29	1.20	0.7	1.20	0.7
387	US357	110	4.1	2.39	0.30	1.67	1.9	1.66	1.9
388	US358	110	3.9	2.08	0.30	0.67	-1.1	0.67	-1.1
389	US359	110	4.1	2.53	0.30	0.98	0	0.98	0
390	US360	110	3.4	0.89	0.30	2.51	3.8	2.52	3.8
391	US361	110	3.9	2.08	0.30	1.04	0.2	1.04	0.2
392	US362	110	3.3	0.71	0.30	1.34	1.1	1.33	1.1
393	US363	110	3.6	1.24	0.37	1.22	0.6	1.21	0.6
394	US364	110	3.0	-0.06	0.30	0.93	-0.1	0.93	-0.1
395	US365	110	2.2	-1.67	0.37	2.02	2.4	2.00	2.4
396	US366	110	3.9	1.93	0.30	4.31	6.3	4.30	6.2
397	US367	110	3.5	1.11	0.30	0.74	-0.8	0.74	-0.8
398	US368	110	4.3	2.83	0.29	0.85	-0.4	0.84	-0.4
399	US369	110	3.1	0.30	0.30	0.77	-0.7	0.77	-0.7
400	US370	110	3.6	1.30	0.30	3.19	4.9	3.19	4.8
401	US371	110	4.2	2.63	0.30	0.77	-0.7	0.76	-0.7
402	US372	110	3.2	0.46	0.30	1.54	1.7	1.55	1.7
403	US373	110	3.3	0.64	0.30	0.88	-0.3	0.87	-0.3
404	US374	110	2.8	-0.35	0.30	1.02	0.1	1.02	0.1
405	US375	110	3.0	0.10	0.30	0.89	-0.3	0.89	-0.3
406	US376	110	2.6	-0.78	0.33	1.99	2.6	1.99	2.6
407	US377	110	3.3	0.73	0.30	1.04	0.2	1.04	0.2
408	US378	110	3.3	0.73	0.30	1.28	0.9	1.28	0.9
409	US379	110	3.3	0.73	0.30	1.01	0.1	1.01	0.1
410	US380	110	4.3	2.80	0.29	0.67	-1.2	0.65	-1.2
411	US381	110	3.3	0.63	0.30	0.72	-0.9	0.72	-1.0
412	US382	110	3.1	0.18	0.30	0.47	-2.2	0.47	-2.2
413	US383	110	3.4	0.79	0.33	1.37	1.1	1.37	1.1
414	US384	110	3.7	1.45	0.30	0.43	-2.3	0.43	-2.3
415	US385	110	4.4	3.06	0.29	1.25	0.9	1.23	0.8
416	US386	110	4.0	2.28	0.30	0.81	-0.5	0.82	-0.5
417	US387	110	4.6	3.40	0.29	1.16	0.6	1.17	0.6
418	US388	110	3.8	1.64	0.30	0.97	0	0.97	0
419	US389	110	4.6	3.49	0.29	1.58	1.9	1.56	1.8
420	US390	110	4.1	2.38	0.30	1.00	0	0.99	0
421	US391	110	4.1	2.47	0.30	0.55	-1.7	0.54	-1.7

422	US392	110	4.2	2.65	0.30	1.68	2.0	1.66	1.9
423	US393	110	4.7	3.66	0.29	1.55	1.8	1.57	1.9
424	US394	110	3.6	1.38	0.30	0.78	-0.6	0.78	-0.6
425	US395	110	3.6	1.38	0.30	1.25	0.8	1.24	0.8
426	US396	110	5.3	4.78	0.31	0.95	0	0.92	-0.2
427	US397	110	4.5	3.24	0.29	0.85	-0.4	0.83	-0.5
428	US398	110	4.0	2.20	0.30	0.75	-0.8	0.76	-0.7
429	US399	106	3.5	-0.52	0.30	0.76	-0.7	0.76	-0.7
430	US400	106	3.7	-0.09	0.30	0.72	-0.9	0.72	-0.9
431	US401	106	4.3	1.14	0.29	2.23	3.4	2.28	3.4
432	US402	106	3.8	0.18	0.30	1.38	1.2	1.38	1.2
433	US403	106	3.3	-0.92	0.30	0.82	-0.5	0.83	-0.5
434	US404	106	2.2	-3.42	0.31	0.40	-2.7	0.40	-2.8
436	US405	106	3.7	-0.02	0.30	0.77	-0.7	0.78	-0.7
437	US406	106	3.4	-0.65	0.30	0.95	0	0.95	0
438	US407	106	2.9	-1.83	0.30	0.73	-0.9	0.72	-0.9
439	US408	106	4.1	0.78	0.29	0.50	-2.0	0.51	-2.0
440	US409	106	4.5	1.55	0.29	0.96	0	0.96	0
441	US410	106	5.2	3.20	0.33	1.27	0.9	1.14	0.5
442	US411	106	4.1	0.87	0.29	1.32	1.1	1.35	1.2
443	US412	106	3.4	-0.74	0.30	0.41	-2.5	0.42	-2.4
444	US413	106	2.8	-1.92	0.30	0.57	-1.7	0.57	-1.7
445	US414	106	2.9	-1.83	0.30	1.38	1.3	1.38	1.2
446	US415	106	3.5	-0.47	0.30	1.37	1.2	1.35	1.1
447	US416	106	4.2	1.04	0.29	0.67	-1.2	0.67	-1.2
448	US417	106	3.2	-1.20	0.30	0.47	-2.2	0.47	-2.2
449	US418	104	3.1	-2.50	0.30	0.66	-1.2	0.66	-1.2
450	US419	104	2.8	-3.12	0.30	0.68	-1.1	0.68	-1.1
452	US420	104	2.2	-4.57	0.31	0.30	-3.4	0.30	-3.5
454	US421	104	3.6	-1.45	0.30	1.56	1.7	1.57	1.7
455	US422	104	3.8	-0.99	0.30	0.99	0	0.99	0
456	US423	104	3.1	-2.54	0.30	0.89	-0.2	0.89	-0.2
457	US424	104	3.5	-1.73	0.30	0.48	-2.1	0.49	-2.1
458	US425	104	3.9	-0.90	0.30	0.37	-2.6	0.37	-2.6
459	US426	104	3.6	-1.45	0.30	0.66	-1.2	0.66	-1.2
460	US427	104	3.3	-2.09	0.30	0.85	-0.4	0.84	-0.4
462	US428	104	4.3	-0.09	0.30	0.99	0	0.99	0
463	US429	104	2.9	-3.08	0.30	0.48	-2.1	0.49	-2.1

465	US430	104	3.0	-2.72	0.30	0.64	-1.3	0.64	-1.3
466	US431	104	2.6	-3.61	0.30	0.46	-2.3	0.45	-2.3
467	US432	104	2.2	-4.43	0.37	0.62	-1.1	0.63	-1.1
468	US433	104	3.7	-1.27	0.30	1.42	1.3	1.42	1.3
469	US434	104	3.4	-1.91	0.33	0.89	-0.2	0.89	-0.2
470	US435	104	3.0	-2.90	0.30	1.34	1.1	1.34	1.1
471	US436	104	4.3	0.08	0.29	0.94	-0.1	0.97	0
473	US437	104	2.3	-4.24	0.30	0.96	0	0.96	0
474	US438	106	3.5	-0.31	0.30	1.13	0.5	1.13	0.5
475	US439	106	3.3	-0.86	0.30	0.94	-0.1	0.94	-0.1
478	US440	106	5.0	2.77	0.31	0.56	-1.8	0.56	-1.8
479	US441	106	3.6	-0.22	0.30	0.92	-0.1	0.93	-0.1
480	US442	106	4.4	1.51	0.29	2.47	3.9	2.42	3.8
481	US443	106	4.3	1.25	0.29	0.97	0	0.98	0
483	US444	106	4.8	2.19	0.30	1.23	0.8	1.24	0.9
484	US445	106	3.4	-0.70	0.30	0.54	-1.8	0.54	-1.8
485	US446	106	4.3	1.34	0.29	2.22	3.3	2.26	3.4
486	US447	106	4.5	1.76	0.29	1.49	1.6	1.48	1.6
487	US448	106	2.5	-2.61	0.30	1.12	0.4	1.12	0.5
489	US449	106	4.7	2.02	0.29	0.93	-0.1	0.92	-0.2
490	US450	106	3.2	-1.07	0.30	1.01	0.1	1.01	0.1
492	US451	106	4.0	0.64	0.30	1.85	2.4	1.90	2.5
493	US452	106	4.5	1.76	0.29	1.34	1.2	1.34	1.2
494	US453	106	4.0	0.64	0.30	0.81	-0.6	0.81	-0.5
495	US454	106	4.0	0.73	0.30	0.85	-0.4	0.84	-0.4
497	US455	106	2.9	-1.69	0.30	0.67	-1.2	0.66	-1.2
498	US456	106	3.4	-0.56	0.30	1.56	1.7	1.57	1.7
499	US457	106	3.7	-0.31	0.33	1.04	0.2	1.04	0.2
500	US458	106	3.9	0.43	0.30	0.64	-1.3	0.64	-1.3
501	US459	106	2.8	-1.83	0.30	1.07	0.3	1.06	0.3
503	US460	106	2.8	-1.95	0.30	0.77	-0.7	0.77	-0.8
505	US461	106	3.3	-0.77	0.30	0.99	0	0.99	0
506	US462	106	3.4	-0.59	0.30	1.23	0.8	1.23	0.8
508	US463	106	3.5	-0.32	0.30	0.88	-0.3	0.88	-0.3
510	US464	106	3.9	0.50	0.30	1.51	1.6	1.51	1.6
511	US465	106	2.3	-2.88	0.31	1.39	1.3	1.33	1.1
512	US466	106	4.0	0.85	0.30	0.92	-0.1	0.94	-0.1
514	US467	106	3.6	-0.13	0.30	1.23	0.8	1.23	0.8

515	US468	106	3.8	0.41	0.30	0.96	0	0.97	0
516	US469	106	3.5	-0.22	0.30	0.64	-1.3	0.64	-1.2
517	US470	106	3.8	0.32	0.30	0.41	-2.4	0.41	-2.4
519	US471	106	4.2	1.11	0.29	0.79	-0.6	0.79	-0.7
520	US472	106	3.5	-0.44	0.30	0.60	-1.4	0.60	-1.4
521	US473	112	4.0	2.37	0.30	0.59	-1.5	0.59	-1.5
522	US474	112	3.9	2.19	0.30	0.57	-1.6	0.57	-1.6
523	US475	112	3.8	2.02	0.30	0.87	-0.3	0.86	-0.3
524	US476	112	3.9	2.12	0.30	1.11	0.4	1.12	0.4
525	US477	112	4.1	2.65	0.30	0.82	-0.5	0.82	-0.5
526	US478	112	4.1	2.56	0.30	1.09	0.3	1.09	0.4
527	US479	112	4.2	2.74	0.30	0.54	-1.8	0.54	-1.7
528	US480	112	3.6	1.47	0.30	1.21	0.7	1.22	0.7
529	US481	112	4.3	2.91	0.29	0.72	-0.9	0.72	-0.9
530	US482	112	4.1	2.65	0.30	1.35	1.1	1.36	1.1
531	US483	112	4.5	3.42	0.29	1.41	1.4	1.40	1.3
532	US484	112	3.7	1.75	0.30	1.18	0.6	1.19	0.6
533	US485	112	3.8	1.84	0.30	0.80	-0.6	0.80	-0.6
534	US486	112	4.5	3.51	0.29	0.76	-0.8	0.74	-0.9
535	US487	112	3.8	1.81	0.33	0.65	-1.1	0.65	-1.0
536	US488	112	4.2	2.74	0.30	0.61	-1.4	0.60	-1.4
537	US489	112	4.7	3.84	0.29	0.87	-0.4	0.85	-0.5
538	US490	112	3.8	1.93	0.30	0.71	-0.9	0.70	-1.0
539	US491	112	4.0	2.29	0.30	1.46	1.4	1.47	1.4
540	US492	112	2.8	-0.24	0.30	0.74	-0.8	0.75	-0.8
541	US493	112	4.1	2.55	0.30	1.39	1.2	1.37	1.2
542	US494	112	3.4	1.01	0.30	0.92	-0.2	0.92	-0.1
543	US495	112	3.5	1.20	0.30	0.84	-0.4	0.84	-0.4
544	US496	112	3.3	0.83	0.30	0.97	0	0.97	0
545	US497	112	3.5	1.20	0.30	1.36	1.1	1.36	1.2
546	US498	112	3.2	0.56	0.30	0.44	-2.4	0.44	-2.4
547	US499	112	2.4	-1.14	0.30	2.35	3.6	2.32	3.5
548	US500	112	3.3	0.74	0.30	1.20	0.7	1.20	0.7
549	US501	112	3.9	2.11	0.30	0.93	-0.1	0.93	-0.1
550	US502	112	2.6	-0.69	0.30	0.53	-1.9	0.54	-1.9
551	US503	112	3.6	1.47	0.30	0.54	-1.7	0.54	-1.8
552	US504	112	2.1	-1.70	0.31	1.27	0.9	1.23	0.8
553	US505	112	2.3	-1.20	0.33	1.47	1.4	1.45	1.3

554	US506	112	2.8	-0.24	0.30	1.25	0.9	1.25	0.9
555	US507	112	3.3	0.93	0.30	0.27	-3.6	0.27	-3.6
556	US508	112	3.4	1.11	0.30	1.27	0.9	1.27	0.9
557	US509	112	3.4	1.02	0.30	1.31	1.0	1.32	1.0
558	US510	112	3.8	1.84	0.30	1.23	0.8	1.23	0.8
559	US511	112	2.9	-0.02	0.33	1.32	1.0	1.32	1.0
560	US512	112	3.5	1.28	0.30	0.65	-1.2	0.65	-1.2
561	US513	112	3.7	1.75	0.30	3.90	5.8	3.92	5.9
562	US514	112	3.4	1.11	0.30	0.58	-1.6	0.58	-1.6
563	US515	112	2.9	0.03	0.30	0.72	-0.9	0.73	-0.9
564	US516	112	3.1	0.42	0.30	2.14	3.1	2.14	3.1
565	US517	112	5.1	4.73	0.30	1.11	0.5	1.12	0.5
566	US518	112	3.8	1.97	0.30	1.59	1.7	1.60	1.7
567	US519	112	3.8	1.97	0.30	0.93	-0.1	0.92	-0.1
568	US520	112	3.3	0.78	0.30	0.91	-0.2	0.90	-0.2
569	US521	112	4.4	3.21	0.29	0.99	0	0.98	0
570	US522	112	4.1	2.60	0.30	0.57	-1.6	0.57	-1.6
571	US523	112	5.6	5.90	0.39	0.79	-0.6	0.90	-0.2
572	US524	112	4.3	2.95	0.30	0.90	-0.2	0.91	-0.2
573	US525	112	4.5	3.47	0.29	0.30	-3.4	0.30	-3.4
574	US526	112	4.3	3.04	0.29	0.78	-0.7	0.76	-0.7
575	US527	112	4.2	2.86	0.30	1.03	0.1	1.03	0.2
576	US528	112	4.3	3.04	0.29	1.87	2.5	1.79	2.3
577	US529	112	3.8	1.97	0.30	1.09	0.4	1.09	0.3
578	US530	112	3.9	2.24	0.30	1.00	0.1	1.00	0.1
579	US531	112	4.6	3.74	0.29	1.63	2.0	1.65	2.1
580	US532	112	4.1	2.71	0.30	0.96	0	0.96	0
581	US533	112	5.2	4.91	0.31	1.14	0.6	1.10	0.4
582	US534	112	4.3	2.94	0.29	0.51	-2.0	0.50	-2.0
583	US535	112	3.8	1.86	0.30	0.98	0	0.98	0
584	US536	112	4.2	2.90	0.29	1.31	1.0	1.30	1.0
585	US537	112	4.0	2.46	0.30	1.74	2.1	1.72	2.0
586	US538	112	3.5	1.37	0.30	0.91	-0.2	0.90	-0.2
587	US539	112	4.9	4.26	0.29	1.96	2.9	1.96	3.0
588	US540	112	5.2	4.89	0.31	1.95	2.9	2.09	3.3
589	US541	112	4.0	2.37	0.30	1.09	0.4	1.09	0.3
590	US542	112	3.9	2.28	0.30	1.06	0.3	1.07	0.3
591	US543	112	4.4	3.33	0.29	0.74	-0.9	0.74	-0.9

592	US544	112	4.6	3.67	0.29	0.67	-1.3	0.68	-1.2
593	US545	112	4.1	2.63	0.30	0.87	-0.3	0.89	-0.2
594	US546	112	4.1	2.72	0.30	0.50	-2.0	0.50	-1.9
595	US547	112	3.8	1.91	0.30	0.29	-3.3	0.29	-3.3
596	US548	112	4.3	2.98	0.29	0.59	-1.6	0.59	-1.5
597	US549	112	3.3	1.01	0.30	0.54	-1.8	0.54	-1.8
598	US550	112	3.9	2.28	0.33	0.65	-1.1	0.64	-1.1
599	US551	112	3.5	1.27	0.30	0.71	-0.9	0.72	-0.9
600	US552	112	4.5	3.58	0.29	0.61	-1.5	0.63	-1.4
601	US553	112	4.4	3.26	0.29	0.53	-1.9	0.53	-1.9
602	US554	112	3.6	1.48	0.30	0.71	-1.0	0.71	-0.9
603	US555	112	4.9	4.38	0.32	0.55	-1.8	0.56	-1.7
604	US556	112	3.9	2.21	0.30	0.62	-1.3	0.61	-1.4
605	US557	112	3.5	1.28	0.30	0.68	-1.1	0.68	-1.1
606	US558	112	2.5	-0.87	0.30	2.32	3.5	2.30	3.5
607	US559	112	3.1	0.47	0.30	0.98	0	0.98	0
608	US560	112	3.3	0.83	0.30	0.41	-2.5	0.41	-2.5
609	US561	112	2.2	-1.51	0.31	0.61	-1.5	0.60	-1.6
610	US562	112	3.3	0.92	0.30	1.01	0.1	1.00	0.1
611	US563	112	2.5	-0.78	0.30	0.57	-1.7	0.57	-1.7
612	US564	112	3.5	1.28	0.30	1.19	0.7	1.20	0.7
613	US565	112	3.9	2.10	0.30	1.19	0.6	1.19	0.7
614	US566	112	3.2	0.56	0.30	0.74	-0.9	0.73	-0.9
615	US567	112	3.8	1.83	0.30	1.38	1.2	1.38	1.2
616	US568	112	2.3	-1.28	0.30	2.51	3.9	2.48	3.9
617	US569	112	2.9	0.07	0.30	0.74	-0.9	0.74	-0.9
618	US570	112	2.5	-0.92	0.30	1.97	2.8	1.96	2.8
619	US571	104	4.7	0.76	0.29	0.85	-0.5	0.85	-0.5
620	US572	104	3.2	-2.44	0.30	1.37	1.2	1.36	1.2
621	US573	104	4.5	0.51	0.29	1.03	0.1	1.02	0.1
622	US574	104	3.5	-1.62	0.30	1.02	0.1	1.01	0.1
623	US575	104	3.8	-1.16	0.30	0.49	-1.9	0.49	-1.9
624	US576	104	4.0	-0.62	0.30	1.08	0.3	1.09	0.3
625	US577	104	3.0	-2.79	0.30	2.32	3.4	2.32	3.4
627	US578	104	5.0	1.50	0.32	1.44	1.5	1.44	1.5
628	US579	104	3.8	-0.98	0.30	4.87	6.9	4.87	6.9
629	US580	104	3.5	-1.71	0.30	1.08	0.3	1.08	0.3
630	US581	104	3.1	-2.53	0.30	1.60	1.8	1.60	1.8

631	US582	104	3.9	-1.00	0.33	1.03	0.1	1.03	0.1
632	US583	104	4.0	-0.62	0.30	0.91	-0.1	0.92	-0.1
633	US584	104	3.8	-1.07	0.30	0.69	-1.0	0.69	-1.0
634	US585	104	3.5	-1.71	0.30	1.56	1.7	1.57	1.7
635	US586	104	2.9	-3.06	0.30	1.03	0.1	1.03	0.1
637	US587	104	2.7	-3.42	0.30	0.92	-0.2	0.92	-0.2
638	US588	104	3.5	-1.85	0.33	0.87	-0.3	0.87	-0.3
639	US589	104	3.5	-1.76	0.30	0.62	-1.4	0.62	-1.4

Separation: 6.69 Reliability: 0.98

Fixed (all same) chi-square: 26019.5 significance p: .00

Note: 104 = Grade 4 US student, 106 = Grade 6 US student, 108 = Grade 8 US student, 110 = Grade 10 US student, 112 = Grade 12 US student

Calibration of the students for US student essays using US raters (85 misfitting students)

Student			Observed	Measure	S.E.	Infit		Outfit	
Number		Group	score	(Logit)		MnSq	ZStd	MnSq	ZStd
531	US483	112	4.5	3.42	0.29	1.41	1.4	1.40	1.3
24	US021	104	3.3	-2.03	0.30	1.42	1.3	1.41	1.3
468	US433	104	3.7	-1.27	0.30	1.42	1.3	1.42	1.3
184	US169	108	2.4	-1.91	0.33	1.43	1.3	1.44	1.3
223	US208	108	4.8	3.33	0.29	1.43	1.5	1.44	1.5
297	US269	104	5.0	1.36	0.29	1.43	1.6	1.42	1.5
324	US294	106	4.4	1.31	0.32	1.44	1.4	1.43	1.3
627	US578	104	5.0	1.50	0.32	1.44	1.5	1.44	1.5
204	US189	108	3.9	1.62	0.30	1.45	1.4	1.45	1.4
298	US270	104	3.9	-0.77	0.33	1.45	1.2	1.45	1.3
62	US075	106	4.3	1.12	0.29	1.46	1.5	1.46	1.5
353	US323	110	2.8	-0.35	0.30	1.46	1.5	1.46	1.5
539	US491	112	4.0	2.29	0.30	1.46	1.4	1.47	1.4
216	US201	108	3.9	1.47	0.33	1.47	1.3	1.46	1.3
381	US351	110	4.6	3.56	0.29	1.47	1.6	1.47	1.6
385	US355	110	3.9	2.03	0.30	1.47	1.4	1.47	1.4
553	US505	112	2.3	-1.20	0.33	1.47	1.4	1.45	1.3
486	US447	106	4.5	1.76	0.29	1.49	1.6	1.48	1.6
68	US057	104	3.8	-1.01	0.37	1.50	1.3	1.50	1.3
510	US464	106	3.9	0.50	0.30	1.51	1.6	1.51	1.6
239	US221	104	2.5	-3.95	0.30	1.52	1.7	1.51	1.6

347	US317	110	4.4	3.03	0.29	1.53	1.7	1.53	1.7
402	US372	110	3.2	0.46	0.30	1.54	1.7	1.55	1.7
423	US393	110	4.7	3.66	0.29	1.55	1.8	1.57	1.9
247	US228	104	3.0	-2.63	0.30	1.56	1.7	1.56	1.7
454	US421	104	3.6	-1.45	0.30	1.56	1.7	1.57	1.7
498	US456	106	3.4	-0.56	0.30	1.56	1.7	1.57	1.7
634	US585	104	3.5	-1.71	0.30	1.56	1.7	1.57	1.7
19	US017	104	1.6	-6.00	0.39	1.58	1.8	1.48	1.5
148	US133	108	4.8	3.27	0.29	1.58	1.9	1.61	2.0
419	US389	110	4.6	3.49	0.29	1.58	1.9	1.56	1.8
150	US135	108	5.3	4.44	0.32	1.59	1.9	1.77	2.4
566	US518	112	3.8	1.97	0.30	1.59	1.7	1.60	1.7
630	US581	104	3.1	-2.53	0.30	1.60	1.8	1.60	1.8
292	US265	104	3.9	-0.78	0.30	1.61	1.8	1.62	1.8
316	US286	104	4.3	0.04	0.29	1.63	1.9	1.65	1.9
579	US531	112	4.6	3.74	0.29	1.63	2.0	1.65	2.1
218	US203	108	3.9	1.43	0.30	1.64	1.8	1.65	1.9
357	US327	110	3.1	0.34	0.30	1.64	1.9	1.64	1.9
21	US018	104	3.8	-1.04	0.30	1.65	1.9	1.65	1.9
370	US340	110	3.5	1.26	0.30	1.65	1.9	1.64	1.9
44	US035	106	4.3	1.09	0.29	1.66	2.0	1.72	2.2
254	US235	106	2.7	-2.08	0.30	1.66	2.0	1.72	2.2
387	US357	110	4.1	2.39	0.30	1.67	1.9	1.66	1.9
422	US392	110	4.2	2.65	0.30	1.68	2.0	1.66	1.9
111	US096	108	3.1	-0.15	0.30	1.69	2.0	1.68	2.0
242	US224	104	2.4	-3.88	0.37	1.69	1.8	1.69	1.7
373	US343	110	4.5	3.31	0.29	1.74	2.3	1.79	2.4
585	US537	112	4.0	2.46	0.30	1.74	2.1	1.72	2.0
96	US081	106	3.5	-0.59	0.30	1.78	2.2	1.77	2.2
371	US341	110	3.3	0.89	0.30	1.85	2.4	1.84	2.4
492	US451	106	4.0	0.64	0.30	1.85	2.4	1.90	2.5
576	US528	112	4.3	3.04	0.29	1.87	2.5	1.79	2.3
88	US073	106	2.7	-2.35	0.30	1.91	2.6	1.90	2.6
59	US049	104	2.6	-3.43	0.30	1.95	2.7	1.95	2.7
588	US540	112	5.2	4.89	0.31	1.95	2.9	2.09	3.3
203	US188	108	3.3	0.26	0.30	1.96	2.7	1.95	2.6
587	US539	112	4.9	4.26	0.29	1.96	2.9	1.96	3.0
618	US570	112	2.5	-0.92	0.30	1.97	2.8	1.96	2.8

406	US376	110	2.6	-0.78	0.33	1.99	2.6	1.99	2.6
395	US365	110	2.2	-1.67	0.37	2.02	2.4	2.00	2.4
30	US025	1106	2.8	-1.96	0.30	2.06	2.9	2.05	2.9
221	US206	108	4.3	2.32	0.29	2.14	3.0	2.13	3.0
564	US516	112	3.1	0.42	0.30	2.14	3.1	2.14	3.1
485	US446	106	4.3	1.34	0.29	2.22	3.3	2.26	3.4
70	US058	104	2.5	-3.87	0.30	2.23	3.4	2.24	3.4
431	US401	106	4.3	1.14	0.29	2.23	3.4	2.28	3.4
273	US248	106	3.8	0.21	0.30	2.28	3.3	2.28	3.3
606	US558	112	2.5	-0.87	0.30	2.32	3.5	2.30	3.5
625	US577	104	3.0	-2.79	0.30	2.32	3.4	2.32	3.4
547	US499	112	2.4	-1.14	0.30	2.35	3.6	2.32	3.5
352	US322	110	3.0	0.18	0.30	2.38	3.6	2.39	3.6
356	US326	110	3.1	0.39	0.33	2.46	3.4	2.45	3.4
480	US442	106	4.4	1.51	0.29	2.47	3.9	2.42	3.8
390	US360	110	3.4	0.89	0.30	2.51	3.8	2.52	3.8
616	US568	112	2.3	-1.28	0.30	2.51	3.9	2.48	3.9
86	US071	106	3.8	0.01	0.30	2.56	3.8	2.53	3.7
329	US299	106	3.8	0.17	0.33	2.77	3.8	2.72	3.7
176	US161	108	3.4	0.35	0.33	3.15	4.4	3.13	4.4
400	US370	110	3.6	1.30	0.30	3.19	4.9	3.19	4.8
299	US271	104	3.2	-2.44	0.33	3.34	4.8	3.35	4.8
561	US513	112	3.7	1.75	0.30	3.90	5.8	3.92	5.9
396	US366	110	3.9	1.93	0.30	4.31	6.3	4.30	6.2
628	US579	104	3.8	-0.98	0.30	4.87	6.9	4.87	6.9
230	US214	104	2.9	-3.12	0.30	5.60	8.2	5.60	8.2

Separation: 6.69 Reliability: 0.98

Fixed (all same) chi-square: 26019.5 significance *p*: .00

Note: 104 = Grade 4 US student, 106 = Grade 6 US student, 108 = Grade 8 US student, 110 = Grade 10 US student, 112 = Grade 12 US student

Calibration of the prompts for US student essays using US raters

Prompt	Observed score	Measure (Logit)	S.E.	Infit MnSq	Infit ZStd	Outfit MnSq	Outfit ZStd
4AI	3.6	-2.01	0.06	0.80	-3.5	0.79	-3.6
4AN	3.7	-2.09	0.06	0.85	-2.6	0.86	-2.4
4AP	3.6	-2.07	0.06	1.39	5.6	1.39	5.6
4BI	3.7	-1.67	0.05	1.02	0.5	1.01	0.3

4BN	3.8	-1.49	0.05	0.76	-5.9	0.77	-5.8
4BP	3.6	-1.37	0.05	1.31	6.4	1.30	6.3
8AI	3.5	0	0.05	1.08	1.7	1.08	1.7
8AN	3.7	-0.15	0.05	0.77	-5.5	0.78	-5.4
8AP	2.9	1.35	0.05	0.95	-1.1	0.95	-1.1
8BI	3.8	0.46	0.05	1.03	0.6	1.02	0.5
8BN	3.7	0.85	0.05	0.77	-5.3	0.76	-5.4
8BP	3.6	0.73	0.05	1.08	1.6	1.07	1.5
12AI	3.8	1.14	0.05	1.03	0.5	1.03	0.5
12AN	3.9	1.08	0.05	0.92	-1.4	0.93	-1.3
12AP	3.4	1.99	0.05	1.04	0.7	1.04	0.6
12BI	3.5	1.82	0.07	0.98	-0.3	0.97	-0.3
12BN	4.0	0.83	0.07	1.30	3.9	1.31	3.9
12BP	4.0	0.59	0.08	1.26	3.2	1.27	3.4

Separation: 24.19 Reliability: 1.00

Fixed (all same) chi-square: 10971.5 significance *p*: .00

Calibration of the US raters for US student essays using US raters

US rater	Observed score	Measure (Logit)	S.E.	Infit		Outfit	
				MnSq	ZStd	MnSq	ZStd
USR101	3.9	-0.31	0.05	0.85	-3.4	0.86	-3.1
USR102	3.8	-0.15	0.05	0.88	-2.8	0.87	-2.9
USR103	3.6	0.26	0.09	0.83	-2.1	0.82	-2.2
USR104	3.6	0.30	0.05	0.80	-4.7	0.80	-4.6
USR105	3.8	-0.16	0.05	0.90	-2.1	0.90	-2.0
USR106	4.0	-0.68	0.06	0.81	-3.5	0.82	-3.3
USR108	3.3	-1.43	0.34	0.65	-1.0	0.65	-1.1
USR200	3.2	0.97	0.06	1.22	3.6	1.22	3.5
USR201	3.6	0.01	0.05	1.11	2.3	1.10	2.2
USR202	3.6	-0.20	0.06	1.06	1.0	1.06	0.9
USR203	3.6	0.10	0.04	1.17	3.9	1.17	3.8
USR204	3.7	0.12	0.05	0.93	-1.6	0.93	-1.6
USR205	3.5	0.09	0.05	1.03	0.5	1.04	0.7
USR206	3.4	0.36	0.05	0.88	-2.7	0.88	-2.7
USR207	3.7	-0.28	0.05	1.13	2.3	1.12	2.3
USR208	3.7	-0.16	0.05	0.97	-0.4	0.97	-0.5
USR209	3.4	0.59	0.05	1.04	0.7	1.03	0.6
USR210	3.7	0.01	0.05	1.34	6.6	1.33	6.5

USR211	3.3	0.57	0.35	0.65	-1.0	0.63	-1.1
Separation: 3.92				Reliability: 0.94			
Fixed (all same) chi-square: 746.5				significance *p*: .00			

Appendix 12

Calibration of the students for US student essays using US raters, Lexile Analyzer and IEA

Student			Observed	Measure	S.E.	Infit		Outfit	
Number		Group	score	(Logit)		MnSq	ZStd	MnSq	ZStd
1	US001	104	3.7	0.11	0.35	0.90	-0.2	0.90	-0.2
2	US002	104	3.8	0.23	0.39	1.17	0.6	1.17	0.5
3	US003	104	4.1	0.63	0.35	0.99	0	0.98	0
4	US004	104	4.2	0.14	0.34	0.54	-1.8	0.56	-1.7
5	US005	104	3.1	-1.99	0.35	1.51	1.6	1.52	1.7
6	US006	104	4.1	0.23	0.35	0.36	-2.7	0.35	-2.7
8	US007	104	3.3	-1.20	0.35	1.12	0.4	1.12	0.4
9	US008	104	3.2	-0.49	0.35	0.54	-1.8	0.54	-1.8
10	US009	104	4.1	0.82	0.35	1.31	1.0	1.32	1.0
11	US010	104	3.5	-0.55	0.35	0.62	-1.4	0.61	-1.4
12	US011	104	5.3	1.84	0.34	0.79	-0.7	0.76	-0.8
13	US012	104	1.7	-3.54	0.46	0.87	-0.3	1.01	0.1
14	US013	104	4.3	1.35	0.34	0.76	-0.8	0.78	-0.7
16	US014	104	4.6	0.71	0.33	0.61	-1.5	0.58	-1.6
17	US015	104	4.1	-0.06	0.39	0.41	-2.1	0.40	-2.1
18	US016	104	3.4	-1.11	0.35	0.72	-1.0	0.71	-1.0
19	US017	104	1.6	-3.30	0.41	2.29	3.6	2.30	3.6
21	US018	104	3.8	-0.61	0.35	1.41	1.3	1.41	1.3
22	US019	104	3.7	-0.52	0.35	0.59	-1.5	0.59	-1.5
23	US020	104	4.2	-0.56	0.35	1.03	0.1	1.06	0.3
24	US021	104	3.3	-1.17	0.35	2.08	3.0	2.09	3.0
26	US022	106	2.6	-0.98	0.35	0.67	-1.2	0.67	-1.2
28	US023	106	4.0	0.66	0.34	0.74	-0.8	0.75	-0.8
29	US024	106	3.1	0	0.35	0.72	-0.9	0.74	-0.8
30	US025	106	2.8	-0.82	0.35	1.10	0.4	1.10	0.4
31	US026	106	3.4	-0.06	0.35	0.83	-0.5	0.83	-0.4
33	US027	106	4.0	1.17	0.34	0.88	-0.3	0.94	-0.1
35	US028	106	3.3	0.10	0.35	0.68	-1.0	0.67	-1.1
36	US029	106	3.3	0.47	0.35	0.77	-0.7	0.76	-0.8
37	US030	106	3.2	-0.06	0.35	1.61	1.8	1.60	1.8
39	US031	106	3.5	-0.41	0.35	1.95	2.7	1.91	2.6
40	US032	106	3.1	-1.16	0.35	0.79	-0.7	0.77	-0.7

41	US033	106	3.1	0.20	0.35	0.50	-2.1	0.50	-2.1
42	US034	106	3.0	0.79	0.35	1.04	0.2	1.06	0.2
44	US035	106	4.3	0.84	0.33	1.42	1.4	1.39	1.3
45	US036	106	3.7	-0.04	0.34	0.84	-0.5	0.81	-0.6
46	US037	106	3.8	0.30	0.34	0.98	0	1.00	0
47	US038	106	3.3	0.66	0.35	0.66	-1.2	0.65	-1.2
48	US039	106	2.5	-0.77	0.35	1.08	0.3	1.11	0.4
49	US040	106	3.8	-0.04	0.34	0.85	-0.4	0.84	-0.5
50	US041	104	3.8	-0.33	0.35	0.78	-0.7	0.79	-0.6
51	US042	104	2.6	-1.22	0.35	0.45	-2.3	0.44	-2.4
52	US043	104	3.6	0.26	0.35	0.94	-0.1	0.95	0
53	US044	104	3.8	0.28	0.35	1.06	0.3	1.09	0.3
54	US045	104	4.0	0.33	0.35	0.61	-1.4	0.61	-1.4
55	US046	104	2.7	-1.16	0.35	0.62	-1.4	0.62	-1.4
57	US047	104	3.6	0.11	0.35	0.66	-1.2	0.64	-1.2
58	US048	104	3.4	-0.69	0.43	1.08	0.3	1.07	0.3
59	US049	104	2.6	-1.85	0.35	2.89	4.4	2.90	4.4
61	US050	104	3.0	-0.80	0.38	1.16	0.5	1.16	0.5
62	US051	104	4.5	0.94	0.33	1.11	0.4	1.13	0.5
63	US052	104	3.2	-0.80	0.35	1.99	2.8	2.00	2.8
64	US053	104	3.8	0.14	0.39	0.63	-1.2	0.62	-1.2
65	US054	104	3.6	-0.38	0.35	1.33	1.1	1.35	1.2
66	US055	104	3.0	-0.79	0.35	1.08	0.3	1.08	0.3
67	US056	104	2.9	-1.95	0.35	1.37	1.2	1.38	1.2
68	US057	104	3.8	-0.34	0.43	1.30	0.8	1.35	0.9
70	US058	104	2.5	-1.56	0.34	2.20	3.3	2.19	3.3
71	US059	104	3.3	-1.08	0.38	0.46	-2.0	0.46	-2.0
72	US060	104	2.4	-1.73	0.34	1.21	0.8	1.22	0.8
74	US061	104	2.4	-2.42	0.42	0.87	-0.2	0.87	-0.2
75	US062	106	3.7	-0.33	0.36	1.53	1.5	1.51	1.5
76	US063	106	3.9	0.68	0.43	0.73	-0.7	0.69	-0.8
77	US064	106	3.7	-0.50	0.36	0.64	-1.2	0.64	-1.2
79	US065	106	3.7	-0.10	0.35	1.28	0.9	1.30	1.0
80	US066	106	4.5	1.58	0.34	0.85	-0.4	0.84	-0.5
81	US067	106	3.9	1.27	0.34	0.80	-0.6	0.80	-0.6
82	US068	106	3.8	0.37	0.36	1.28	0.9	1.29	0.9
83	US069	106	3.5	-0.15	0.35	0.98	0	0.98	0
84	US070	106	2.4	-0.38	0.36	0.73	-0.9	0.73	-0.9

86	US071	106	3.8	0.22	0.35	1.46	1.4	1.46	1.4
87	US072	106	3.0	-0.19	0.35	1.25	0.8	1.26	0.9
88	US073	106	2.7	-0.73	0.35	1.39	1.2	1.39	1.2
89	US074	106	3.6	0.52	0.35	1.18	0.6	1.20	0.7
90	US075	106	4.3	1.61	0.34	1.36	1.2	1.38	1.2
91	US076	106	4.0	1.04	0.35	0.52	-1.9	0.50	-1.9
92	US077	106	2.5	0.23	0.38	0.67	-1.0	0.67	-1.1
93	US078	106	4.2	-0.02	0.35	1.12	0.5	1.13	0.5
94	US079	106	4.1	0.10	0.35	0.50	-1.9	0.49	-2.0
95	US080	106	2.5	-1.38	0.35	0.64	-1.4	0.64	-1.3
96	US081	106	3.5	0.08	0.35	0.89	-0.3	0.88	-0.3
97	US082	108	3.4	0.64	0.35	0.53	-1.8	0.53	-1.8
98	US083	108	3.1	-0.49	0.35	0.70	-1.0	0.70	-1.0
99	US084	108	1.5	-1.74	0.44	0.93	-0.1	0.91	-0.2
100	US085	108	3.0	-0.02	0.35	0.79	-0.6	0.79	-0.7
101	US086	108	2.8	-0.78	0.35	0.68	-1.1	0.68	-1.1
102	US087	108	3.0	-0.39	0.35	1.17	0.6	1.17	0.6
103	US088	108	3.0	-0.02	0.35	0.55	-1.7	0.54	-1.8
104	US089	108	3.0	0.49	0.35	0.57	-1.6	0.57	-1.6
105	US090	108	3.3	-0.42	0.35	0.58	-1.6	0.58	-1.6
106	US091	108	2.2	-0.44	0.35	1.06	0.3	1.07	0.3
107	US092	108	3.5	-0.33	0.35	0.97	0	0.98	0
108	US093	108	2.8	-0.63	0.35	1.58	1.8	1.60	1.8
109	US094	108	3.7	0.99	0.35	0.57	-1.6	0.57	-1.6
110	US095	108	3.3	0.52	0.35	1.41	1.3	1.39	1.2
111	US096	108	3.1	-0.44	0.35	1.15	0.6	1.14	0.5
112	US097	108	3.6	-0.28	0.35	0.85	-0.4	0.86	-0.4
113	US098	108	2.8	-0.33	0.34	0.63	-1.4	0.62	-1.4
114	US099	108	2.8	-1.16	0.35	0.92	-0.1	0.93	-0.1
115	US100	108	3.0	-0.20	0.35	0.79	-0.7	0.79	-0.6
116	US101	108	2.9	-0.39	0.35	2.00	2.8	1.99	2.8
117	US102	108	3.0	0.70	0.35	0.83	-0.5	0.81	-0.6
118	US103	108	2.6	-0.47	0.35	0.94	-0.1	0.93	-0.1
119	US104	108	3.6	0.23	0.35	0.93	-0.1	0.94	-0.1
120	US105	108	2.9	0.35	0.35	0.56	-1.7	0.56	-1.7
121	US106	108	3.6	0.97	0.35	0.70	-1.0	0.69	-1.1
122	US107	108	3.3	0.20	0.35	0.93	-0.1	0.93	-0.1
123	US108	108	3.7	0.84	0.35	0.70	-1.0	0.70	-1.0

124	US109	108	3.2	-0.36	0.35	0.84	-0.5	0.85	-0.4
125	US110	108	3.2	-0.04	0.38	1.22	0.7	1.22	0.7
126	US111	108	3.4	0.36	0.35	0.53	-1.7	0.53	-1.7
127	US112	108	3.3	0.70	0.35	1.12	0.4	1.12	0.4
128	US113	108	3.4	0.20	0.35	0.89	-0.2	0.90	-0.2
129	US114	108	3.1	0.40	0.35	0.88	-0.3	0.88	-0.3
130	US115	108	3.2	0.30	0.35	0.67	-1.2	0.67	-1.2
131	US116	108	3.1	0	0.35	0.86	-0.3	0.86	-0.3
132	US117	108	2.3	0.15	0.35	0.40	-2.7	0.40	-2.7
133	US118	108	2.3	-2.17	0.35	0.72	-1.0	0.71	-1.0
134	US119	108	3.0	0.25	0.35	0.49	-2.1	0.49	-2.1
135	US120	108	2.8	0.46	0.35	0.65	-1.2	0.65	-1.3
136	US121	108	2.1	-0.95	0.35	0.93	-0.1	0.94	-0.1
137	US122	108	2.6	0.22	0.35	0.45	-2.3	0.45	-2.3
138	US123	108	2.5	-1.05	0.38	0.31	-2.9	0.31	-2.9
139	US124	108	3.2	0.96	0.35	0.37	-2.7	0.37	-2.7
140	US125	108	2.3	-0.32	0.35	0.86	-0.4	0.87	-0.3
141	US126	108	2.5	-0.35	0.35	0.71	-1.0	0.71	-1.0
142	US127	108	2.7	0.82	0.35	0.76	-0.8	0.76	-0.8
143	US128	108	3.1	-0.13	0.35	0.32	-3.1	0.31	-3.2
144	US129	108	3.5	0.75	0.43	0.87	-0.2	0.88	-0.2
145	US130	108	2.4	-0.05	0.38	0.22	-3.6	0.22	-3.6
146	US131	108	4.6	1.34	0.43	0.56	-1.4	0.58	-1.2
147	US132	108	4.1	1.47	0.34	1.02	0.1	1.05	0.2
148	US133	108	4.8	1.54	0.33	1.25	0.9	1.24	0.9
149	US134	108	4.0	1.02	0.35	0.70	-1.0	0.69	-1.0
150	US135	108	5.3	2.82	0.35	1.91	2.8	2.04	3.1
151	US136	108	5.5	2.28	0.38	1.25	0.9	1.11	0.4
152	US137	108	4.9	1.94	0.34	1.58	1.9	1.62	2.0
153	US138	108	3.8	0.81	0.35	0.63	-1.3	0.63	-1.3
154	US139	108	4.5	1.58	0.34	1.35	1.2	1.30	1.0
155	US140	108	4.3	1.63	0.33	0.58	-1.7	0.57	-1.7
156	US141	108	4.1	1.41	0.34	0.55	-1.7	0.53	-1.8
157	US142	108	4.5	1.55	0.33	0.52	-2.0	0.53	-2.0
158	US143	108	4.6	1.12	0.34	0.83	-0.5	0.84	-0.5
158	US144	108	5.4	2.68	0.35	0.85	-0.5	0.89	-0.3
160	US145	108	4.6	1.54	0.34	0.93	-0.1	1.06	0.2
161	US146	108	4.5	1.31	0.33	0.61	-1.6	0.63	-1.5

162	US147	108	4.4	1.30	0.37	1.42	1.2	1.43	1.2
163	US148	108	5.0	2.00	0.33	0.77	-0.8	0.79	-0.7
164	US149	108	5.5	2.98	0.43	1.28	0.9	1.18	0.5
165	US150	108	5.4	2.60	0.36	0.51	-2.2	0.53	-2.0
166	US151	108	4.6	1.43	0.34	1.11	0.4	1.10	0.4
167	US152	108	5.0	1.72	0.37	1.05	0.2	1.05	0.2
168	US153	108	2.3	0.09	0.45	0.55	-1.4	0.63	-1.0
169	US154	108	3.3	-0.11	0.39	1.11	0.4	1.11	0.4
170	US155	108	3.0	0.09	0.35	0.85	-0.4	0.85	-0.4
171	US156	108	3.3	0.76	0.35	0.52	-1.9	0.52	-1.9
172	US157	108	3.4	0.77	0.35	0.71	-1.0	0.72	-0.9
173	US158	108	3.5	0.05	0.38	0.46	-1.9	0.47	-1.9
174	US159	108	3.4	-0.15	0.35	0.73	-0.9	0.73	-0.9
175	US160	108	2.4	-0.19	0.34	0.37	-2.9	0.36	-2.9
176	US161	108	3.4	0.13	0.38	1.78	2.0	1.79	2.0
177	US162	108	4.0	0.87	0.35	1.07	0.3	1.08	0.3
178	US163	108	4.1	1.36	0.39	0.79	-0.5	0.80	-0.4
179	US164	108	2.1	-0.71	0.35	0.83	-0.5	0.82	-0.5
180	US165	108	2.5	-0.25	0.36	0.95	0	0.93	-0.1
181	US166	108	1.8	-0.24	0.44	0.89	-0.2	0.91	-0.1
182	US167	108	3.1	0.84	0.34	0.62	-1.4	0.63	-1.4
183	US168	108	3.5	-0.61	0.35	0.75	-0.9	0.75	-0.9
184	US169	108	2.4	-0.44	0.38	0.81	-0.5	0.82	-0.5
185	US170	108	2.5	-0.71	0.35	0.69	-1.1	0.68	-1.1
186	US171	108	3.1	-0.48	0.35	1.00	0.1	1.00	0.1
187	US172	108	3.2	0.63	0.35	1.07	0.3	1.08	0.3
188	US173	108	3.5	1.03	0.35	0.88	-0.3	0.88	-0.3
189	US174	108	3.2	0.33	0.35	0.47	-2.1	0.47	-2.1
190	US175	108	4.0	1.06	0.38	0.63	-1.1	0.62	-1.2
191	US176	108	4.1	0.48	0.35	0.51	-1.8	0.52	-1.8
192	US177	108	3.8	1.16	0.35	0.87	-0.3	0.89	-0.2
193	US178	108	4.2	1.37	0.34	1.03	0.1	1.03	0.2
194	US179	108	4.0	2.12	0.38	0.90	-0.2	0.90	-0.2
195	US180	108	3.5	1.12	0.35	1.44	1.3	1.46	1.4
196	US181	108	3.3	0.29	0.35	1.01	0.1	1.02	0.1
197	US182	108	3.3	0.68	0.39	1.40	1.1	1.39	1.1
198	US183	108	5.0	1.92	0.34	1.45	1.5	1.43	1.5
199	US184	108	3.3	0.42	0.35	0.86	-0.4	0.85	-0.4

200	US185	108	3.6	0.76	0.38	0.86	-0.3	0.87	-0.3
201	US186	108	4.2	1.51	0.35	1.62	1.8	1.63	1.8
202	US187	108	3.4	0.11	0.35	0.66	-1.2	0.66	-1.2
203	US188	108	3.3	0.85	0.35	1.77	2.3	1.78	2.3
204	US189	108	3.9	0.81	0.34	1.51	1.6	1.49	1.6
205	US190	108	3.9	1.32	0.38	0.54	-1.6	0.52	-1.6
206	US191	108	3.6	0.66	0.35	0.70	-1.0	0.69	-1.0
207	US192	108	2.8	0.04	0.35	0.65	-1.3	0.65	-1.3
208	US193	108	3.0	0.47	0.35	0.81	-0.5	0.81	-0.5
209	US194	108	3.6	0.57	0.35	0.56	-1.6	0.56	-1.6
210	US195	108	3.6	-0.38	0.38	2.02	2.6	2.03	2.6
211	US196	108	4.2	0.90	0.35	1.12	0.4	1.13	0.5
212	US197	108	3.6	-0.46	0.35	0.96	0	0.97	0
213	US198	108	3.7	0.39	0.35	0.40	-2.5	0.39	-2.6
214	US199	108	3.5	0.17	0.35	0.83	-0.5	0.82	-0.5
215	US200	108	3.1	-0.34	0.35	0.41	-2.5	0.41	-2.5
216	US201	108	3.9	1.19	0.38	1.13	0.5	1.11	0.4
217	US202	108	3.4	-0.24	0.35	0.66	-1.2	0.66	-1.2
218	US203	108	3.9	0.80	0.35	0.98	0	0.99	0
219	US204	108	3.4	0.68	0.35	0.64	-1.3	0.64	-1.3
220	US205	108	4.0	0.45	0.35	0.73	-0.9	0.73	-0.9
221	US206	108	4.3	1.36	0.35	1.25	0.8	1.26	0.8
222	US207	108	4.1	2.37	0.35	1.15	0.5	1.15	0.5
223	US208	108	4.8	1.67	0.33	1.23	0.9	1.22	0.8
224	US209	108	4.0	0.13	0.35	1.10	0.4	1.11	0.4
225	US210	108	3.8	0.98	0.39	1.11	0.4	1.12	0.4
226	US211	104	4.0	-0.18	0.36	0.95	0	0.95	0
227	US212	104	3.5	-0.13	0.35	0.73	-0.9	0.73	-0.9
229	US213	104	4.1	0.47	0.42	0.55	-1.3	0.58	-1.2
230	US214	104	2.9	-1.98	0.34	2.06	2.9	2.06	2.9
231	US215	104	3.9	0.69	0.35	0.48	-2.1	0.46	-2.1
233	US216	104	3.7	0.59	0.35	0.63	-1.3	0.63	-1.3
234	US217	104	3.5	-0.42	0.35	1.22	0.8	1.21	0.7
235	US218	104	4.2	0.97	0.35	0.96	0	0.93	-0.1
237	US219	104	4.3	1.09	0.35	0.42	-2.2	0.41	-2.3
238	US220	104	3.6	-0.21	0.35	0.97	0	0.99	0
239	US221	104	2.5	-1.76	0.34	1.26	0.9	1.27	0.9
240	US222	104	4.1	0.01	0.34	0.64	-1.3	0.61	-1.4

241	US223	104	3.8	0.19	0.35	0.81	-0.5	0.80	-0.6
242	US224	104	2.4	-0.97	0.42	2.32	3.0	2.31	2.9
243	US225	104	3.3	-0.90	0.35	0.94	0	0.95	0
244	US226	104	4.3	0.32	0.34	0.73	-0.9	0.77	-0.7
245	US227	104	3.3	-0.59	0.38	0.74	-0.8	0.74	-0.8
247	US228	104	3.0	-1.12	0.35	0.83	-0.5	0.83	-0.5
248	US229	104	3.5	-0.29	0.42	1.07	0.3	1.07	0.3
249	US230	106	3.0	-0.56	0.35	1.28	0.9	1.29	1.0
250	US231	106	3.1	-0.51	0.35	0.64	-1.3	0.63	-1.3
251	US232	106	3.3	0.28	0.35	0.56	-1.6	0.55	-1.7
252	US233	106	2.7	-0.58	0.35	0.70	-1.1	0.70	-1.1
253	US234	106	3.0	-0.83	0.35	0.54	-1.8	0.54	-1.8
254	US235	106	2.7	0.21	0.35	0.83	-0.5	0.84	-0.5
256	US236	106	2.4	-0.73	0.35	0.84	-0.5	0.84	-0.5
258	US237	106	3.0	-1.18	0.35	1.63	1.9	1.62	1.9
260	US238	106	3.3	-0.26	0.35	0.85	-0.4	0.83	-0.5
261	US239	106	3.5	0.11	0.35	1.12	0.4	1.09	0.4
262	US240	106	3.1	0.34	0.35	0.57	-1.7	0.56	-1.7
263	US241	106	3.5	-0.04	0.35	1.52	1.6	1.52	1.6
265	US242	106	3.7	0.86	0.35	1.27	0.9	1.34	1.1
266	US243	106	3.8	0.52	0.35	1.17	0.6	1.18	0.6
268	US244	106	3.8	0.82	0.35	0.78	-0.7	0.78	-0.7
269	US245	106	2.4	-0.80	0.35	0.66	-1.2	0.67	-1.2
270	US246	106	2.9	-1.28	0.35	0.72	-0.9	0.73	-0.9
272	US247	106	3.7	0.61	0.35	1.03	0.2	1.05	0.2
273	US248	106	3.8	0.55	0.34	1.99	2.8	1.97	2.7
275	US249	106	3.7	0.23	0.35	1.13	0.5	1.13	0.5
276	US250	104	3.4	0.08	0.35	0.44	-2.4	0.43	-2.4
277	US251	104	3.2	-1.29	0.35	0.84	-0.4	0.85	-0.4
278	US252	104	3.4	-0.90	0.35	1.23	0.8	1.23	0.8
279	US253	104	3.1	-0.80	0.35	0.95	0	0.95	0
280	US254	104	3.9	0.99	0.35	0.53	-1.8	0.53	-1.8
281	US255	104	2.7	-1.10	0.35	0.55	-1.8	0.54	-1.8
282	US256	104	1.9	-2.49	0.39	1.40	1.2	1.39	1.2
284	US257	104	3.4	-1.87	0.35	0.83	-0.5	0.83	-0.5
285	US258	104	3.5	0.22	0.35	1.29	0.9	1.27	0.9
286	US259	104	4.4	0.63	0.34	0.55	-1.7	0.52	-1.9
287	US260	104	3.0	-1.42	0.35	1.69	2.0	1.69	2.0

288	US261	104	4.1	0.22	0.35	0.91	-0.1	0.92	-0.1
289	US262	104	2.2	-2.57	0.38	0.59	-1.5	0.58	-1.5
290	US263	104	2.1	-2.20	0.35	0.99	0	0.98	0
291	US264	104	2.9	-0.79	0.34	0.81	-0.6	0.83	-0.5
292	US265	104	3.9	-0.20	0.35	1.47	1.5	1.45	1.4
293	US266	104	2.4	-1.53	0.34	1.30	1.0	1.31	1.0
295	US267	104	2.5	-1.04	0.35	1.22	0.8	1.21	0.7
296	US268	104	2.2	-2.48	0.35	0.99	0	0.98	0
297	US269	104	5.0	1.38	0.32	1.86	2.7	1.88	2.7
298	US270	104	3.9	0.18	0.39	1.66	1.7	1.65	1.7
299	US271	104	3.2	-1.36	0.38	4.92	6.7	4.94	6.7
300	US272	104	4.5	-0.94	0.34	1.77	2.2	1.76	2.2
301	US273	104	4.1	0.90	0.35	0.38	-2.6	0.37	-2.6
303	US274	104	3.6	0.12	0.35	0.79	-0.6	0.79	-0.6
304	US275	104	4.7	1.28	0.32	0.45	-2.5	0.45	-2.5
305	US276	104	3.9	0.35	0.35	1.08	0.3	1.05	0.2
306	US277	104	4.4	1.04	0.34	0.71	-1.0	0.74	-0.9
307	US278	104	4.8	1.45	0.33	1.57	1.9	1.53	1.7
308	US279	104	3.5	0.97	0.35	0.51	-2.0	0.50	-2.0
309	US280	104	4.4	0.15	0.41	1.60	1.5	1.58	1.4
310	US281	104	4.2	-0.28	0.35	2.16	2.9	2.25	3.1
311	US282	104	4.1	0.69	0.35	0.41	-2.4	0.40	-2.4
312	US283	104	3.8	-0.57	0.35	1.83	2.2	1.85	2.3
313	US284	104	4.1	-0.39	0.35	1.25	0.8	1.27	0.8
314	US285	104	1.2	-6.12	0.61	1.88	1.7	3.49	2.2
316	US286	104	4.3	0.08	0.34	1.02	0.1	1.00	0
317	US287	104	4.3	0.52	0.34	0.82	-0.5	0.83	-0.5
318	US288	104	4.5	0.89	0.33	1.58	1.8	1.60	1.8
319	US289	104	3.0	-0.32	0.35	0.67	-1.1	0.68	-1.1
320	US290	104	3.4	-1.27	0.43	0.49	-1.6	0.49	-1.6
321	US291	106	3.3	0.24	0.35	1.44	1.4	1.44	1.4
322	US292	106	5.2	1.81	0.36	1.90	2.6	1.59	1.5
323	US293	106	4.3	1.26	0.33	0.73	-0.9	0.72	-1.0
324	US294	106	4.4	1.06	0.36	1.19	0.6	1.18	0.6
325	US295	106	3.1	-0.18	0.35	0.73	-0.9	0.74	-0.8
326	US296	106	4.2	0.67	0.34	1.32	1.1	1.30	1.0
327	US297	106	3.9	0.60	0.35	1.26	0.8	1.27	0.8
328	US298	106	4.7	0.91	0.33	1.19	0.7	1.18	0.7

329	US299	106	3.8	0.32	0.37	2.38	3.3	2.56	3.6
330	US300	106	4.4	1.91	0.34	0.92	-0.2	0.92	-0.1
331	US301	106	3.5	-0.50	0.35	1.23	0.8	1.21	0.7
332	US302	106	5.2	1.82	0.39	0.87	-0.3	0.77	-0.5
333	US303	106	4.1	0.04	0.34	0.76	-0.8	0.74	-0.9
334	US304	106	5.0	2.15	0.36	0.96	0	0.98	0
335	US305	106	4.5	0.90	0.37	0.89	-0.2	0.95	0
336	US306	106	3.0	-1.05	0.43	2.42	2.9	2.50	3.0
337	US307	110	4.8	2.20	0.33	0.50	-2.2	0.49	-2.2
338	US308	110	3.9	0.65	0.35	0.63	-1.3	0.61	-1.4
339	US309	110	4.1	1.62	0.34	0.51	-1.9	0.50	-1.9
340	US310	110	3.8	1.02	0.35	0.94	-0.1	0.92	-0.1
341	US311	110	4.9	1.45	0.34	0.92	-0.1	0.91	-0.2
342	US312	110	5.0	2.70	0.33	1.09	0.4	1.07	0.3
343	US313	110	4.7	2.25	0.33	0.76	-0.8	0.74	-0.9
344	US314	110	4.5	1.52	0.33	0.70	-1.1	0.68	-1.1
345	US315	110	5.0	1.93	0.34	1.08	0.3	1.06	0.3
346	US316	110	4.0	1.32	0.35	0.97	0	0.94	0
347	US317	110	4.4	1.30	0.35	1.12	0.4	1.08	0.3
348	US318	110	4.1	1.04	0.35	1.45	1.4	1.47	1.4
349	US319	110	4.8	1.76	0.33	0.82	-0.6	0.80	-0.7
350	US320	110	5.3	1.85	0.35	1.04	0.2	1.06	0.2
351	US321	110	3.1	0.26	0.35	0.55	-1.8	0.55	-1.8
352	US322	110	3.0	0.01	0.35	1.11	0.4	1.10	0.4
353	US323	110	2.8	0.07	0.35	1.13	0.5	1.12	0.5
354	US324	110	4.3	1.09	0.34	0.67	-1.2	0.67	-1.2
355	US325	110	3.3	0.07	0.35	0.50	-2.0	0.50	-2.1
356	US326	110	3.1	-0.41	0.38	1.33	1.0	1.32	1.0
357	US327	110	3.1	-0.74	0.35	1.23	0.8	1.24	0.8
358	US328	110	3.9	1.11	0.35	0.68	-1.1	0.68	-1.0
359	US329	110	3.0	-0.53	0.35	0.89	-0.3	0.89	-0.3
360	US330	110	3.2	0.67	0.39	0.06	-5.2	0.05	-5.3
361	US331	110	4.0	0.63	0.35	1.28	0.9	1.25	0.8
362	US332	110	2.6	-1.20	0.42	0.90	-0.1	0.90	-0.1
363	US333	110	3.3	0.68	0.35	0.79	-0.6	0.80	-0.6
364	US334	110	4.0	0.39	0.35	1.17	0.6	1.16	0.6
365	US335	110	4.9	1.40	0.33	1.12	0.5	1.11	0.5
366	US336	110	4.3	0.99	0.35	0.69	-1.0	0.70	-1.0

367	US337	110	3.4	0.61	0.35	0.66	-1.1	0.66	-1.2
368	US338	110	3.5	-0.12	0.35	0.71	-1.0	0.71	-1.0
369	US339	110	4.5	1.21	0.34	0.89	-0.3	0.90	-0.2
370	US340	110	3.5	-0.15	0.35	1.15	0.5	1.15	0.5
371	US341	110	3.3	-0.16	0.35	1.40	1.3	1.39	1.3
372	US342	110	4.8	2.27	0.36	0.97	0	0.99	0
373	US343	110	4.5	1.27	0.33	1.70	2.1	1.63	1.9
374	US344	110	4.8	1.66	0.33	1.51	1.7	1.58	1.9
375	US345	110	4.7	1.68	0.33	0.68	-1.2	0.71	-1.1
376	US346	110	4.1	0.69	0.35	1.05	0.2	1.06	0.2
377	US347	110	4.1	1.39	0.34	0.69	-1.1	0.70	-1.0
378	US348	110	4.1	1.26	0.35	0.68	-1.0	0.67	-1.1
379	US349	110	3.8	0.63	0.35	1.47	1.5	1.45	1.4
380	US350	110	4.0	1.16	0.35	0.80	-0.6	0.79	-0.6
381	US351	110	4.6	1.98	0.33	1.29	1.0	1.32	1.1
382	US352	110	3.4	-0.39	0.35	1.31	1.0	1.30	1.0
383	US353	110	4.8	2.19	0.33	1.43	1.4	1.44	1.5
384	US354	110	4.1	0.91	0.35	1.30	0.9	1.31	1.0
385	US355	110	3.9	0.48	0.35	1.19	0.7	1.20	0.7
386	US356	110	4.5	1.24	0.33	1.49	1.6	1.54	1.7
387	US357	110	4.1	0.86	0.34	1.88	2.5	1.92	2.6
388	US358	110	3.9	0.68	0.35	0.75	-0.7	0.75	-0.7
389	US359	110	4.1	0.42	0.34	1.27	0.9	1.28	0.9
390	US360	110	3.4	-0.97	0.35	1.64	1.9	1.68	1.9
391	US361	110	3.9	0.86	0.35	1.27	0.9	1.30	0.9
392	US362	110	3.3	-0.80	0.35	2.22	3.2	2.22	3.2
393	US363	110	3.6	1.02	0.43	0.86	-0.2	0.88	-0.2
394	US364	110	3.0	-0.44	0.35	0.86	-0.4	0.86	-0.4
395	US365	110	2.2	0.03	0.44	1.02	0.1	0.99	0
396	US366	110	3.9	0.29	0.35	2.80	4.0	2.75	3.9
397	US367	110	3.5	0.52	0.35	0.72	-1.0	0.72	-1.0
398	US368	110	4.3	1.49	0.34	1.00	0	1.02	0.1
399	US369	110	3.1	0.16	0.35	0.68	-1.1	0.70	-1.0
400	US370	110	3.6	0.48	0.35	1.44	1.3	1.42	1.3
401	US371	110	4.2	1.62	0.35	0.65	-1.2	0.62	-1.3
402	US372	110	3.2	0.68	0.34	1.00	0	0.98	0
403	US373	110	3.3	-0.32	0.35	0.88	-0.3	0.88	-0.3
404	US374	110	2.8	0.14	0.35	0.60	-1.6	0.59	-1.6

405	US375	110	3.0	-0.22	0.35	0.42	-2.4	0.42	-2.4
406	US376	110	2.6	-1.29	0.4	0.89	-0.2	0.81	-0.4
407	US377	110	3.3	0.68	0.35	0.70	-1.0	0.70	-1.0
408	US378	110	3.3	-0.11	0.35	0.83	-0.5	0.83	-0.4
409	US379	110	3.3	0.36	0.35	0.81	-0.5	0.82	-0.5
410	US380	110	4.3	1.79	0.34	1.16	0.6	1.12	0.4
411	US381	110	3.3	-0.32	0.35	0.67	-1.1	0.66	-1.2
412	US382	110	3.1	-0.22	0.35	0.71	-1.0	0.71	-1.0
413	US383	110	3.4	0.29	0.39	1.21	0.7	1.20	0.6
414	US384	110	3.7	0.85	0.35	0.53	-1.8	0.53	-1.8
415	US385	110	4.4	1.64	0.34	0.97	0	0.93	-0.1
416	US386	110	4.0	1.21	0.35	1.02	0.1	1.06	0.2
417	US387	110	4.6	1.71	0.33	1.60	1.9	1.60	1.9
418	US388	110	3.8	1.05	0.36	1.15	0.5	1.15	0.5
419	US389	110	4.6	2.14	0.33	0.94	-0.1	0.94	-0.1
420	US390	110	4.1	0.67	0.34	1.25	0.8	1.25	0.8
421	US391	110	4.1	0.51	0.34	0.94	-0.1	0.97	0
422	US392	110	4.2	1.03	0.34	1.13	0.5	1.12	0.4
423	US393	110	4.7	2.17	0.33	1.37	1.3	1.40	1.3
424	US394	110	3.6	0.46	0.35	1.14	0.5	1.14	0.5
425	US395	110	3.6	0.77	0.35	0.85	-0.3	0.87	-0.3
426	US396	110	5.3	2.38	0.37	0.78	-0.7	0.80	-0.6
427	US397	110	4.5	1.75	0.33	0.87	-0.4	0.87	-0.3
428	US398	110	4.0	0.11	0.35	0.77	-0.7	0.77	-0.6
429	US399	106	3.5	0.39	0.35	0.64	-1.3	0.63	-1.4
430	US400	106	3.7	0.89	0.35	0.96	0	0.98	0
431	US401	106	4.3	1.25	0.33	1.53	1.7	1.54	1.7
432	US402	106	3.8	1.24	0.34	0.93	-0.1	0.96	0
433	US403	106	3.3	-0.01	0.34	0.47	-2.2	0.47	-2.2
434	US404	106	2.2	-0.40	0.36	0.52	-2.0	0.51	-2.1
436	US405	106	3.7	0.67	0.35	1.28	0.9	1.33	1.1
437	US406	106	3.4	0.15	0.35	1.10	0.4	1.11	0.4
438	US407	106	2.9	-0.22	0.35	0.75	-0.8	0.75	-0.8
439	US408	106	4.1	0.61	0.34	0.48	-2.1	0.48	-2.0
440	US409	106	4.5	1.36	0.36	0.63	-1.3	0.65	-1.2
441	US410	106	5.2	2.38	0.37	1.12	0.4	1.00	0.1
442	US411	106	4.1	1.36	0.34	1.06	0.2	1.11	0.4
443	US412	106	3.4	-0.01	0.35	0.44	-2.3	0.43	-2.3

444	US413	106	2.8	-0.17	0.35	0.81	-0.5	0.81	-0.5
445	US414	106	2.9	-1.05	0.35	1.11	0.4	1.10	0.4
446	US415	106	3.5	-0.10	0.35	0.56	-1.6	0.56	-1.7
447	US416	106	4.2	1.20	0.34	0.57	-1.7	0.58	-1.7
448	US417	106	3.2	0.23	0.35	0.62	-1.4	0.61	-1.4
449	US418	104	3.1	-0.30	0.35	1.22	0.8	1.22	0.8
450	US419	104	2.8	-0.58	0.35	0.92	-0.1	0.92	-0.1
452	US420	104	2.2	-1.77	0.35	0.28	-3.5	0.27	-3.6
454	US421	104	3.6	-0.12	0.35	1.26	0.9	1.27	0.9
455	US422	104	3.8	-0.28	0.35	1.15	0.6	1.13	0.5
456	US423	104	3.1	0.80	0.35	1.20	0.7	1.18	0.6
457	US424	104	3.5	0.46	0.35	0.54	-1.8	0.54	-1.9
458	US425	104	3.9	0.24	0.35	0.46	-2.0	0.46	-2.1
459	US426	104	3.6	0.11	0.35	0.99	0	0.98	0
460	US427	104	3.3	-0.57	0.35	1.03	0.2	1.01	0.1
462	US428	104	4.3	0.47	0.34	0.87	-0.3	0.84	-0.4
463	US429	104	2.9	-0.70	0.35	0.46	-2.2	0.46	-2.2
465	US430	104	3.0	-0.78	0.35	0.93	-0.1	0.93	-0.1
466	US431	104	2.6	-0.89	0.35	0.61	-1.5	0.61	-1.5
467	US432	104	2.2	-1.48	0.42	1.05	0.2	1.06	0.2
468	US433	104	3.7	-0.11	0.35	1.47	1.5	1.51	1.6
469	US434	104	3.4	-0.66	0.39	0.77	-0.6	0.77	-0.6
470	US435	104	3.0	-1.12	0.35	1.42	1.3	1.42	1.3
471	US436	104	4.3	0.08	0.34	1.00	0.1	0.99	0
473	US437	104	2.3	-1.73	0.35	1.28	0.9	1.27	0.9
474	US438	106	3.5	0.62	0.35	0.98	0	0.98	0
475	US439	106	3.3	0.35	0.35	0.81	-0.6	0.80	-0.6
478	US440	106	5.0	2.03	0.36	0.94	-0.1	0.98	0
479	US441	106	3.6	0.66	0.35	0.70	-1.0	0.71	-0.9
480	US442	106	4.4	0.37	0.35	1.14	0.5	1.09	0.3
481	US443	106	4.3	1.59	0.34	0.78	-0.7	0.78	-0.7
483	US444	106	4.8	1.67	0.33	1.32	1.1	1.28	1.0
484	US445	106	3.4	0.96	0.35	0.90	-0.2	0.89	-0.2
485	US446	106	4.3	1.08	0.34	0.99	0	0.98	0
486	US447	106	4.5	1.14	0.33	1.63	2.0	1.61	1.9
487	US448	106	2.5	-1.39	0.35	1.16	0.6	1.15	0.5
489	US449	106	4.7	1.29	0.34	1.06	0.3	1.11	0.4
490	US450	106	3.2	0.10	0.35	0.69	-1.1	0.68	-1.1

492	US451	106	4.0	0.36	0.34	2.22	3.1	2.23	3.1
493	US452	106	4.5	0.96	0.33	1.08	0.3	1.06	0.3
494	US453	106	4.0	1.57	0.35	0.72	-0.9	0.75	-0.8
495	US454	106	4.0	-0.44	0.35	1.33	1.0	1.31	1.0
497	US455	106	2.9	-0.18	0.35	0.95	0	0.96	0
498	US456	106	3.4	-0.25	0.35	1.69	2.0	1.73	2.1
499	US457	106	3.7	0.35	0.38	1.05	0.2	1.07	0.3
500	US458	106	3.9	0.69	0.35	1.05	0.2	1.04	0.2
501	US459	106	2.8	0.12	0.35	1.03	0.1	1.02	0.1
503	US460	106	2.8	-0.32	0.36	1.19	0.7	1.18	0.6
505	US461	106	3.3	0.92	0.35	0.92	-0.1	0.94	-0.1
506	US462	106	3.4	-0.46	0.35	1.27	0.9	1.28	0.9
508	US463	106	3.5	0.25	0.35	1.45	1.4	1.43	1.3
510	US464	106	3.9	1.05	0.34	1.55	1.7	1.52	1.6
511	US465	106	2.3	-2.69	0.37	3.30	5.1	3.21	4.6
512	US466	106	4.0	1.07	0.34	0.93	-0.1	0.95	0
514	US467	106	3.6	-0.27	0.35	0.83	-0.5	0.83	-0.5
515	US468	106	3.8	0.62	0.35	0.89	-0.2	0.91	-0.2
516	US469	106	3.5	-0.42	0.35	0.94	0	0.93	-0.1
517	US470	106	3.8	1.29	0.35	0.46	-2.2	0.45	-2.2
519	US471	106	4.2	1.14	0.34	1.01	0.1	1.01	0.1
520	US472	106	3.5	0.75	0.35	0.81	-0.5	0.80	-0.6
521	US473	112	4.0	0.63	0.34	0.55	-1.7	0.55	-1.7
522	US474	112	3.9	0.75	0.34	0.61	-1.4	0.61	-1.4
523	US475	112	3.8	1.12	0.35	0.72	-0.9	0.72	-0.9
524	US476	112	3.9	1.01	0.35	0.83	-0.5	0.80	-0.6
525	US477	112	4.1	1.55	0.34	0.95	0	0.91	-0.2
526	US478	112	4.1	1.21	0.35	1.34	1.1	1.36	1.1
527	US479	112	4.2	1.39	0.34	0.46	-2.2	0.46	-2.2
528	US480	112	3.6	0.21	0.35	0.88	-0.3	0.86	-0.4
529	US481	112	4.3	0.84	0.34	0.76	-0.8	0.74	-0.9
530	US482	112	4.1	1.30	0.34	0.90	-0.2	0.88	-0.3
531	US483	112	4.5	1.03	0.34	1.14	0.5	1.09	0.3
532	US484	112	3.7	0.81	0.35	1.01	0.1	1.01	0.1
533	US485	112	3.8	1.07	0.35	0.88	-0.3	0.87	-0.3
534	US486	112	4.5	1.44	0.34	0.77	-0.7	0.74	-0.9
535	US487	112	3.8	0.48	0.39	0.86	-0.3	0.86	-0.3
536	US488	112	4.2	1.06	0.34	0.79	-0.6	0.75	-0.8

537	US489	112	4.7	2.23	0.33	0.90	-0.2	0.87	-0.4
538	US490	112	3.8	0.70	0.35	1.32	1.0	1.30	1.0
539	US491	112	4.0	0.96	0.34	1.31	1.0	1.34	1.1
540	US492	112	2.8	0.31	0.35	0.97	0	0.97	0
541	US493	112	4.1	0.98	0.34	1.13	0.5	1.12	0.5
542	US494	112	3.4	-0.42	0.35	1.03	0.1	1.04	0.2
543	US495	112	3.5	0.90	0.35	0.83	-0.5	0.82	-0.5
544	US496	112	3.3	0.27	0.35	0.93	-0.1	0.93	-0.1
545	US497	112	3.5	0.12	0.35	0.75	-0.8	0.75	-0.8
546	US498	112	3.2	0.54	0.35	0.69	-1.1	0.69	-1.1
547	US499	112	2.4	-0.27	0.35	1.03	0.1	1.03	0.1
548	US500	112	3.3	0.01	0.35	0.69	-1.1	0.70	-1.0
549	US501	112	3.9	-0.14	0.35	1.44	1.4	1.45	1.4
550	US502	112	2.6	0.09	0.35	0.43	-2.4	0.42	-2.5
551	US503	112	3.6	1.06	0.35	0.69	-1.0	0.69	-1.1
552	US504	112	2.1	-0.79	0.35	0.52	-2.0	0.51	-2.0
553	US505	112	2.3	-0.20	0.38	1.04	0.2	1.02	0.1
554	US506	112	2.8	-0.49	0.35	1.36	1.2	1.37	1.2
555	US507	112	3.3	0.23	0.35	0.56	-1.6	0.55	-1.7
556	US508	112	3.4	0.34	0.35	0.79	-0.7	0.80	-0.6
557	US509	112	3.4	0.34	0.35	0.92	-0.1	0.91	-0.2
558	US510	112	3.8	0.79	0.35	0.76	-0.8	0.76	-0.8
559	US511	112	2.9	0.24	0.38	0.94	0	0.94	0
560	US512	112	3.5	0.92	0.35	0.97	0	0.97	0
561	US513	112	3.7	-0.05	0.34	1.62	1.9	1.61	1.8
562	US514	112	3.4	1.23	0.35	0.77	-0.7	0.78	-0.6
563	US515	112	2.9	0.36	0.35	0.66	-1.3	0.65	-1.3
564	US516	112	3.1	0.58	0.35	1.16	0.6	1.15	0.5
565	US517	112	5.1	2.58	0.35	0.78	-0.7	0.85	-0.5
566	US518	112	3.8	1.45	0.35	1.28	0.9	1.28	0.9
567	US519	112	3.8	1.24	0.35	0.79	-0.6	0.80	-0.6
568	US520	112	3.3	0.86	0.35	0.58	-1.6	0.57	-1.6
569	US521	112	4.4	1.39	0.34	1.01	0.1	0.98	0
570	US522	112	4.1	1.13	0.35	0.51	-1.9	0.51	-1.9
571	US523	112	5.6	3.81	0.42	0.61	-1.3	0.80	-0.5
572	US524	112	4.3	1.69	0.34	0.72	-0.9	0.72	-0.9
573	US525	112	4.5	2.11	0.33	0.73	-0.9	0.74	-0.9
574	US526	112	4.3	1.13	0.34	0.72	-1.0	0.70	-1.1

575	US527	112	4.2	1.35	0.34	0.93	-0.1	0.93	-0.1
576	US528	112	4.3	1.68	0.34	0.82	-0.5	0.82	-0.5
577	US529	112	3.8	0.91	0.35	0.56	-1.7	0.55	-1.7
578	US530	112	3.9	0.13	0.35	0.98	0	1.00	0
579	US531	112	4.6	1.86	0.33	1.04	0.2	1.05	0.2
580	US532	112	4.1	1.71	0.35	0.98	0	0.98	0
581	US533	112	5.2	3.00	0.35	1.12	0.5	1.18	0.6
582	US534	112	4.3	0.98	0.35	1.20	0.7	1.15	0.5
583	US535	112	3.8	0.71	0.36	0.94	0	0.95	0
584	US536	112	4.2	0.82	0.34	1.05	0.2	1.04	0.2
585	US537	112	4.0	1.23	0.35	0.93	-0.1	0.96	0
586	US538	112	3.5	0.75	0.35	0.75	-0.8	0.74	-0.8
587	US539	112	4.9	1.59	0.33	1.51	1.7	1.49	1.7
588	US540	112	5.2	2.77	0.33	1.28	1.0	1.31	1.1
589	US541	112	4.0	1.41	0.34	0.93	-0.1	0.96	0
590	US542	112	3.9	1.00	0.35	0.70	-1.0	0.71	-1.0
591	US543	112	4.4	1.75	0.34	0.67	-1.2	0.68	-1.1
592	US544	112	4.6	1.80	0.33	0.81	-0.6	0.84	-0.5
593	US545	112	4.1	0.43	0.34	0.82	-0.5	0.85	-0.4
594	US546	112	4.1	1.15	0.35	0.62	-1.3	0.61	-1.3
595	US547	112	3.8	0.49	0.35	0.58	-1.5	0.58	-1.5
596	US548	112	4.3	0.53	0.35	0.47	-2.1	0.46	-2.1
597	US549	112	3.3	-0.01	0.35	0.38	-2.6	0.37	-2.7
598	US550	112	3.9	0.75	0.39	0.89	-0.2	0.90	-0.2
599	US551	112	3.5	-0.09	0.35	1.28	0.9	1.26	0.9
600	US552	112	4.5	1.78	0.34	0.68	-1.2	0.71	-1.0
601	US553	112	4.4	2.02	0.33	0.80	-0.6	0.81	-0.6
602	US554	112	3.6	0.57	0.35	0.90	-0.2	0.91	-0.2
603	US555	112	4.9	2.05	0.35	0.54	-1.8	0.54	-1.8
604	US556	112	3.9	0.74	0.35	0.52	-1.8	0.52	-1.8
605	US557	112	3.5	0.83	0.35	0.67	-1.2	0.66	-1.2
606	US558	112	2.5	-0.81	0.35	1.02	0.1	1.04	0.2
607	US559	112	3.1	0.66	0.35	0.67	-1.1	0.67	-1.1
608	US560	112	3.3	-0.20	0.35	0.44	-2.2	0.44	-2.2
609	US561	112	2.2	-0.61	0.35	0.45	-2.4	0.45	-2.4
610	US562	112	3.3	0.73	0.35	0.76	-0.7	0.75	-0.8
611	US563	112	2.5	-0.44	0.35	0.85	-0.4	0.85	-0.4
612	US564	112	3.5	1.24	0.35	0.64	-1.3	0.64	-1.3

613	US565	112	3.9	0.87	0.35	0.75	-0.8	0.76	-0.7
614	US566	112	3.2	0.04	0.35	0.70	-1.0	0.69	-1.0
615	US567	112	3.8	0.38	0.35	1.25	0.8	1.24	0.8
616	US568	112	2.3	-1.21	0.35	1.25	0.9	1.25	0.9
617	US569	112	2.9	-0.46	0.35	0.85	-0.4	0.85	-0.4
618	US570	112	2.5	-1.06	0.35	1.08	0.3	1.08	0.3
619	US571	104	4.7	1.45	0.33	0.90	-0.3	0.90	-0.2
620	US572	104	3.2	-0.89	0.35	3.21	5.1	3.19	5.1
621	US573	104	4.5	0.71	0.33	1.79	2.3	1.87	2.5
622	US574	104	3.5	-0.12	0.35	0.92	-0.1	0.93	-0.1
623	US575	104	3.8	-0.04	0.36	0.70	-1.0	0.70	-1.0
624	US576	104	4.0	0.61	0.35	0.74	-0.8	0.72	-0.9
625	US577	104	3.0	-1.77	0.35	4.59	6.9	4.62	6.9
627	US578	104	5.0	1.95	0.37	0.94	0	0.93	-0.1
628	US579	104	3.8	-1.11	0.35	8.56	9.0	7.79	9.0
629	US580	104	3.5	-0.15	0.35	1.34	1.1	1.35	1.1
630	US581	104	3.1	-1.41	0.35	2.42	3.5	2.42	3.5
631	US582	104	3.9	0.30	0.38	0.70	-0.9	0.69	-1.0
632	US583	104	4.0	0.88	0.35	0.86	-0.3	0.84	-0.4
633	US584	104	3.8	-0.37	0.35	0.53	-1.9	0.53	-1.9
634	US585	104	3.5	-2.16	0.34	1.06	0.3	1.09	0.4
635	US586	104	2.9	-0.12	0.35	1.43	1.4	1.46	1.4
637	US587	104	2.7	-1.37	0.34	0.89	-0.3	0.88	-0.3
638	US588	104	3.5	1.17	0.39	0.47	-1.9	0.47	-1.9
639	US589	104	3.5	-0.81	0.35	0.75	-0.8	0.75	-0.8

Separation: 2.80 Reliability: 0.89

Fixed (all same) chi-square: 4912.1 significance *p*: .00

Note: 104 = Grade 4 US student, 106 = Grade 6 US student, 108 = Grade 8 US student, 110 = Grade 10 US student, 112 = Grade 12 US student

Calibration of the students for US student essays using US raters, Lexile Analyzer and IEA (76 misfitting students)

Student			Observed	Measure	S.E.	Infit		Outfit	
	Number	Group	score	(Logit)		MnSq	ZStd	MnSq	ZStd
21	US018	104	3.8	-0.61	0.35	1.41	1.3	1.41	1.3
110	US095	108	3.3	0.52	0.35	1.41	1.3	1.39	1.2
44	US435	106	4.3	0.84	0.33	1.42	1.4	1.39	1.3
162	US035	108	4.4	1.30	0.37	1.42	1.2	1.43	1.2

470	US147	104	3.0	-1.12	0.35	1.42	1.3	1.42	1.3
383	US586	110	4.8	2.19	0.33	1.43	1.4	1.44	1.5
635	US353	104	2.9	-0.12	0.35	1.43	1.4	1.46	1.4
195	US291	108	3.5	1.12	0.35	1.44	1.3	1.46	1.4
321	US180	106	3.3	0.24	0.35	1.44	1.4	1.44	1.4
400	US370	110	3.6	0.48	0.35	1.44	1.3	1.42	1.3
549	US501	112	3.9	-0.14	0.35	1.44	1.4	1.45	1.4
198	US463	108	5.0	1.92	0.34	1.45	1.5	1.43	1.5
348	US183	110	4.1	1.04	0.35	1.45	1.4	1.47	1.4
508	US318	106	3.5	0.25	0.35	1.45	1.4	1.43	1.3
86	US071	106	3.8	0.22	0.35	1.46	1.4	1.46	1.4
292	US265	104	3.9	-0.20	0.35	1.47	1.5	1.45	1.4
379	US433	110	3.8	0.63	0.35	1.47	1.5	1.45	1.4
468	US349	104	3.7	-0.11	0.35	1.47	1.5	1.51	1.6
386	US356	110	4.5	1.24	0.33	1.49	1.6	1.54	1.7
5	US005	104	3.1	-1.99	0.35	1.51	1.6	1.52	1.7
204	US189	108	3.9	0.81	0.34	1.51	1.6	1.49	1.6
374	US344	110	4.8	1.66	0.33	1.51	1.7	1.58	1.9
587	US539	112	4.9	1.59	0.33	1.51	1.7	1.49	1.7
263	US241	106	3.5	-0.04	0.35	1.52	1.6	1.52	1.6
75	US062	106	3.7	-0.33	0.36	1.53	1.5	1.51	1.5
431	US401	106	4.3	1.25	0.33	1.53	1.7	1.54	1.7
510	US464	106	3.9	1.05	0.34	1.55	1.7	1.52	1.6
307	US278	104	4.8	1.45	0.33	1.57	1.9	1.53	1.7
108	US288	108	2.8	-0.63	0.35	1.58	1.8	1.60	1.8
152	US093	108	4.9	1.94	0.34	1.58	1.9	1.62	2.0
318	US137	104	4.5	0.89	0.33	1.58	1.8	1.60	1.8
309	US280	104	4.4	0.15	0.41	1.60	1.5	1.58	1.4
417	US387	110	4.6	1.71	0.33	1.60	1.9	1.60	1.9
37	US030	106	3.2	-0.06	0.35	1.61	1.8	1.60	1.8
201	US186	108	4.2	1.51	0.35	1.62	1.8	1.63	1.8
561	US513	112	3.7	-0.05	0.34	1.62	1.9	1.61	1.8
258	US237	106	3.0	-1.18	0.35	1.63	1.9	1.62	1.9
486	US447	106	4.5	1.14	0.33	1.63	2.0	1.61	1.9
390	US360	110	3.4	-0.97	0.35	1.64	1.9	1.68	1.9
298	US270	104	3.9	0.18	0.39	1.66	1.7	1.65	1.7
287	US260	104	3.0	-1.42	0.35	1.69	2.0	1.69	2.0
498	US456	106	3.4	-0.25	0.35	1.69	2.0	1.73	2.1

373	US343	110	4.5	1.27	0.33	1.70	2.1	1.63	1.9
203	US272	108	3.3	0.85	0.35	1.77	2.3	1.78	2.3
300	US188	104	4.5	-0.94	0.34	1.77	2.2	1.76	2.2
176	US161	108	3.4	0.13	0.38	1.78	2.0	1.79	2.0
621	US573	104	4.5	0.71	0.33	1.79	2.3	1.87	2.5
312	US283	104	3.8	-0.57	0.35	1.83	2.2	1.85	2.3
297	US269	104	5.0	1.38	0.32	1.86	2.7	1.88	2.7
314	US285	104	1.2	-6.12	0.61	1.88	1.7	3.49	2.2
387	US357	110	4.1	0.86	0.34	1.88	2.5	1.92	2.6
322	US292	106	5.2	1.81	0.36	1.90	2.6	1.59	1.5
150	US135	108	5.3	2.82	0.35	1.91	2.8	2.04	3.1
39	US031	106	3.5	-0.41	0.35	1.95	2.7	1.91	2.6
63	US052	104	3.2	-0.80	0.35	1.99	2.8	2.00	2.8
273	US248	106	3.8	0.55	0.34	1.99	2.8	1.97	2.7
116	US101	108	2.9	-0.39	0.35	2.00	2.8	1.99	2.8
210	US195	108	3.6	-0.38	0.38	2.02	2.6	2.03	2.6
230	US214	104	2.9	-1.98	0.34	2.06	2.9	2.06	2.9
24	US021	104	3.3	-1.17	0.35	2.08	3.0	2.09	3.0
310	US281	104	4.2	-0.28	0.35	2.16	2.9	2.25	3.1
70	US058	104	2.5	-1.56	0.34	2.20	3.3	2.19	3.3
392	US451	110	3.3	-0.80	0.35	2.22	3.2	2.22	3.2
492	US362	106	4.0	0.36	0.34	2.22	3.1	2.23	3.1
19	US017	104	1.6	-3.30	0.41	2.29	3.6	2.30	3.6
242	US224	104	2.4	-0.97	0.42	2.32	3.0	2.31	2.9
329	US299	106	3.8	0.32	0.37	2.38	3.3	2.56	3.6
336	US581	106	3.0	-1.05	0.43	2.42	2.9	2.50	3.0
630	US306	104	3.1	-1.41	0.35	2.42	3.5	2.42	3.5
396	US366	110	3.9	0.29	0.35	2.80	4.0	2.75	3.9
59	US049	104	2.6	-1.85	0.35	2.89	4.4	2.90	4.4
620	US572	104	3.2	-0.89	0.35	3.21	5.1	3.19	5.1
511	US465	106	2.3	-2.69	0.37	3.30	5.1	3.21	4.6
625	US577	104	3.0	-1.77	0.35	4.59	6.9	4.62	6.9
299	US271	104	3.2	-1.36	0.38	4.92	6.7	4.94	6.7
628	US579	104	3.8	-1.11	0.35	8.56	9.0	7.79	9.0

Separation: 2.80 Reliability: 0.89

Fixed (all same) chi-square: 4912.1 significance *p*: .00

Note: 104 = Grade 4 US student, 106 = Grade 6 US student, 108 = Grade 8 US student, 110 = Grade 10 US student, 112 = Grade 12 US student

Appendix 13

```
+---------------------------------------+
|Measr|-Prompt|-HK Rater|+Student  |R6SCA|
|-----+-------+---------+----------+-----|
|   7 +       +         +          + (6) |
|     |       |         |          |     |
|     |       |         |          |     |
|     |       |         | .        |     |
|   6 +       +         +          +     |
|     |       |         |          |     |
|     |       |         |          |     |
|     |       |         |          |     |
|   5 +       +         +          +     |
|     |       |         |          |     |
|     |       |         |          | --- |
|     |       |         | *        |     |
|   4 +       +         + .        +     |
|     |       |         | .        |     |
|     |       |         | *        |  5  |
|     |       |         | *.       |     |
|   3 +       +         + *.       +     |
|     |       |         | .        |     |
|     |       |         |          | --- |
|     |       |         |          |     |
|   2 +       +         + ****.    +     |
|     |       |         | ****.    |  4  |
|     |       |         | *.       |     |
|     |       |         | *****.   |     |
|   1 +       +         + ****     + --- |
|     |       | HK004   | ***.     |     |
|     |       |         | *****    |     |
|     | 4-4BI | HK001   | ******   |     |
*   0 * 5-4BN * HK002   * ******.  *  3  *
|     | 6-4BP |         | **.      |     |
|     |       |         | *.       |     |
|     |       |         | **.      | --- |
|  -1 +       + HK003   + *        +     |
|     |       |         | **.      |     |
|     |       |         | *        |     |
|     |       |         | **.      |  2  |
|  -2 +       +         + *.       +     |
|     |       |         | .        |     |
|     |       |         | .        | --- |
|     |       |         | *        |     |
|  -3 +       +         + .        +     |
|     |       |         | *        |     |
|     |       |         | *        |  1  |
|     |       |         | *        |     |
|  -4 +       +         +          + (0) |
|-----+-------+---------+----------+-----|
|Measr|-Prompt|-HK Rater| * = 2    |R6SCA|
+---------------------------------------+
```

Vertical map for the HK student essays with scoring of HK raters

Calibration of the HK students for HK student essays using HK raters

Student Number		Observed score	Measure (Logit)	S.E.	Infit MnSq	Infit ZStd	Outfit MnSq	Outfit ZStd
640	HK001	1.0	-3.57	0.74	2.19	1.5	2.01	1.4
641	HK002	3.2	0.51	0.53	0.46	-0.9	0.46	-0.9
642	HK003	3.8	1.36	0.52	0.59	-0.6	0.59	-0.6
643	HK004	3.3	0.77	0.52	0.50	-0.8	0.51	-0.8

644	HK005	1.3	-3.21	0.71	0.21	-1.5	0.19	-1.6
645	HK006	3.5	1.05	0.52	0.50	-0.9	0.50	-0.9
646	HK007	5.3	4.14	0.65	0.12	-2.4	0.16	-2.1
647	HK008	3.2	0.51	0.53	0.44	-1.0	0.45	-1.0
648	HK009	3.7	1.09	0.52	1.69	1.2	1.70	1.2
649	HK010	4.8	3.28	0.56	0.78	-0.2	0.81	-0.1
650	HK011	2.3	-1.15	0.54	1.63	1.1	1.60	1.0
651	HK012	4.7	2.98	0.55	0.44	-1.1	0.42	-1.1
652	HK013	3.7	1.10	0.52	0.48	-0.9	0.48	-0.9
653	HK014	4.2	2.12	0.52	1.17	0.4	1.20	0.5
654	HK015	5.3	4.13	0.65	0.52	-0.8	0.70	-0.3
655	HK016	4.2	2.12	0.52	0.11	-2.6	0.11	-2.6
656	HK017	3.0	0	0.53	0.24	-1.7	0.24	-1.7
657	HK018	3.3	0.77	0.52	0.28	-1.6	0.27	-1.6
658	HK019	3.8	1.37	0.52	1.49	0.9	1.48	0.9
659	HK020	3.0	0.23	0.53	0.51	-0.8	0.50	-0.8
660	HK021	3.8	1.36	0.52	4.68	3.8	4.68	3.7
661	HK022	1.3	-2.79	0.58	0.98	0.1	0.97	0.1
662	HK023	4.0	1.63	0.52	1.41	0.8	1.41	0.8
663	HK024	4.8	3.29	0.56	0.77	-0.2	0.72	-0.3
664	HK025	4.7	2.95	0.54	0.78	-0.2	0.81	-0.1
665	HK026	3.7	1.09	0.52	0.18	-2.1	0.18	-2.1
666	HK027	2.8	-0.06	0.53	1.08	0.3	1.04	0.2
667	HK028	3.2	0.28	0.53	0.26	-1.6	0.26	-1.6
668	HK029	3.0	0.22	0.53	0.56	-0.7	0.57	-0.6
669	HK030	4.2	1.91	0.52	0.74	-0.3	0.74	-0.3
670	HK031	3.2	0.51	0.53	0.52	-0.8	0.53	-0.8
671	HK032	4.8	3.06	0.56	0.76	-0.2	0.76	-0.3
672	HK033	3.0	0.22	0.53	0.32	-1.4	0.32	-1.4
673	HK034	4.2	1.90	0.52	0.87	0	0.87	0
674	HK035	3.3	0.77	0.52	0.28	-1.6	0.27	-1.6
675	HK036	3.2	0.28	0.53	7.49	5.1	7.55	5.2
676	HK037	3.0	0.40	0.65	0.62	-0.3	0.63	-0.3
677	HK038	3.7	1.09	0.52	0.32	-1.4	0.32	-1.4
678	HK039	3.0	0.22	0.53	0.71	-0.3	0.68	-0.4
679	HK040	4.0	1.63	0.52	0.31	-1.5	0.32	-1.5
680	HK041	2.8	-0.06	0.53	0.70	-0.3	0.70	-0.3
681	HK042	4.2	1.91	0.52	0.66	-0.5	0.64	-0.5

682	HK043	3.0	0.23	0.53	0.22	-1.8	0.22	-1.8
683	HK044	3.8	1.36	0.52	2.05	1.6	2.06	1.6
684	HK045	3.2	0.50	0.53	0.32	-1.4	0.32	-1.4
685	HK046	2.8	-0.28	0.53	0.41	-1.1	0.41	-1.1
686	HK047	3.5	1.05	0.52	0.20	-2.0	0.20	-2.0
687	HK048	5.2	3.75	0.61	2.23	1.7	2.64	2.2
688	HK049	3.7	1.32	0.52	1.53	1.0	1.53	1.0
689	HK050	3.5	1.05	0.64	0.60	-0.4	0.59	-0.4
690	HK051	2.8	-0.06	0.53	0.61	-0.5	0.62	-0.5
691	HK052	3.2	0.28	0.53	0.26	-1.6	0.26	-1.6
692	HK053	2.8	-0.06	0.53	0.70	-0.3	0.70	-0.3
693	HK054	3.8	1.37	0.52	1.57	1.0	1.60	1.0
694	HK055	3.8	1.58	0.52	1.41	0.8	1.41	0.8
695	HK056	2.5	-0.86	0.54	0.87	0	0.87	0
696	HK057	2.0	-1.53	0.55	0.24	-1.7	0.24	-1.7
697	HK058	3.3	0.55	0.52	1.56	1.0	1.53	0.9
698	HK059	3.3	0.77	0.52	1.59	1.0	1.60	1.0
699	HK060	2.7	-0.57	0.54	0.34	-1.3	0.33	-1.4
700	HK061	2.8	-0.06	0.53	0.35	-1.3	0.35	-1.3
701	HK062	2.7	-0.57	0.54	2.25	1.8	2.17	1.7
702	HK063	4.2	2.12	0.52	1.63	1.1	1.64	1.1
703	HK064	2.5	-0.86	0.54	1.32	0.6	1.30	0.6
704	HK065	4.0	1.86	0.52	1.57	1.0	1.66	1.1
705	HK066	3.2	0.28	0.53	0.73	-0.3	0.71	-0.3
706	HK067	4.0	1.85	0.52	0.67	-0.4	0.66	-0.5
707	HK068	2.8	-0.28	0.53	0.53	-0.7	0.52	-0.8
708	HK069	2.8	-0.06	0.53	1.02	0.2	1.03	0.2
709	HK070	4.0	1.63	0.52	0.61	-0.6	0.61	-0.6
710	HK071	2.3	-0.94	0.54	0.62	-0.5	0.60	-0.6
711	HK072	1.8	-2.08	0.68	2.38	1.6	2.35	1.6
712	HK073	2.8	-0.06	0.53	1.88	1.4	1.91	1.4
713	HK074	2.0	-1.75	0.55	1.61	1.0	1.58	1.0
714	HK075	4.2	2.12	0.52	0.96	0.1	0.95	0
715	HK076	1.8	-2.06	0.56	0.96	0.1	0.94	0
716	HK077	1.8	-1.84	0.56	1.60	1.0	1.61	1.0
717	HK078	2.0	-1.76	0.55	2.82	2.3	2.87	2.4
718	HK079	3.8	1.58	0.52	0.37	-1.3	0.37	-1.3
719	HK080	3.0	0	0.53	2.48	2.0	2.53	2.0

720	HK081	1.0	-3.48	1.03	0.06	-1.6	0.06	-1.6
721	HK082	1.0	-3.64	1.03	0.02	-2.0	0.02	-2.0
722	HK083	3.5	1.05	0.52	0.98	0.1	0.97	0.1
723	HK084	4.3	1.77	0.65	2.85	2.0	2.75	1.9
724	HK085	2.5	-0.63	0.54	0.32	-1.4	0.32	-1.4
725	HK086	1.8	-2.06	0.56	5.36	4.1	5.18	4.0
726	HK087	3.7	1.32	0.52	1.22	0.5	1.22	0.5
727	HK088	1.5	-2.69	0.57	0.40	-1.1	0.40	-1.2
728	HK089	1.5	-2.47	0.57	0.98	0.1	0.97	0.1
729	HK090	3.3	0.55	0.52	2.30	1.8	2.33	1.8
730	HK091	2.2	-1.23	0.55	0.72	-0.3	0.72	-0.3
731	HK092	2.3	-1.15	0.54	0.83	-0.1	0.82	-0.1
732	HK093	2.7	-0.35	0.54	0.19	-2.0	0.19	-2.0
733	HK094	4.0	1.63	0.52	1.47	0.9	1.45	0.9
734	HK095	5.0	3.61	0.58	1.62	1.1	1.63	1.1
735	HK096	3.5	0.82	0.52	0.18	-2.0	0.18	-2.0
736	HK097	3.3	0.78	0.52	0.86	0	0.87	0
737	HK098	1.0	-3.84	0.75	0.31	-1.1	0.34	-1.1
738	HK099	2.8	-0.06	0.53	0.32	-1.4	0.33	-1.4
739	HK100	2.3	-1.15	0.54	0.53	-0.7	0.54	-0.7
740	HK101	3.7	1.32	0.52	0.42	-1.1	0.42	-1.1
741	HK102	2.3	-1.16	0.54	0.45	-1.0	0.45	-1.0
742	HK103	3.8	1.58	0.52	0.70	-0.4	0.7	-0.4
743	HK104	2.2	-1.45	0.55	2.91	2.4	2.86	2.3
744	HK105	2.7	-0.35	0.54	0.19	-2.0	0.19	-2.0
745	HK106	2.0	-1.75	0.55	0.71	-0.3	0.71	-0.3
746	HK107	2.5	-0.64	0.54	0.61	-0.5	0.62	-0.5
747	HK108	1.2	-3.36	0.59	0.86	0	0.87	0
748	HK109	4.0	1.85	0.52	0.20	-2.0	0.20	-2.0
749	HK110	2.8	-0.28	0.53	0.97	0.1	0.93	0
750	HK111	2.5	-0.64	0.54	0.44	-1.0	0.42	-1.0
751	HK112	3.3	0.55	0.52	2.39	1.9	2.45	2.0
752	HK113	4.8	3.29	0.56	1.40	0.8	1.30	0.6
753	HK114	3.8	1.37	0.52	2.64	2.2	2.63	2.2
754	HK115	3.2	0.51	0.53	0.96	0.1	0.96	0.1
755	HK116	4.2	1.91	0.52	0.33	-1.4	0.35	-1.4
756	HK117	3.0	0.22	0.53	1.61	1.0	1.64	1.1
757	HK118	3.8	1.36	0.52	0.50	-0.9	0.50	-0.9

758	HK119	3.7	1.32	0.52	0.47	-0.9	0.48	-0.9
759	HK120	3.0	0	0.53	1.41	0.8	1.45	0.8
760	HK121	3.0	0.23	0.53	1.41	0.8	1.38	0.7
761	HK122	2.0	-1.75	0.55	0.43	-1.0	0.43	-1.0
762	HK123	3.3	0.77	0.52	0.19	-2.0	0.18	-2.0
763	HK124	1.3	-3.01	0.57	0.12	-2.5	0.12	-2.5
764	HK125	3.2	0.50	0.53	0.81	-0.1	0.80	-0.1
765	HK126	4.7	2.76	0.55	0.87	0	0.97	0.1
766	HK127	5.0	3.57	0.57	0.56	-0.7	0.55	-0.7
767	HK128	2.7	-0.57	0.54	1.34	0.7	1.29	0.6
768	HK129	5.2	3.95	0.61	0.41	-1.1	0.50	-0.9
769	HK130	4.2	1.90	0.52	0.99	0.1	0.96	0.1
770	HK131	1.8	-2.18	0.68	0.19	-1.6	0.19	-1.6
771	HK132	4.0	1.64	0.52	0.32	-1.5	0.33	-1.4
772	HK133	3.0	0.23	0.53	1.07	0.3	1.05	0.2
773	HK134	3.0	0	0.53	0.13	-2.4	0.13	-2.4
774	HK135	5.8	6.21	1.07	0.87	0.1	0.66	0
775	HK136	3.0	0	0.53	0.51	-0.8	0.51	-0.8
776	HK137	2.3	-0.94	0.54	0.83	-0.1	0.82	-0.1

Separation: 2.99 Reliability: 0.90

Fixed (all same) chi-square: 1143.6 significance p: .00

Calibration of the HK students for HK student essays using HK raters (34 misfitting students)

Student		Observed score	Measure (Logit)	S.E.	Infit MnSq	Infit ZStd	Outfit MnSq	Outfit ZStd
662	HK023	4.0	1.63	0.52	1.41	0.8	1.41	0.8
694	HK055	3.8	1.58	0.52	1.41	0.8	1.41	0.8
759	HK120	3.0	0	0.53	1.41	0.8	1.45	0.8
760	HK121	3.0	0.23	0.53	1.41	0.8	1.38	0.7
733	HK094	4.0	1.63	0.52	1.47	0.9	1.45	0.9
658	HK019	3.8	1.37	0.52	1.49	0.9	1.48	0.9
688	HK049	3.7	1.32	0.52	1.53	1.0	1.53	1.0
697	HK058	3.3	0.55	0.52	1.56	1.0	1.53	0.9
693	HK054	3.8	1.37	0.52	1.57	1.0	1.60	1.0
704	HK065	4.0	1.86	0.52	1.57	1.0	1.66	1.1
698	HK059	3.3	0.77	0.52	1.59	1.0	1.60	1.0

716	HK077	1.8	-1.84	0.56	1.60	1.0	1.61	1.0
713	HK074	2.0	-1.75	0.55	1.61	1.0	1.58	1.0
756	HK117	3.0	0.22	0.53	1.61	1.0	1.64	1.1
734	HK095	5.0	3.61	0.58	1.62	1.1	1.63	1.1
650	HK011	2.3	-1.15	0.54	1.63	1.1	1.60	1.0
702	HK063	4.2	2.12	0.52	1.63	1.1	1.64	1.1
648	HK009	3.7	1.09	0.52	1.69	1.2	1.70	1.2
712	HK073	2.8	-0.06	0.53	1.88	1.4	1.91	1.4
683	HK044	3.8	1.36	0.52	2.05	1.6	2.06	1.6
640	HK001	1.0	-3.57	0.74	2.19	1.5	2.01	1.4
687	HK048	5.2	3.75	0.61	2.23	1.7	2.64	2.2
701	HK062	2.7	-0.57	0.54	2.25	1.8	2.17	1.7
729	HK090	3.3	0.55	0.52	2.30	1.8	2.33	1.8
711	HK072	1.8	-2.08	0.68	2.38	1.6	2.35	1.6
751	HK112	3.3	0.55	0.52	2.39	1.9	2.45	2.0
719	HK080	3.0	0	0.53	2.48	2.0	2.53	2.0
753	HK114	3.8	1.37	0.52	2.64	2.2	2.63	2.2
717	HK078	2.0	-1.76	0.55	2.82	2.3	2.87	2.4
723	HK084	4.3	1.77	0.65	2.85	2.0	2.75	1.9
743	HK104	2.2	-1.45	0.55	2.91	2.4	2.86	2.3
660	HK021	3.8	1.36	0.52	4.68	3.8	4.68	3.7
725	HK086	1.8	-2.06	0.56	5.36	4.1	5.18	4.0
675	HK036	3.2	0.28	0.53	7.49	5.1	7.55	5.2

Separation: 2.99 Reliability: 0.90
Fixed (all same) chi-square: 1143.6 significance *p*: .00

Calibration of the prompts for HK student essays using HK raters

Prompt	Observed score	Measure (Logit)	S.E.	Infit MnSq	Infit ZStd	Outfit MnSq	Outfit ZStd
4BI	3.1	0.28	0.08	0.97	-0.3	0.96	-0.4
4BN	3.3	-0.05	0.08	0.90	-1.1	0.90	-1.1
4BP	3.4	-0.24	0.08	1.11	1.3	1.12	1.3

Separation: 2.45 Reliability: 0.86
Fixed (all same) chi-square: 21.1 significance *p*: .00

Calibration of the HK raters for HK student essays using HK raters

HK rater	Observed score	Measure (Logit)	S.E.	Infit		Outfit	
				MnSq	ZStd	MnSq	ZStd
HKR001	3.0	0.34	0.09	0.91	-0.8	0.92	-0.7
HKR002	3.3	-0.02	0.09	0.82	-1.8	0.82	-1.8
HKR003	3.8	-1.00	0.09	1.34	3.1	1.35	3.1
HKR004	2.9	0.68	0.09	0.91	-0.8	0.89	-1.1

Separation: 6.65 Reliability: 0.98

Fixed (all same) chi-square: 180.1 significance *p*: .00

Appendix 14

Calibration of the HK students for HK student essays using HK raters and Lexile Analyzer

Student	Number	Observed score	Measure (Logit)	S.E.	Ifit MnSq	Ifit ZStd	Outfit MnSq	Outfit ZStd
640	HK001	1.0	-2.68	0.75	2.30	1.6	2.12	1.5
641	HK002	3.2	0.91	0.54	0.50	-0.8	0.50	-0.8
642	HK003	3.8	1.17	0.53	0.66	-0.4	0.65	-0.5
643	HK004	3.3	1.28	0.54	0.42	-1.1	0.43	-1.1
644	HK005	1.3	-2.19	0.73	0.27	-1.3	0.26	-1.3
645	HK006	3.5	1.19	0.54	0.56	-0.7	0.55	-0.7
646	HK007	5.3	3.85	0.66	0.22	-1.9	0.25	-1.7
647	HK008	3.2	1.06	0.54	0.40	-1.1	0.41	-1.1
648	HK009	3.7	1.21	0.53	1.35	0.7	1.35	0.7
649	HK010	4.8	3.06	0.58	0.48	-0.9	0.51	-0.9
650	HK011	2.3	-0.06	0.56	1.87	1.3	1.84	1.3
651	HK012	4.7	3.06	0.56	0.58	-0.7	0.55	-0.7
652	HK013	3.7	1.68	0.54	0.70	-0.4	0.71	-0.3
653	HK014	4.2	2.48	0.54	1.02	0.2	1.05	0.2
654	HK015	5.3	3.55	0.66	0.51	-0.8	0.62	-0.6
655	HK016	4.2	2.04	0.54	0.21	-1.9	0.21	-1.9
656	HK017	3.0	0.25	0.54	0.26	-1.7	0.26	-1.7
657	HK018	3.3	0.88	0.54	0.09	-2.6	0.09	-2.6
658	HK019	3.8	1.52	0.54	2.22	1.8	2.19	1.7
659	HK020	3.0	0.62	0.54	0.45	-1.0	0.45	-1.0
660	HK021	3.8	1.25	0.53	5.67	4.3	5.76	4.4
661	HK022	1.3	-1.79	0.59	0.93	0	0.92	0
662	HK023	4.0	1.21	0.54	0.98	0.1	1.01	0.2
663	HK024	4.8	3.54	0.58	0.71	-0.3	0.67	-0.4
664	HK025	4.7	2.87	0.56	0.64	-0.5	0.67	-0.4
665	HK026	3.7	1.74	0.53	0.25	-1.7	0.25	-1.7
666	HK027	2.8	0.40	0.55	1.32	0.6	1.28	0.6
667	HK028	3.2	0.69	0.54	0.23	-1.8	0.23	-1.8
668	HK029	3.0	0.69	0.54	0.60	-0.6	0.61	-0.6
669	HK030	4.2	1.63	0.54	0.47	-0.9	0.47	-1.0
670	HK031	3.2	0.47	0.54	0.39	-1.2	0.38	-1.2
671	HK032	4.8	1.46	0.61	2.35	1.8	2.81	2.3

672	HK033	3.0	0.97	0.54	0.18	-2.0	0.18	-2.0
673	HK034	4.2	2.26	0.54	0.88	0	0.87	0
674	HK035	3.3	0.98	0.54	0.22	-1.8	0.22	-1.9
675	HK036	3.2	-1.07	0.56	3.04	2.5	3.02	2.5
676	HK037	3.0	0.02	0.67	0.57	-0.4	0.57	-0.4
677	HK038	3.7	1.24	0.53	0.27	-1.6	0.27	-1.6
678	HK039	3.0	0.96	0.54	0.61	-0.5	0.59	-0.6
679	HK040	4.0	1.97	0.53	0.33	-1.4	0.33	-1.4
680	HK041	2.8	0.17	0.55	0.64	-0.5	0.64	-0.5
681	HK042	4.2	1.45	0.54	0.71	-0.3	0.69	-0.4
682	HK043	3.0	0.56	0.54	0.18	-2.0	0.18	-2.0
683	HK044	3.8	1.90	0.53	2.20	1.7	2.21	1.8
684	HK045	3.2	0.84	0.54	0.31	-1.4	0.32	-1.4
685	HK046	2.8	0.29	0.55	0.27	-1.6	0.27	-1.6
686	HK047	3.5	1.42	0.54	0.26	-1.7	0.26	-1.7
687	HK048	5.2	4.30	0.63	2.55	2.0	2.98	2.4
688	HK049	3.7	1.70	0.53	1.82	1.3	1.82	1.3
689	HK050	3.5	1.23	0.65	0.52	-0.6	0.51	-0.6
690	HK051	2.8	0.47	0.55	0.62	-0.5	0.62	-0.5
691	HK052	3.2	0.62	0.54	0.25	-1.7	0.25	-1.7
692	HK053	2.8	0.39	0.55	0.80	-0.1	0.79	-0.1
693	HK054	3.8	1.91	0.54	1.66	1.1	1.71	1.2
694	HK055	3.8	1.58	0.53	1.32	0.7	1.31	0.6
695	HK056	2.5	0.04	0.55	0.92	0	0.91	0
696	HK057	2.0	-0.88	0.57	0.39	-1.2	0.39	-1.2
697	HK058	3.3	1.05	0.54	1.76	1.2	1.73	1.2
698	HK059	3.3	1.28	0.54	1.70	1.2	1.71	1.2
699	HK060	2.7	0.14	0.55	0.41	-1.1	0.41	-1.1
700	HK061	2.8	0.40	0.55	0.37	-1.2	0.37	-1.2
701	HK062	2.7	-0.28	0.55	2.37	1.9	2.27	1.8
702	HK063	4.2	1.67	0.54	2.20	1.8	2.21	1.8
703	HK064	2.5	-0.51	0.55	1.29	0.6	1.28	0.6
704	HK065	4.0	2.13	0.54	1.76	1.2	1.89	1.4
705	HK066	3.2	1.23	0.54	0.72	-0.3	0.71	-0.3
706	HK067	4.0	1.86	0.53	0.81	-0.1	0.80	-0.2
707	HK068	2.8	0.10	0.55	0.52	-0.8	0.52	-0.8
708	HK069	2.8	0.17	0.55	1.16	0.4	1.18	0.4
709	HK070	4.0	1.96	0.53	0.76	-0.2	0.75	-0.2

710	HK071	2.3	-0.82	0.56	0.64	-0.5	0.61	-0.5
711	HK072	1.8	-1.11	0.70	2.48	1.7	2.45	1.7
712	HK073	2.8	0.60	0.55	1.44	0.8	1.46	0.8
713	HK074	2.0	-1.11	0.57	1.58	1.0	1.52	0.9
714	HK075	4.2	1.98	0.54	0.78	-0.2	0.77	-0.2
715	HK076	1.8	-1.43	0.57	1.21	0.5	1.19	0.5
716	HK077	1.8	-1.63	0.57	1.94	1.4	1.96	1.5
717	HK078	2.0	-1.59	0.57	2.37	1.9	2.43	1.9
718	HK079	3.8	1.73	0.53	0.46	-1.0	0.46	-1.0
719	HK080	3.0	0.46	0.54	2.27	1.8	2.32	1.8
720	HK081	1.0	-2.57	1.05	0.06	-1.6	0.06	-1.6
721	HK082	1.0	-2.73	1.05	0.02	-2.0	0.02	-2.0
722	HK083	3.5	1.17	0.54	0.81	-0.1	0.80	-0.1
723	HK084	4.3	1.99	0.67	2.84	2.0	2.73	1.9
724	HK085	2.5	0.07	0.55	0.20	-1.9	0.20	-1.9
725	HK086	1.8	-1.44	0.57	5.20	4.0	4.96	3.8
726	HK087	3.7	1.78	0.53	1.17	0.4	1.18	0.4
727	HK088	1.5	-1.89	0.58	0.47	-0.9	0.47	-1.0
728	HK089	1.5	-1.87	0.58	1.05	0.2	1.03	0.2
729	HK090	3.3	0.97	0.54	2.28	1.8	2.27	1.8
730	HK091	2.2	-0.98	0.56	0.86	0	0.86	0
731	HK092	2.3	-0.27	0.56	0.93	0	0.92	0
732	HK093	2.7	-0.04	0.55	0.19	-1.9	0.19	-1.9
733	HK094	4.0	1.28	0.53	1.43	0.8	1.41	0.8
734	HK095	5.0	3.60	0.59	1.48	0.9	1.42	0.8
735	HK096	3.5	1.04	0.54	0.15	-2.3	0.15	-2.3
736	HK097	3.3	1.28	0.54	0.90	0	0.91	0
737	HK098	1.0	-2.96	0.76	0.34	-1.1	0.37	-1.0
738	HK099	2.8	0.29	0.55	0.45	-1.0	0.45	-1.0
739	HK100	2.3	-0.27	0.56	0.56	-0.7	0.56	-0.7
740	HK101	3.7	1.56	0.53	0.43	-1.1	0.43	-1.1
741	HK102	2.3	-0.83	0.56	0.31	-1.4	0.31	-1.4
742	HK103	3.8	1.76	0.53	0.85	0	0.85	-0.1
743	HK104	2.2	-0.86	0.56	2.50	2.0	2.46	2.0
744	HK105	2.7	0.03	0.55	0.24	-1.7	0.24	-1.7
745	HK106	2.0	-1.25	0.57	0.55	-0.7	0.55	-0.7
746	HK107	2.5	-0.01	0.55	0.99	0.1	1.01	0.2
747	HK108	1.2	-2.59	0.60	1.14	0.4	1.13	0.4

748	HK109	4.0	2.41	0.53	0.21	-2.0	0.21	-2.0
749	HK110	2.8	0.03	0.55	0.92	0	0.89	0
750	HK111	2.5	-0.28	0.55	0.48	-0.9	0.46	-0.9
751	HK112	3.3	1.17	0.54	1.99	1.5	2.03	1.5
752	HK113	4.8	3.38	0.58	1.23	0.5	1.21	0.5
753	HK114	3.8	1.77	0.53	2.78	2.3	2.76	2.3
754	HK115	3.2	0.86	0.54	1.02	0.2	1.02	0.2
755	HK116	4.2	2.41	0.54	0.40	-1.2	0.43	-1.1
756	HK117	3.0	0.63	0.54	1.83	1.3	1.85	1.3
757	HK118	3.8	1.69	0.53	0.52	-0.8	0.52	-0.8
758	HK119	3.7	1.62	0.54	0.49	-0.9	0.50	-0.9
759	HK120	3.0	0.51	0.54	0.92	0	0.96	0.1
760	HK121	3.0	0.77	0.54	1.59	1.0	1.56	1.0
761	HK122	2.0	-1.18	0.57	0.59	-0.6	0.61	-0.6
762	HK123	3.3	1.21	0.54	0.25	-1.7	0.24	-1.7
763	HK124	1.3	-2.02	0.59	0.14	-2.3	0.14	-2.4
764	HK125	3.2	0.73	0.54	0.88	0	0.89	0
765	HK126	4.7	2.90	0.57	1.03	0.2	1.33	0.7
766	HK127	5.0	3.70	0.60	1.04	0.2	0.99	0.1
767	HK128	2.7	0.35	0.55	1.57	1.0	1.52	0.9
768	HK129	5.2	3.81	0.63	0.38	-1.2	0.56	-0.7
769	HK130	4.2	1.92	0.54	0.87	0	0.84	-0.1
770	HK131	1.8	-1.22	0.70	0.20	-1.5	0.20	-1.5
771	HK132	4.0	1.98	0.54	0.43	-1.1	0.45	-1.0
772	HK133	3.0	0.91	0.54	0.88	0	0.87	0
773	HK134	3.0	0.46	0.54	0.14	-2.3	0.14	-2.3
774	HK135	5.8	10.46	1.21	0.90	0.2	0.31	3.1
775	HK136	3.0	0.32	0.54	0.48	-0.9	0.49	-0.9
776	HK137	2.3	-0.45	0.56	0.79	-0.2	0.78	-0.2

Separation: 2.76 Reliability: 0.88

Fixed (all same) chi-square: 871.6 significance *p*: .00

Appendix 15

Student measures of 270 Grades 4 and 6 US student essays

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Student
143	27	3	90.38	18.44		MAXIMUM MEASURE			.00	.00	100.0	100.0	272
163	27	3	90.38	18.44		MAXIMUM MEASURE			.00	.00	100.0	100.0	292
173	27	3	90.38	18.44		MAXIMUM MEASURE			.00	.00	100.0	100.0	302
189	27	3	90.38	18.44		MAXIMUM MEASURE			.00	.00	100.0	100.0	410
219	27	3	90.38	18.44		MAXIMUM MEASURE			.00	.00	100.0	100.0	440
226	27	3	90.38	18.44		MAXIMUM MEASURE			.00	.00	100.0	100.0	447
164	26	3	78.05	10.19	.45	-.1	.39	-.2	.83	.19	66.7	71.7	293
35	25	3	70.85	7.23	.90	.3	.79	.2	.83	.26	33.3	46.7	035
159	25	3	70.85	7.23	.15	-.9	.21	-.7	.90	.26	100.0	46.7	288
167	25	3	70.85	7.23	1.49	.8	1.44	.7	.07	.26	33.3	46.7	296
169	25	3	70.85	7.23	.15	-.9	.21	-.7	.90	.26	100.0	46.7	298
175	25	3	70.85	7.23	.15	-.9	.21	-.7	.90	.26	100.0	46.7	304
187	25	3	70.85	7.23	.90	.3	.79	.2	.83	.26	33.3	46.7	408
245	25	3	70.85	7.23	.15	-.9	.21	-.7	.90	.26	100.0	46.7	466
14	24	3	66.63	5.90	.51	-.3	.51	-.3	.44	.31	33.3	36.6	014
65	24	3	66.63	5.90	1.35	.7	1.57	.8	-1.00	.31	33.3	36.6	065
130	24	3	66.63	5.90	.55	-.2	.63	-.1	.56	.31	33.3	36.6	259
188	24	3	66.63	5.90	1.35	.7	1.19	.5	.83	.31	.0	36.6	409
223	24	3	66.63	5.90	.11	-1.2	.12	-1.2	.00	.31	100.0	36.6	444
15	16	2	64.26	7.22	.05	-1.1	.06	-1.1	.00	.23	100.0	37.7	015
20	23	3	63.62	5.13	.33	-.7	.29	-.8	.07	.35	33.3	34.0	020
56	23	3	63.62	5.13	.33	-.7	.29	-.8	.07	.35	33.3	34.0	056
78	23	3	63.62	5.13	.04	-1.8	.04	-1.8	.83	.35	100.0	34.0	078
140	23	3	63.62	5.13	.54	-.3	.55	-.3	.97	.35	33.3	34.0	269
180	23	3	63.62	5.13	1.13	.5	1.19	.5	.39	.35	.0	34.0	401
195	23	3	63.62	5.13	.31	-.7	.42	-.5	.90	.35	33.3	34.0	416
228	23	3	63.62	5.13	.54	-.3	.55	-.3	.97	.35	33.3	34.0	449
231	23	3	63.62	5.13	.82	.1	1.05	.4	-.07	.35	33.3	34.0	452
237	23	3	63.62	5.13	.31	-.7	.42	-.5	.90	.35	33.3	34.0	458
250	23	3	63.62	5.13	.31	-.7	.42	-.5	.90	.35	33.3	34.0	471
63	8	1	61.56	10.21	.00	-1.6	.00	-1.6	.00	.00	100.0	38.8	063
100	8	1	61.56	10.21	.00	-1.6	.00	-1.6	.00	.00	100.0	38.8	229
11	22	3	61.25	4.63	.35	-.7	.46	-.5	.99	.39	33.3	30.2	011
26	22	3	61.25	4.63	.92	.2	.89	.2	.94	.39	.0	30.2	026
165	22	3	61.25	4.63	.78	.0	.68	-.1	.07	.39	33.3	30.2	294
172	22	3	61.25	4.63	.78	.0	.68	-.1	.07	.39	33.3	30.2	301
176	22	3	61.25	4.63	.92	.2	.89	.2	.94	.39	.0	30.2	305
215	22	3	61.25	4.63	.05	-1.9	.05	-1.9	.90	.39	100.0	30.2	436
36	21	3	59.27	4.30	2.93	1.8	2.60	1.6	.83	.42	.0	29.2	036
110	21	3	59.27	4.30	4.69	2.7	4.43	2.6	.07	.42	.0	29.2	239
119	21	3	59.27	4.30	1.15	.5	1.26	.6	.56	.42	.0	29.2	248
120	21	3	59.27	4.30	.46	-.5	.40	-.7	.83	.42	33.3	29.2	249
170	21	3	59.27	4.30	2.55	1.6	2.52	1.6	.26	.42	.0	29.2	299
171	21	3	59.27	4.30	1.15	.5	1.26	.6	.56	.42	.0	29.2	300
178	21	3	59.27	4.30	1.65	.9	1.42	.7	.71	.42	33.3	29.2	399
207	21	3	59.27	4.30	1.15	.5	1.26	.6	.56	.42	.0	29.2	428
221	21	3	59.27	4.30	.57	-.3	.64	-.2	1.00	.42	33.3	29.2	442
233	21	3	59.27	4.30	1.10	.4	1.07	.4	.44	.42	.0	29.2	454
243	21	3	59.27	4.30	.21	-1.2	.24	-1.1	.00	.42	33.3	29.2	464
16	20	3	57.53	4.06	.07	-2.0	.09	-1.8	.83	.44	66.7	28.3	016
37	20	3	57.53	4.06	.44	-.6	.42	-.7	.60	.44	.0	28.3	037
40	20	3	57.53	4.06	.91	.2	.95	.2	.98	.44	33.3	28.3	040
45	20	3	57.53	4.06	.74	-.1	.69	-.2	.39	.44	33.3	28.3	045
54	20	3	57.53	4.06	.44	-.6	.42	-.7	.60	.44	.0	28.3	054
66	20	3	57.53	4.06	1.58	.9	1.97	1.2	-.71	.44	.0	28.3	066
68	20	3	57.53	4.06	.12	-1.6	.13	-1.6	.90	.44	66.7	28.3	068
75	20	3	57.53	4.06	.12	-1.6	.13	-1.6	.90	.44	66.7	28.3	075
114	20	3	57.53	4.06	.16	-1.5	.15	-1.5	.97	.44	66.7	28.3	243
132	20	3	57.53	4.06	.12	-1.6	.13	-1.6	.90	.44	66.7	28.3	261
144	20	3	57.53	4.06	.74	-.1	.69	-.2	.39	.44	33.3	28.3	273
158	20	3	57.53	4.06	.83	.0	.96	.2	.77	.44	33.3	28.3	287
161	20	3	57.53	4.06	.68	-.2	.83	.1	-.07	.44	33.3	28.3	290
184	20	3	57.53	4.06	.86	.1	.75	-.1	.83	.44	33.3	28.3	405
201	20	3	57.53	4.06	.68	-.2	.83	.1	-.07	.44	33.3	28.3	422
222	20	3	57.53	4.06	.54	-.4	.67	-.2	.98	.44	.0	28.3	443
225	20	3	57.53	4.06	2.22	1.4	1.93	1.2	.73	.44	33.3	28.3	446

248	20	3	57.53	4.06	.44	-.6	.42	-.7	.60	.44	.0	28.3	469
254	20	3	57.53	4.06	1.94	1.2	1.81	1.1	.90	.44	.0	28.3	573
4	19	3	55.94	3.90	1.28	.6	1.52	.8	-.39	.46	.0	23.8	004
18	19	3	55.94	3.90	1.37	.7	1.23	.5	.83	.46	.0	23.8	018
27	19	3	55.94	3.90	1.37	.7	1.23	.5	.83	.46	.0	23.8	027
53	19	3	55.94	3.90	.76	-.1	.86	.1	1.00	.46	33.3	23.8	053
76	19	3	55.94	3.90	.44	-.7	.41	-.7	.94	.46	.0	23.8	076
79	19	3	55.94	3.90	.38	-.8	.45	-.6	-.07	.46	.0	23.8	079
103	19	3	55.94	3.90	.44	-.7	.41	-.7	.94	.46	.0	23.8	232
112	19	3	55.94	3.90	1.33	.6	1.43	.7	.65	.46	.0	23.8	241
113	19	3	55.94	3.90	2.53	1.7	2.52	1.7	.39	.46	.0	23.8	242
146	19	3	55.94	3.90	.12	-1.7	.14	-1.6	.99	.46	66.7	23.8	275
174	19	3	55.94	3.90	.44	-.7	.41	-.7	.94	.46	.0	23.8	303
181	19	3	55.94	3.90	.38	-.8	.45	-.6	-.07	.46	.0	23.8	402
190	19	3	55.94	3.90	2.91	1.9	2.58	1.7	.75	.46	.0	23.8	411
230	19	3	55.94	3.90	1.56	.9	1.46	.8	-.60	.46	.0	23.8	451
235	19	3	55.94	3.90	.09	-1.9	.08	-1.9	.90	.46	66.7	23.8	456
236	19	3	55.94	3.90	2.78	1.8	3.29	2.1	-.46	.46	.0	23.8	457
246	19	3	55.94	3.90	.40	-.8	.41	-.7	.71	.46	.0	23.8	467
252	19	3	55.94	3.90	.74	-.1	.87	.1	.90	.46	33.3	23.8	571
6	18	3	54.47	3.79	.28	-1.1	.25	-1.2	.56	.48	33.3	24.4	006
10	18	3	54.47	3.79	2.05	1.3	2.31	1.5	-.90	.48	.0	24.4	010
23	18	3	54.47	3.79	2.61	1.8	2.80	1.9	-1.00	.48	33.3	24.4	023
24	18	3	54.47	3.79	.30	-1.1	.32	-1.0	.00	.48	33.3	24.4	024
28	18	3	54.47	3.79	.29	-1.1	.27	-1.1	.83	.48	33.3	24.4	028
41	18	3	54.47	3.79	.27	-1.1	.30	-1.1	.90	.48	33.3	24.4	041
51	18	3	54.47	3.79	.89	.1	1.01	.3	-.56	.48	.0	24.4	051
62	18	3	54.47	3.79	.00	-3.5	.00	-3.4	1.00	.48	100.0	24.4	062
101	18	3	54.47	3.79	1.14	.4	1.03	.3	.07	.48	33.3	24.4	230
136	18	3	54.47	3.79	1.97	1.3	2.01	1.3	.56	.48	.0	24.4	265
147	18	3	54.47	3.79	.28	-1.1	.30	-1.1	1.00	.48	33.3	24.4	276
149	18	3	54.47	3.79	1.98	1.3	1.86	1.2	.26	.48	.0	24.4	278
179	18	3	54.47	3.79	2.05	1.3	2.10	1.4	.12	.48	.0	24.4	400
182	18	3	54.47	3.79	1.16	.5	1.06	.3	.71	.48	33.3	24.4	403
193	18	3	54.47	3.79	3.75	2.5	3.42	2.3	.76	.48	.0	24.4	414
247	18	3	54.47	3.79	.27	-1.1	.30	-1.1	.90	.48	33.3	24.4	468
259	18	3	54.47	3.79	4.63	2.9	4.97	3.0	-1.00	.48	33.3	24.4	578
260	18	3	54.47	3.79	.86	.1	.83	.0	-.44	.48	.0	24.4	579
1	17	3	53.05	3.72	.71	-.2	.76	-.1	-.07	.49	33.3	23.9	001
3	17	3	53.05	3.72	.39	-.8	.38	-.8	.07	.49	.0	23.9	003
32	17	3	53.05	3.72	.12	-1.8	.13	-1.7	.83	.49	66.7	23.9	032
82	17	3	53.05	3.72	1.27	.6	1.28	.6	.17	.49	.0	23.9	211
93	17	3	53.05	3.72	2.38	1.6	2.32	1.6	.26	.49	.0	23.9	222
109	17	3	53.05	3.72	1.44	.8	1.39	.7	.90	.49	.0	23.9	238
117	17	3	53.05	3.72	1.26	.6	1.28	.6	-.83	.49	.0	23.9	246
157	17	3	53.05	3.72	1.86	1.2	1.98	1.3	-.31	.49	33.3	23.9	286
191	17	3	53.05	3.72	.32	-1.0	.35	-.9	.98	.49	.0	23.9	412
194	17	3	53.05	3.72	1.86	1.2	1.98	1.3	-.31	.49	33.3	23.9	415
84	10	2	52.58	4.40	.88	.2	.88	.2	1.00	.29	.0	22.1	213
151	10	2	52.58	4.40	.10	-1.5	.10	-1.5	1.00	.29	.0	22.1	280
263	11	2	52.22	4.62	2.50	1.5	2.51	1.5	1.00	.58	.0	24.8	582
2	16	3	51.68	3.69	4.13	2.7	4.30	2.8	-.71	.49	.0	23.2	002
9	16	3	51.68	3.69	.12	-1.8	.11	-1.8	.83	.49	66.7	23.2	009
21	16	3	51.68	3.69	.67	-.3	.64	-.3	.67	.49	33.3	23.2	021
64	16	3	51.68	3.69	.06	-2.2	.06	-2.2	.99	.49	66.7	23.2	064
69	16	3	51.68	3.69	1.13	.4	1.08	.4	.31	.49	.0	23.2	069
123	16	3	51.68	3.69	1.54	.9	1.49	.8	.33	.49	.0	23.2	252
141	16	3	51.68	3.69	.67	-.3	.64	-.3	.67	.49	33.3	23.2	270
155	16	3	51.68	3.69	.32	-1.0	.31	-1.1	.71	.49	.0	23.2	284
162	16	3	51.68	3.69	.44	-.7	.46	-.7	-.07	.49	.0	23.2	291
241	16	3	51.68	3.69	.11	-1.9	.11	-1.9	.90	.49	66.7	23.2	462
256	16	3	51.68	3.69	.06	-2.2	.06	-2.2	.99	.49	66.7	23.2	575
13	15	3	50.32	3.69	.21	-1.4	.20	-1.4	.90	.50	33.3	23.7	013
22	15	3	50.32	3.69	.34	-1.0	.34	-1.0	.44	.50	33.3	23.7	022
31	15	3	50.32	3.69	.00	-3.1	.00	-3.1	1.00	.50	100.0	23.7	031
44	15	3	50.32	3.69	1.81	1.1	1.77	1.1	.26	.50	.0	23.7	044
67	15	3	50.32	3.69	1.76	1.1	1.70	1.0	.56	.50	.0	23.7	067
74	15	3	50.32	3.69	.29	-1.1	.29	-1.1	.83	.50	33.3	23.7	074
107	15	3	50.32	3.69	2.40	1.6	2.40	1.6	-.12	.50	33.3	23.7	236
115	15	3	50.32	3.69	.99	.2	1.03	.3	1.00	.50	33.3	23.7	244
148	15	3	50.32	3.69	.29	-1.1	.29	-1.1	.83	.50	33.3	23.7	277
152	15	3	50.32	3.69	.29	-1.1	.29	-1.1	.83	.50	33.3	23.7	281
153	15	3	50.32	3.69	.23	-1.3	.24	-1.3	1.00	.50	33.3	23.7	282
154	15	3	50.32	3.69	3.33	2.3	3.44	2.3	-.71	.50	.0	23.7	283
166	15	3	50.32	3.69	1.23	.5	1.20	.5	-.07	.50	33.3	23.7	295

196	15	3	50.32	3.69	.96	.2	.93	.2	.80	.50	33.3	23.7	417
217	15	3	50.32	3.69	.21	-1.4	.20	-1.4	.90	.50	33.3	23.7	438
249	15	3	50.32	3.69	1.06	.3	1.06	.3	.07	.50	33.3	23.7	470
251	15	3	50.32	3.69	.33	-1.0	.34	-1.0	.00	.50	33.3	23.7	472
265	15	3	50.32	3.69	.21	-1.4	.20	-1.4	.90	.50	33.3	23.7	584
48	11	2	50.24	4.45	.30	-.8	.31	-.8	1.00	.37	.0	23.2	048
25	14	3	48.94	3.72	2.65	1.8	2.56	1.7	.49	.49	.0	23.2	025
33	14	3	48.94	3.72	1.05	.3	1.01	.3	.90	.49	.0	23.2	033
55	14	3	48.94	3.72	1.67	1.0	1.75	1.1	.99	.49	33.3	23.2	055
71	14	3	48.94	3.72	1.34	.7	1.31	.6	.46	.49	.0	23.2	071
90	14	3	48.94	3.72	.06	-2.2	.06	-2.2	.90	.49	66.7	23.2	219
96	14	3	48.94	3.72	.54	-.5	.52	-.5	.77	.49	33.3	23.2	225
97	14	3	48.94	3.72	.60	-.4	.64	-.3	.98	.49	33.3	23.2	226
168	14	3	48.94	3.72	.09	-1.9	.10	-1.9	.97	.49	66.7	23.2	297
205	14	3	48.94	3.72	.09	-1.9	.10	-1.9	.97	.49	66.7	23.2	426
253	14	3	48.94	3.72	2.69	1.8	2.68	1.8	.23	.49	.0	23.2	572
255	14	3	48.94	3.72	1.58	.9	1.50	.8	.71	.49	33.3	23.2	574
19	13	3	47.54	3.78	.09	-1.9	.10	-1.9	.90	.49	66.7	24.9	019
38	13	3	47.54	3.78	1.09	.4	1.04	.3	.65	.49	.0	24.9	038
59	13	3	47.54	3.78	.78	-.1	.73	-.1	.26	.49	33.3	24.9	059
104	13	3	47.54	3.78	1.70	1.0	1.87	1.2	-.17	.49	33.3	24.9	233
108	13	3	47.54	3.78	1.09	.4	1.04	.3	.65	.49	.0	24.9	237
118	13	3	47.54	3.78	.76	-.1	.79	.0	.67	.49	33.3	24.9	247
145	13	3	47.54	3.78	1.24	.6	1.35	.7	.97	.49	.0	24.9	274
229	13	3	47.54	3.78	1.99	1.3	1.86	1.2	-.07	.49	33.3	24.9	450
232	13	3	47.54	3.78	.07	-2.1	.07	-2.1	.99	.49	66.7	24.9	453
240	13	3	47.54	3.78	.31	-1.0	.30	-1.0	.71	.49	.0	24.9	461
261	13	3	47.54	3.78	.57	-.4	.53	-.5	.90	.49	33.3	24.9	580
266	13	3	47.54	3.78	.07	-2.1	.07	-2.1	.99	.49	66.7	24.9	585
30	12	3	46.07	3.87	.86	.1	.81	.0	.97	.48	.0	25.6	030
39	12	3	46.07	3.87	1.33	.6	1.22	.5	-.07	.48	33.3	25.6	039
88	12	3	46.07	3.87	.32	-1.0	.34	-.9	1.00	.48	33.3	25.6	217
89	12	3	46.07	3.87	.82	.0	.93	.2	-.44	.48	.0	25.6	218
105	12	3	46.07	3.87	1.95	1.2	2.24	1.5	-.44	.48	.0	25.6	234
111	12	3	46.07	3.87	2.67	1.8	2.78	1.8	-1.00	.48	33.3	25.6	240
192	12	3	46.07	3.87	2.22	1.4	2.02	1.3	-.56	.48	.0	25.6	413
204	12	3	46.07	3.87	2.00	1.3	1.83	1.1	.90	.48	.0	25.6	425
209	12	3	46.07	3.87	1.08	.4	.99	.2	.80	.48	33.3	25.6	430
212	12	3	46.07	3.87	.29	-1.0	.30	-1.0	.00	.48	33.3	25.6	433
218	12	3	46.07	3.87	.01	-2.9	.01	-2.9	1.00	.48	100.0	25.6	439
224	12	3	46.07	3.87	.23	-1.2	.26	-1.1	.56	.48	33.3	25.6	445
234	12	3	46.07	3.87	1.08	.4	.99	.2	.80	.48	33.3	25.6	455
242	12	3	46.07	3.87	.37	-.8	.34	-.9	.44	.48	33.3	25.6	463
244	12	3	46.07	3.87	2.30	1.5	2.22	1.4	.12	.48	.0	25.6	465
258	12	3	46.07	3.87	.32	-1.0	.34	-.9	1.00	.48	33.3	25.6	577
264	12	3	46.07	3.87	.38	-.8	.42	-.7	.83	.48	33.3	25.6	583
142	8	2	45.73	4.79	.00	-2.6	.00	-2.5	1.00	.55	100.0	26.0	271
5	11	3	44.52	4.01	1.70	1.0	1.55	.9	-.71	.46	.0	24.9	005
7	11	3	44.52	4.01	1.19	.5	1.42	.7	-.26	.46	.0	24.9	007
47	11	3	44.52	4.01	1.37	.7	1.23	.5	.90	.46	.0	24.9	047
49	11	3	44.52	4.01	.54	-.4	.56	-.4	.60	.46	.0	24.9	049
85	11	3	44.52	4.01	4.61	2.8	4.19	2.6	.90	.46	.0	24.9	214
91	11	3	44.52	4.01	1.24	.5	1.39	.7	-.83	.46	.0	24.9	220
102	11	3	44.52	4.01	.45	-.6	.43	-.7	.98	.46	.0	24.9	231
200	11	3	44.52	4.01	.61	-.3	.68	-.2	.39	.46	33.3	24.9	421
262	11	3	44.52	4.01	.67	-.2	.61	-.3	.77	.46	33.3	24.9	581
52	10	3	42.83	4.21	1.31	.6	1.44	.7	.31	.44	.0	26.9	052
80	10	3	42.83	4.21	.05	-2.1	.05	-2.0	.90	.44	100.0	26.9	080
81	10	3	42.83	4.21	3.90	2.4	3.93	2.4	.08	.44	.0	26.9	081
83	10	3	42.83	4.21	.87	.1	.77	.0	.90	.44	.0	26.9	212
86	10	3	42.83	4.21	.19	-1.3	.21	-1.2	.83	.44	33.3	26.9	215
99	10	3	42.83	4.21	2.78	1.8	2.51	1.6	-.77	.44	.0	26.9	228
150	10	3	42.83	4.21	.41	-.7	.37	-.8	-.07	.44	33.3	26.9	279
186	10	3	42.83	4.21	1.55	.9	1.88	1.1	-.60	.44	.0	26.9	407
214	10	3	42.83	4.21	2.34	1.5	2.20	1.4	-.07	.44	.0	26.9	435
239	10	3	42.83	4.21	.19	-1.3	.21	-1.2	.83	.44	33.3	26.9	460
257	10	3	42.83	4.21	1.35	.7	1.28	.6	-.90	.44	.0	26.9	576
185	9	3	40.94	4.48	.82	.1	1.02	.3	-.44	.42	.0	30.5	406
220	9	3	40.94	4.48	.47	-.5	.41	-.6	.90	.42	33.3	30.5	441
177	5	2	40.16	6.20	1.22	.6	1.24	.6	1.00	.21	.0	33.7	306
50	6	2	39.12	5.45	1.17	.5	1.21	.5	-1.00	.30	.0	27.6	050
8	8	3	38.78	4.84	1.32	.6	1.59	.9	-.83	.39	.0	31.7	008
43	8	3	38.78	4.84	.08	-1.6	.07	-1.7	.83	.39	100.0	31.7	043
46	8	3	38.78	4.84	1.21	.5	1.07	.4	.77	.39	33.3	31.7	046
57	8	3	38.78	4.84	.65	-.1	.62	-.2	-.90	.39	33.3	31.7	057

70	8	3	38.78	4.84	.44	-.5	.55	-.3	.97	.39	33.3	31.7	070
72	8	3	38.78	4.84	.85	.1	.94	.2	.39	.39	.0	31.7	072
73	8	3	38.78	4.84	.28	-.9	.36	-.7	.07	.39	33.3	31.7	073
94	8	3	38.78	4.84	.44	-.5	.55	-.3	.97	.39	33.3	31.7	223
116	8	3	38.78	4.84	.81	.1	.96	.3	.60	.39	33.3	31.7	245
124	8	3	38.78	4.84	2.52	1.5	2.23	1.3	.90	.39	.0	31.7	253
126	8	3	38.78	4.84	.18	-1.2	.15	-1.3	.90	.39	33.3	31.7	255
128	8	3	38.78	4.84	.18	-1.2	.15	-1.3	.90	.39	33.3	31.7	257
131	8	3	38.78	4.84	1.21	.5	1.07	.4	.77	.39	33.3	31.7	260
137	8	3	38.78	4.84	1.01	.3	.97	.3	.98	.39	.0	31.7	266
206	8	3	38.78	4.84	.08	-1.6	.07	-1.7	.83	.39	100.0	31.7	427
270	8	3	38.78	4.84	.28	-.9	.36	-.7	.07	.39	33.3	31.7	589
34	7	3	36.19	5.36	.40	-.5	.51	-.3	.83	.35	33.3	36.0	034
87	7	3	36.19	5.36	.02	-2.0	.03	-2.0	.90	.35	100.0	36.0	216
122	7	3	36.19	5.36	1.95	1.1	2.52	1.5	-.60	.35	.0	36.0	251
125	7	3	36.19	5.36	.40	-.5	.51	-.3	.83	.35	33.3	36.0	254
129	7	3	36.19	5.36	.02	-2.0	.03	-2.0	.90	.35	100.0	36.0	258
133	7	3	36.19	5.36	3.44	2.0	3.42	2.0	-.07	.35	.0	36.0	262
202	7	3	36.19	5.36	3.44	2.0	3.42	2.0	-.07	.35	.0	36.0	423
203	7	3	36.19	5.36	.39	-.5	.35	-.7	-.07	.35	33.3	36.0	424
60	6	3	32.92	6.14	2.54	1.4	2.53	1.4	-.07	.31	.0	38.1	060
92	6	3	32.92	6.14	1.43	.7	1.55	.8	-1.00	.31	33.3	38.1	221
98	6	3	32.92	6.14	1.43	.7	1.27	.6	.90	.31	.0	38.1	227
121	6	3	32.92	6.14	.49	-.3	.50	-.3	.56	.31	33.3	38.1	250
160	6	3	32.92	6.14	1.23	.6	1.08	.4	-.56	.31	33.3	38.1	289
208	6	3	32.92	6.14	.30	-.7	.34	-.6	1.00	.31	33.3	38.1	429
213	6	3	32.92	6.14	1.43	.7	1.27	.6	.90	.31	.0	38.1	434
227	6	3	32.92	6.14	.30	-.7	.34	-.6	1.00	.31	33.3	38.1	448
238	6	3	32.92	6.14	.11	-1.3	.11	-1.3	.00	.31	100.0	38.1	459
268	6	3	32.92	6.14	.66	-.1	.76	.1	.44	.31	33.3	38.1	587
61	4	2	32.50	7.52	.46	-.2	.52	-.1	1.00	.36	.0	37.2	061
29	5	3	28.39	7.44	.19	-.9	.25	-.7	.83	.26	100.0	46.6	029
42	5	3	28.39	7.44	1.67	.9	1.66	.9	-.07	.26	33.3	46.6	042
58	5	3	28.39	7.44	1.67	.9	1.66	.9	-.07	.26	33.3	46.6	058
197	5	3	28.39	7.44	1.67	.9	1.66	.9	-.07	.26	33.3	46.6	418
199	5	3	28.39	7.44	1.67	.9	1.66	.9	-.07	.26	33.3	46.6	420
267	5	3	28.39	7.44	.39	-.4	.49	-.3	.07	.26	33.3	46.6	586
269	3	2	24.81	10.45	1.04	.5	1.19	.6	-1.00	.27	.0	55.1	588
12	3	2	23.47	10.44	.81	.3	.85	.3	-1.00	.16	.0	51.4	012
77	4	3	20.88	10.34	1.01	.4	1.29	.6	-.83	.19	66.7	71.4	077
134	4	3	20.88	10.34	.44	-.2	.39	-.2	.90	.19	66.7	71.4	263
139	4	3	20.88	10.34	.81	.3	.81	.3	-.07	.19	66.7	71.4	268
183	4	3	20.88	10.34	1.01	.4	1.29	.6	-.83	.19	66.7	71.4	404
198	4	3	20.88	10.34	.81	.3	.81	.3	-.07	.19	66.7	71.4	419
216	4	3	20.88	10.34	1.01	.4	1.29	.6	-.83	.19	66.7	71.4	437
95	2	2	15.18	18.58	MINIMUM MEASURE				.00	.00	100.0	100.0	224
17	3	3	8.35	18.52	MINIMUM MEASURE				.00	.00	100.0	100.0	017
106	3	3	8.35	18.52	MINIMUM MEASURE				.00	.00	100.0	100.0	235
127	3	3	8.35	18.52	MINIMUM MEASURE				.00	.00	100.0	100.0	256
135	3	3	8.35	18.52	MINIMUM MEASURE				.00	.00	100.0	100.0	264
138	3	3	8.35	18.52	MINIMUM MEASURE				.00	.00	100.0	100.0	267
156	3	3	8.35	18.52	MINIMUM MEASURE				.00	.00	100.0	100.0	285
210	3	3	8.35	18.52	MINIMUM MEASURE				.00	.00	100.0	100.0	431
211	3	3	8.35	18.52	MINIMUM MEASURE				.00	.00	100.0	100.0	432
MEAN	14.5	2.9	49.09	5.47	.98	-.1	.99	-.1			29.0	30.0	
S.D.	6.3	.3	14.10	3.49	.92	1.2	.91	1.2			29.9	9.5	

Student measures of Grade 6 HK student essays

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Student
33	27	3	93.25	18.60	MAXIMUM MEASURE				.00	.00	100.0	100.0	033
43	27	3	93.25	18.60	MAXIMUM MEASURE				.00	.00	100.0	100.0	043
136	27	3	93.25	18.60	MAXIMUM MEASURE				.00	.00	100.0	100.0	136
23	26	3	80.50	10.50	.68	.1	.64	.0	.37	.12	66.7	70.8	023
64	26	3	80.50	10.50	.62	.0	.57	.0	.62	.12	66.7	70.8	064
76	26	3	80.50	10.50	.68	.1	.64	.0	.37	.12	66.7	70.8	076
3	25	3	72.66	7.65	1.45	.7	1.37	.7	.37	.16	33.3	46.6	003
7	25	3	72.66	7.65	1.45	.7	1.37	.7	.37	.16	33.3	46.6	007
10	25	3	72.66	7.65	.55	-.2	.54	-.2	-.62	.16	33.3	46.6	010
15	25	3	72.66	7.65	.22	-.9	.25	-.8	.99	.16	100.0	46.6	015
26	25	3	72.66	7.65	1.45	.7	1.37	.7	.37	.16	33.3	46.6	026
95	25	3	72.66	7.65	.49	-.3	.51	-.3	-.37	.16	33.3	46.6	095
130	25	3	72.66	7.65	.22	-.9	.25	-.8	.99	.16	100.0	46.6	130
31	24	3	67.83	6.36	2.26	1.3	2.14	1.2	.37	.19	.0	41.4	031
37	24	3	67.83	6.36	.58	-.2	.62	-.2	.79	.19	33.3	41.4	037
85	24	3	67.83	6.36	1.18	.5	1.30	.6	-.93	.19	33.3	41.4	085
96	24	3	67.83	6.36	2.26	1.3	2.14	1.2	.37	.19	.0	41.4	096
114	24	3	67.83	6.36	1.18	.5	1.30	.6	-.93	.19	33.3	41.4	114
51	16	2	67.29	7.79	.06	-1.2	.06	-1.2	.00	.23	100.0	41.1	051
38	16	2	66.81	7.79	.04	-1.3	.04	-1.3	.00	.20	100.0	41.5	038
12	23	3	64.30	5.57	.19	-1.1	.17	-1.2	.37	.22	33.3	36.4	012
16	23	3	64.30	5.57	.48	-.4	.55	-.3	.99	.22	33.3	36.4	016
21	23	3	64.30	5.57	3.08	1.8	2.93	1.7	.37	.22	.0	36.4	021
131	23	3	64.30	5.57	.13	-1.3	.11	-1.4	.62	.22	100.0	36.4	131
6	22	3	61.51	5.02	.30	-.8	.33	-.7	-.37	.24	33.3	33.8	006
68	22	3	61.51	5.02	.77	.0	.86	.1	.89	.24	33.3	33.8	068
104	22	3	61.51	5.02	.36	-.7	.37	-.6	-.62	.24	33.3	33.8	104
24	21	3	59.20	4.62	.75	.0	.86	.1	.99	.26	.0	29.4	024
39	21	3	59.20	4.62	.07	-1.7	.08	-1.7	.00	.26	100.0	29.4	039
102	21	3	59.20	4.62	.23	-1.1	.23	-1.1	.79	.26	33.3	29.4	102
128	21	3	59.20	4.62	2.57	1.6	2.54	1.6	.54	.26	.0	29.4	128
36	20	3	57.20	4.32	.94	.2	.85	.1	.05	.28	33.3	26.5	036
56	20	3	57.20	4.32	.73	-.1	.77	.0	-.37	.28	33.3	26.5	056
97	20	3	57.20	4.32	.14	-1.5	.13	-1.5	.37	.28	66.7	26.5	097
18	19	3	55.44	4.09	1.46	.8	1.33	.6	.14	.30	33.3	27.3	018
19	19	3	55.44	4.09	.89	.1	.87	.1	-.05	.30	.0	27.3	019
32	19	3	55.44	4.09	1.06	.3	1.20	.5	.99	.30	.0	27.3	032
35	19	3	55.44	4.09	.34	-.9	.36	-.8	-.62	.30	33.3	27.3	035
66	19	3	55.44	4.09	.28	-1.0	.25	-1.1	.62	.30	33.3	27.3	066
84	19	3	55.44	4.09	.89	.1	.87	.1	-.05	.30	.0	27.3	084
127	19	3	55.44	4.09	3.89	2.4	3.74	2.3	.72	.30	.0	27.3	127
9	18	3	53.84	3.92	3.32	2.2	3.32	2.2	.61	.31	.0	26.6	009
14	18	3	53.84	3.92	.47	-.6	.45	-.6	-.14	.31	.0	26.6	014
17	18	3	53.84	3.92	.13	-1.6	.13	-1.7	.79	.31	66.7	26.6	017
41	18	3	53.84	3.92	.13	-1.6	.13	-1.7	.79	.31	66.7	26.6	041
42	18	3	53.84	3.92	2.10	1.3	1.94	1.2	.19	.31	33.3	26.6	042
71	18	3	53.84	3.92	2.10	1.3	1.94	1.2	.19	.31	33.3	26.6	071
80	18	3	53.84	3.92	5.04	3.0	5.60	3.3	-.99	.31	.0	26.6	080
119	18	3	53.84	3.92	.07	-2.0	.07	-2.0	.93	.31	66.7	26.6	119
120	18	3	53.84	3.92	1.80	1.1	1.66	1.0	.46	.31	33.3	26.6	120
44	17	3	52.34	3.80	.42	-.7	.46	-.6	-.99	.32	.0	23.4	044
72	17	3	52.34	3.80	.42	-.7	.46	-.6	-.99	.32	.0	23.4	072
94	17	3	52.34	3.80	.14	-1.7	.13	-1.7	.37	.32	66.7	23.4	094
137	17	3	52.34	3.80	.31	-1.0	.30	-1.0	.84	.32	.0	23.4	137
49	16	3	50.92	3.73	.98	.2	1.01	.3	.97	.33	33.3	24.0	049
70	16	3	50.92	3.73	1.23	.5	1.17	.5	.14	.33	33.3	24.0	070
79	16	3	50.92	3.73	5.53	3.4	5.85	3.5	-1.00	.33	.0	24.0	079
92	16	3	50.92	3.73	2.72	1.9	2.80	1.9	-.88	.33	33.3	24.0	092
100	16	3	50.92	3.73	3.85	2.6	4.04	2.7	-.98	.33	.0	24.0	100
106	16	3	50.92	3.73	.34	-1.0	.32	-1.0	.37	.33	33.3	24.0	106
117	16	3	50.92	3.73	.23	-1.3	.22	-1.4	.62	.33	33.3	24.0	117
133	16	3	50.92	3.73	.80	.0	.79	-.1	-.05	.33	.0	24.0	133
46	15	3	49.55	3.69	.63	-.3	.63	-.3	.79	.34	.0	23.6	046
48	15	3	49.55	3.69	.52	-.6	.52	-.6	.93	.34	33.3	23.6	048
50	15	3	49.55	3.69	.54	-.5	.53	-.5	.62	.34	33.3	23.6	050
53	15	3	49.55	3.69	.13	-1.8	.14	-1.8	.00	.34	33.3	23.6	053
78	15	3	49.55	3.69	1.82	1.2	1.86	1.2	-.79	.34	.0	23.6	078
111	15	3	49.55	3.69	.11	-1.9	.10	-1.9	.79	.34	33.3	23.6	111
115	15	3	49.55	3.69	.35	-1.0	.35	-1.0	.14	.34	33.3	23.6	115
116	15	3	49.55	3.69	.13	-1.8	.14	-1.8	.00	.34	33.3	23.6	116

13	14	3	48.18	3.69	1.53	.9	1.54	.9	.98	.35	33.3	22.9	013
29	14	3	48.18	3.69	2.47	1.7	2.48	1.7	-.99	.35	.0	22.9	029
63	14	3	48.18	3.69	.28	-1.2	.28	-1.2	.84	.35	.0	22.9	063
88	14	3	48.18	3.69	.09	-2.0	.09	-2.0	.62	.35	66.7	22.9	088
91	14	3	48.18	3.69	2.74	1.9	2.72	1.9	-.05	.35	.0	22.9	091
103	14	3	48.18	3.69	1.95	1.3	1.96	1.3	-.79	.35	33.3	22.9	103
112	14	3	48.18	3.69	.14	-1.7	.14	-1.7	.37	.35	66.7	22.9	112
118	14	3	48.18	3.69	1.04	.3	1.04	.3	.62	.35	.0	22.9	118
121	14	3	48.18	3.69	1.01	.3	1.02	.3	.88	.35	.0	22.9	121
126	14	3	48.18	3.69	1.53	.9	1.54	.9	.98	.35	33.3	22.9	126
20	13	3	46.80	3.73	.73	-.2	.71	-.2	.99	.35	.0	23.5	020
45	13	3	46.80	3.73	.20	-1.4	.20	-1.4	.98	.35	33.3	23.5	045
47	13	3	46.80	3.73	2.27	1.5	2.26	1.5	.96	.35	33.3	23.5	047
60	13	3	46.80	3.73	.26	-1.3	.25	-1.3	.89	.35	33.3	23.5	060
62	13	3	46.80	3.73	.26	-1.3	.25	-1.3	.89	.35	33.3	23.5	062
65	13	3	46.80	3.73	.34	-1.0	.35	-1.0	.37	.35	33.3	23.5	065
98	13	3	46.80	3.73	.01	-3.1	.01	-3.1	.99	.35	100.0	23.5	098
2	12	3	45.38	3.81	1.13	.4	1.09	.4	1.00	.34	.0	22.9	002
27	12	3	45.38	3.81	2.27	1.5	2.17	1.4	.99	.34	.0	22.9	027
69	12	3	45.38	3.81	.74	-.1	.78	-.1	.37	.34	33.3	22.9	069
113	12	3	45.38	3.81	1.66	1.0	1.66	1.0	.93	.34	33.3	22.9	113
122	12	3	45.38	3.81	.13	-1.7	.13	-1.7	.00	.34	66.7	22.9	122
124	12	3	45.38	3.81	.59	-.4	.62	-.3	.62	.34	33.3	22.9	124
4	11	3	43.88	3.94	.46	-.6	.43	-.7	-.99	.34	.0	26.3	004
28	11	3	43.88	3.94	1.80	1.1	1.91	1.2	-.41	.34	.0	26.3	028
30	11	3	43.88	3.94	.65	-.3	.62	-.3	1.00	.34	33.3	26.3	030
105	11	3	43.88	3.94	.65	-.3	.62	-.3	1.00	.34	33.3	26.3	105
107	11	3	43.88	3.94	.66	-.2	.71	-.2	.33	.34	33.3	26.3	107
110	11	3	43.88	3.94	.09	-1.9	.09	-1.9	.62	.34	66.7	26.3	110
134	11	3	43.88	3.94	.65	-.3	.62	-.3	1.00	.34	33.3	26.3	134
34	10	3	42.26	4.12	.39	-.8	.43	-.6	.37	.32	33.3	26.9	034
54	10	3	42.26	4.12	.32	-.9	.30	-1.0	-.37	.32	33.3	26.9	054
55	10	3	42.26	4.12	.39	-.8	.43	-.6	.37	.32	33.3	26.9	055
74	10	3	42.26	4.12	2.41	1.6	2.27	1.5	-1.00	.32	.0	26.9	074
81	10	3	42.26	4.12	1.02	.3	.95	.2	-.99	.32	.0	26.9	081
108	10	3	42.26	4.12	.94	.2	1.02	.3	.79	.32	.0	26.9	108
135	10	3	42.26	4.12	.36	-.8	.35	-.8	.89	.32	33.3	26.9	135
8	9	3	40.46	4.37	.12	-1.5	.13	-1.5	.93	.31	66.7	26.1	008
52	9	3	40.46	4.37	.10	-1.6	.10	-1.7	.00	.31	66.7	26.1	052
59	9	3	40.46	4.37	1.52	.8	1.54	.8	-.37	.31	.0	26.1	059
123	9	3	40.46	4.37	3.00	1.9	3.20	2.0	-.19	.31	.0	26.1	123
138	9	3	40.46	4.37	.80	.0	.74	-.1	-.79	.31	33.3	26.1	138
40	8	3	38.40	4.73	.10	-1.5	.11	-1.5	.62	.29	66.7	33.6	040
87	8	3	38.40	4.73	.63	-.2	.69	-.1	.84	.29	.0	33.6	087
58	7	3	35.93	5.24	.76	.0	.73	.0	.98	.26	33.3	35.1	058
61	7	3	35.93	5.24	.03	-2.0	.02	-2.0	.99	.26	100.0	35.1	061
75	7	3	35.93	5.24	1.26	.6	1.15	.5	-.99	.26	.0	35.1	075
77	7	3	35.93	5.24	.53	-.3	.62	-.2	.62	.26	33.3	35.1	077
73	6	3	32.79	6.01	1.54	.8	1.39	.7	.99	.23	.0	38.0	073
83	6	3	32.79	6.01	1.54	.8	1.39	.7	.99	.23	.0	38.0	083
90	6	3	32.79	6.01	.73	.0	.84	.2	.14	.23	33.3	38.0	090
99	6	3	32.79	6.01	.41	-.4	.43	-.4	.93	.23	33.3	38.0	099
101	6	3	32.79	6.01	2.50	1.4	2.63	1.5	-.37	.23	.0	38.0	101
67	5	3	28.43	7.31	1.65	.9	1.74	.9	-.37	.19	.0	35.2	067
82	5	3	28.43	7.31	1.01	.4	.91	.3	.99	.19	.0	35.2	082
86	5	3	28.43	7.31	.62	-.1	.56	-.1	-.99	.19	66.7	35.2	086
93	5	3	28.43	7.31	1.01	.4	.91	.3	.99	.19	.0	35.2	093
129	5	3	28.43	7.31	.26	-.7	.30	-.6	.62	.19	66.7	35.2	129
1	4	3	21.13	10.22	.81	.3	.85	.3	-.37	.14	66.7	71.5	001
22	4	3	21.13	10.22	.49	-.1	.44	-.2	.99	.14	66.7	71.5	022
57	4	3	21.13	10.22	.49	-.1	.44	-.2	.99	.14	66.7	71.5	057
89	4	3	21.13	10.22	.49	-.1	.44	-.2	.99	.14	66.7	71.5	089
109	4	3	21.13	10.22	.49	-.1	.44	-.2	.99	.14	66.7	71.5	109
125	4	3	21.13	10.22	.49	-.1	.44	-.2	.99	.14	66.7	71.5	125
132	4	3	21.13	10.22	.81	.3	.85	.3	-.37	.14	66.7	71.5	132
25	1	1	17.06	18.63	MINIMUM MEASURE				.00	.00	100.0	100.0	025
5	2	2	12.25	18.49	MINIMUM MEASURE				.00	.00	100.0	100.0	005
11	3	3	8.79	18.43	MINIMUM MEASURE				.00	.00	100.0	100.0	011
MEAN	14.7	3.0	49.56	5.68	.98	-.1	.98	-.1			32.6	32.5	
S.D.	6.5	.2	15.35	3.33	1.00	1.2	1.02	1.2			28.9	13.0	

Appendix 16

```
+-----------------------------------------------------------+
|Measr|-Prompt|-US Rater                    |+Student|R6SCA|
|-----+-------+-----------------------------+--------+-----|
|   6 +       +                             +        + (6) |
|     |       |                             | *      |     |
|     |       |                             |        | --- |
|   5 +       +                             +        +     |
|     |       |                             |        |     |
|     |       |                             |        |     |
|   4 +       +                             +        +  5  |
|     |       |                             |        |     |
|     |       |                             | *      |     |
|   3 +       +                             + **     +     |
|     |       |                             | **     | --- |
|     |       |                             | *      |     |
|   2 +       +                             + *****  +     |
|     |       | US211                       | ***    |     |
|     |       | US200                       | ***    |  4  |
|   1 +       +                             +        +     |
|     |       | US104 US209                 |        |     |
|     | 6-4BP | US103 US204                 | *      |     |
*   0 * 5-4BN * US105 US203 US206 US207     * **     * --- *
|     | 4-4BI | US102 US208 US210           | **     |     |
|     |       | US101 US106 US201 US202     | ***    |     |
|  -1 +       +                             + *      +     |
|     |       |                             | ****   |     |
|     |       | US205                       | ***    |  3  |
|  -2 +       +                             +        +     |
|     |       |                             | ***    |     |
|     |       |                             | *****  |     |
|  -3 +       +                             + **     + --- |
|     |       |                             |        |     |
|     |       |                             | **     |     |
|  -4 +       +                             +        +  2  |
|     |       |                             | **     |     |
|     |       |                             |        |     |
|  -5 +       +                             +        + --- |
|     |       |                             | *      |     |
|     |       |                             |        |     |
|  -6 +       +                             +        +     |
|     |       |                             |        |     |
|     |       |                             |        |     |
|  -7 +       +                             + *      +     |
|     |       |                             |        |     |
|     |       |                             |        |     |
|  -8 +       +                             +        + (1) |
|-----+-------+-----------------------------+--------+-----|
|Measr|-Prompt|-US Rater                    | * = 1  |R6SCA|
+-----------------------------------------------------------+
```

Vertical map for a sub-sample of US student essays with scoring of US raters

Calibration of the students for sub-sample of US student essays using US raters

	Student		Observed	Measure	S.E.	Infit		Outfit	
	Number	Group	score	(Logit)		MnSq	ZStd	MnSq	ZStd
7	US006	304	3.9	1.19	0.48	0.30	-2.1	0.31	-2.1
15	US013	304	4.4	2.54	0.47	0.94	0	0.96	0
20	US017	304	1.4	-5.31	0.52	1.60	1.3	1.30	0.6
25	US021	304	3.1	-1.43	0.49	2.08	2.0	2.17	2.1
27	US022	306	3.1	-1.67	0.49	0.84	-0.2	0.81	-0.2

32	US026	306	4.1	1.19	0.48	0.56	-1.1	0.56	-1.1
34	US027	306	4.4	2.09	0.47	1.03	0.2	1.05	0.2
38	US030	306	3.5	-0.45	0.49	0.69	-0.7	0.70	-0.6
43	US034	306	3.1	-1.67	0.49	0.59	-0.8	0.59	-0.8
56	US046	304	2.5	-2.80	0.44	0.93	0	0.97	0
60	US049	304	2.6	-2.61	0.45	1.97	2.0	1.94	1.9
69	US057	304	3.5	-0.55	0.60	1.04	0.2	1.03	0.2
73	US060	304	2.2	-3.72	0.41	2.00	2.4	2.02	2.4
78	US064	306	4.2	1.63	0.48	0.77	-0.4	0.77	-0.4
85	US070	306	2.7	-2.62	0.45	0.42	-1.6	0.41	-1.6
228	US212	304	3.4	-0.62	0.49	0.62	-0.9	0.61	-1.0
232	US215	304	3.7	0.09	0.49	0.55	-1.2	0.55	-1.2
236	US218	304	4.3	1.94	0.47	1.21	0.6	1.21	0.6
246	US227	304	2.8	-2.43	0.57	0.39	-1.3	0.36	-1.3
255	US235	306	2.4	-2.97	0.43	0.50	-1.5	0.50	-1.4
257	US236	306	2.7	-2.39	0.45	0.58	-1.0	0.60	-0.9
259	US237	306	3.0	-1.49	0.49	1.10	0.3	1.12	0.4
264	US241	306	4.3	1.68	0.47	1.26	0.7	1.26	0.7
267	US243	306	4.2	1.46	0.48	0.81	-0.3	0.82	-0.3
271	US246	306	3.5	-0.41	0.49	1.41	1.0	1.44	1.1
274	US248	306	4.8	3.21	0.47	2.65	3.0	2.64	3.0
283	US256	304	1.9	-4.37	0.52	0.65	-0.8	0.65	-0.8
294	US266	304	2.0	-4.19	0.42	0.98	0	0.97	0
302	US273	304	4.2	1.51	0.48	0.87	-0.2	0.87	-0.2
315	US285	304	1.1	-7.13	1.04	0.90	0.2	0.74	0
435	US404	306	2.6	-3.03	0.44	0.50	-1.4	0.51	-1.3
451	US419	304	2.7	-2.46	0.46	0.33	-1.9	0.32	-1.9
453	US420	304	2.2	-3.72	0.42	0.63	-1.1	0.67	-0.9
461	US427	304	3.3	-0.94	0.49	1.22	0.6	1.15	0.4
464	US429	304	2.6	-2.77	0.45	0.33	-2.0	0.32	-2.0
472	US436	304	4.3	2.09	0.47	1.64	1.4	1.66	1.5
476	US439	306	3.6	0.32	0.49	1.38	0.9	1.40	1.0
478	US440	306	5.5	5.63	0.54	0.69	-0.7	0.74	-0.6
482	US443	306	4.5	2.85	0.47	0.96	0	0.97	0
488	US448	306	2.5	-2.65	0.44	0.36	-2.0	0.35	-2.0
491	US450	306	3.4	-0.15	0.49	1.60	1.3	1.60	1.3
496	US454	306	4.4	2.63	0.47	1.31	0.8	1.31	0.8
502	US459	306	3.3	-0.56	0.49	1.61	1.3	1.63	1.3

504	US460	306	3.0	-1.28	0.49	1.00	0.1	1.03	0.2
507	US462	306	4.2	2.03	0.48	0.97	0	0.98	0
509	US463	306	4.3	2.26	0.47	0.60	-1.0	0.59	-1.0
513	US466	306	4.6	3.14	0.47	1.00	0.1	1.02	0.1
518	US470	306	4.2	2.03	0.48	0.51	-1.3	0.51	-1.3
626	US577	304	3.1	-1.43	0.50	1.33	0.7	1.33	0.7
636	US586	304	3.0	-1.68	0.49	1.07	0.2	1.06	0.2

Separation: 5.08 Reliability: 0.96

Fixed (all same) chi-square: 1317.7 significance *p*: .00

Note: 304 = Grade 4 US sub-sample student, 306 = Grade 6 US sub-sample student

Calibration of the students for a sub-sample of US student essays using US raters (9 misfitting students)

Student			Observed	Measure	S.E.	Infit		Outfit	
Number		Group	score	(Logit)		MnSq	ZStd	MnSq	ZStd
271	US246	306	3.5	-0.41	0.49	1.41	1.0	1.44	1.1
20	US017	304	1.4	-5.31	0.52	1.60	1.3	1.30	0.6
491	US450	306	3.4	-0.15	0.49	1.60	1.3	1.60	1.3
502	US459	306	3.3	-0.56	0.49	1.61	1.3	1.63	1.3
472	US436	304	4.3	2.09	0.47	1.64	1.4	1.66	1.5
60	US049	304	2.6	-2.61	0.45	1.97	2.0	1.94	1.9
73	US060	304	2.2	-3.72	0.41	2.00	2.4	2.02	2.4
25	US021	304	3.1	-1.43	0.49	2.08	2.0	2.17	2.1
274	US248	306	4.8	3.21	0.47	2.65	3.0	2.64	3.0

Separation: 5.08 Reliability: 0.96

Fixed (all same) chi-square: 1317.7 significance *p*: .00

Note: 304 = Grade 4 US sub-sample student, 306 = Grade 6 US sub-sample student

Calibration of the prompts for a sub-sample of US student essays using US raters

Prompt	Observed	Measure	S.E.	Infit		Outfit	
	score	(Logit)		MnSq	ZStd	MnSq	ZStd
4BI	3.5	-0.31	0.12	1.00	0	1.00	0
4BN	3.4	-0.12	0.12	0.73	-2.8	0.73	-2.7
4BP	3.2	0.43	0.12	1.24	2.2	1.25	2.2

Separation: 2.50 Reliability: 0.86

Fixed (all same) chi-square: 21.6 significance *p*: .00

Calibration of the US raters for a sub-sample of US student essays using US raters

US rater	Observed score	Measure (Logit)	S.E.	Infit		Outfit	
				MnSq	ZStd	MnSq	ZStd
USR101	3.6	-0.60	0.31	0.99	0	1.01	0.1
USR102	3.5	-0.43	0.26	0.92	-0.3	0.90	-0.3
USR103	3.4	0.21	0.31	0.61	-1.5	0.62	-1.4
USR104	3.1	0.61	0.29	0.47	-2.4	0.48	-2.3
USR105	3.4	-0.13	0.26	0.59	-2.0	0.55	-2.2
USR106	3.7	-0.66	0.30	0.85	-0.5	0.85	-0.5
USR200	3.1	1.40	0.28	1.28	1.1	1.30	1.1
USR201	3.5	-0.52	0.28	1.12	0.5	1.10	0.4
USR202	3.2	-0.66	0.40	1.41	1.1	1.42	1.1
USR203	3.3	0.09	0.20	1.18	0.9	1.21	1.1
USR204	3.5	0.29	0.25	0.91	-0.3	0.90	-0.3
USR205	3.8	-1.53	0.33	0.97	0	0.94	-0.1
USR206	3.3	-0.08	0.24	1.27	1.2	1.29	1.2
USR207	3.1	0.07	0.30	0.93	-0.1	0.94	-0.1
USR208	3.1	-0.21	0.41	1.15	0.5	1.10	0.3
USR209	3.0	0.77	0.32	1.26	0.9	1.28	0.9
USR210	3.7	-0.42	0.24	0.99	0	0.98	0
USR211	2.5	1.80	0.90	0.89	0.1	0.83	0.2
Separation: 1.90				Reliability: 0.78			
Fixed (all same) chi-square: 84.0				significance *p*: .00			

```
+---------------------------------------------------------------------------------+
|Measr|-Prompt        |-US Rater              |-Lexile Analyzer|-IEA|+Student|R6SCA|
|-----+---------------+-----------------------+----------------+----+--------+-----|
|   5 +               +                       +                +    +        + (6) |
|     |               |                       |                |    |        |     |
|     |               |                       |                |    | *      |     |
|     |               |                       |                |    |        |     |
|     |               |                       |                |    |        |     |
|   4 +               +                       +                +    +        +     |
|     |               |                       |                |    | *      |     |
|     |               |                       |                |    |        |     |
|     |               |                       |                |    | *      | --- |
|     |               |                       |                |    |        |     |
|   3 +               +                       +                +    +        +     |
|     |               |                       |                |    | **     |     |
|     |               |                       |                |    | **     |     |
|     |               |                       |                |    |        |     |
|     |               |                       | 1              |    |        |     |
|   2 +               +                       +                +    +        +     |
|     |               | US200                 |                | 1  | *****  |     |
|     |               |                       | 2              | 2  |        |  4  |
|     |               | US211                 |                |    | **     |     |
|     |               |                       |                | 4  | ****   |     |
|   1 +               + US104                 +                +    + ****   +     |
|     |               | US209                 |                | 3  | **     |     |
|     |               | US103                 |                |    | *      |     |
|     |               |                       | 3              |    | ***    |     |
|     | 5-4BN  6-4BP  | US105  US203  US204   |                |    | **     |     |
*   0 *               * US208                 *                *    * *      *     *
|     |               | US206  US207          | 5              |    | **     |     |
|     | 4-4BI         | US102  US106          |                |    | ***    | --- |
|     |               | US101                 | 8 9            | 7  | *      |     |
|     |               | US201  US210          |                | 5  |        |     |
|  -1 +               + US202                 + 4              +    +        +     |
|     |               |                       |                |    | *****  |     |
|     |               |                       |                |    | *      |     |
|     |               |                       |                |    | *      |     |
|     |               |                       |                | 8  | ***    |     |
|  -2 +               + US205                 + 6              +    +        +     |
|     |               |                       |                | 6  | *      |  3  |
|     |               |                       |                |    |        |     |
|     |               |                       |                |    | *      |     |
|     |               |                       |                |    |        |     |
|  -3 +               +                       +                +    +        +     |
|     |               |                       |                |    |        |     |
|     |               |                       |                |    |        |     |
|     |               |                       |                |    |        |     |
|     |               |                       |                |    |        |     |
|  -4 +               +                       +                +    +        + --- |
|     |               |                       |                |    |        |     |
|     |               |                       |                |    | *      |     |
|     |               |                       |                |    |        |     |
|     |               |                       |                |    |        |     |
|  -5 +               +                       +                +    +        + (1) |
|-----+---------------+-----------------------+----------------+----+--------+-----|
|Measr|-Prompt        |-US Rater              |-Lexile Analyzer|-IEA| * = 1  |R6SCA|
+---------------------------------------------------------------------------------+
```

Vertical map for a sub-sample of US student essays with scoring of US raters, Lexile Analyzer and IEA

Calibration of the students for a sub-sample of US student essays using US raters, Lexile Analyzer and IEA

Student			Observed	Measure	S.E.	Infit		Outfit	
Number		Group	score	(Logit)		MnSq	ZStd	MnSq	ZStd
7	US006	304	3.9	1.06	0.58	0.24	-2.3	0.21	-2.4
15	US013	304	4.4	2.71	0.56	0.72	-0.6	0.74	-0.5
20	US017	304	1.4	-2.62	0.54	2.15	2.2	2.05	1.7
25	US021	304	3.1	-1.11	0.52	1.41	1.0	1.31	0.8
27	US022	306	3.1	-1.70	0.55	1.03	0.2	1.00	0.1
32	US026	306	4.1	0.95	0.57	0.94	0	0.93	0
34	US027	306	4.4	1.89	0.55	0.77	-0.5	0.75	-0.5
38	US030	306	3.5	0.89	0.56	1.10	0.3	1.11	0.3
43	US034	306	3.1	0.21	0.53	1.20	0.5	1.28	0.7
56	US046	304	2.5	-1.28	0.49	0.72	-0.6	0.73	-0.5
60	US049	304	2.6	-1.79	0.48	1.60	1.5	1.46	1.2
69	US057	304	3.5	1.39	0.69	0.95	0	0.97	0.1
73	US060	304	2.2	-1.87	0.47	1.02	0.1	0.91	-0.1
78	US064	306	4.2	0.52	0.56	1.10	0.3	1.16	0.4
85	US070	306	2.7	-0.53	0.50	0.50	-1.2	0.48	-1.2
228	US212	304	3.4	0.93	0.56	0.66	-0.7	0.64	-0.8
232	US215	304	3.7	1.76	0.56	0.70	-0.7	0.68	-0.7
236	US218	304	4.3	2.67	0.56	1.52	1.1	1.48	1.1
246	US227	304	2.8	-0.17	0.63	0.30	-1.6	0.27	-1.6
255	US235	306	2.4	0.02	0.46	0.44	-1.7	0.45	-1.6
257	US236	306	2.7	-1.24	0.51	1.48	1.1	1.51	1.1
259	US237	306	3.0	-1.56	0.56	1.49	1.0	1.55	1.1
264	US241	306	4.3	1.11	0.57	0.69	-0.6	0.69	-0.6
267	US243	306	4.2	1.21	0.57	1.24	0.6	1.29	0.7
271	US246	306	3.5	-0.48	0.55	1.18	0.5	1.16	0.5
274	US248	306	4.8	3.76	0.56	1.79	1.7	1.81	1.7
283	US256	304	1.9	-1.49	0.55	1.13	0.4	1.09	0.3
294	US266	304	2.0	-2.20	0.46	1.03	0.1	0.98	0
302	US273	304	4.2	1.19	0.56	0.54	-1.1	0.55	-1.1
315	US285	304	1.1	-4.39	1.06	1.05	0.3	0.95	0.3
435	US404	306	2.6	-0.34	0.49	0.67	-0.7	0.65	-0.8
451	US419	304	2.7	0.43	0.51	0.51	-1.2	0.54	-1.0
453	US420	304	2.2	-1.11	0.45	0.49	-1.6	0.53	-1.4
461	US427	304	3.3	0.89	0.55	1.86	1.7	1.66	1.4

464	US429	304	2.6	-0.20	0.50	0.32	-2.0	0.29	-2.1
472	US436	304	4.3	1.87	0.56	1.31	0.8	1.24	0.6
476	US439	306	3.6	1.27	0.54	1.36	0.9	1.36	0.9
478	US440	306	5.5	4.88	0.57	0.72	-0.9	0.72	-0.9
482	US443	306	4.5	3.34	0.56	0.98	0	0.98	0.1
488	US448	306	2.5	-1.16	0.49	0.39	-1.8	0.37	-1.8
491	US450	306	3.4	0.30	0.57	0.62	-0.7	0.61	-0.8
496	US454	306	4.4	0.34	0.55	1.57	1.3	1.56	1.3
502	US459	306	3.3	1.76	0.56	1.63	1.3	1.56	1.2
504	US460	306	3.0	0.24	0.55	1.27	0.6	1.35	0.8
507	US462	306	4.2	0.95	0.56	0.95	0	0.94	0
509	US463	306	4.3	2.71	0.56	0.56	-1.1	0.53	-1.2
513	US466	306	4.6	2.56	0.55	1.36	0.9	1.42	1.0
518	US470	306	4.2	1.89	0.55	0.59	-1.1	0.56	-1.2
626	US577	304	3.1	-0.48	0.55	1.33	0.8	1.40	0.9
636	US586	304	3.0	1.33	0.56	0.97	0	0.95	0

Separation: 2.97 Reliability: 0.90

Fixed (all same) chi-square: 475.5 significance *p*: .00

Note: 304 = Grade 4 US sub-sample student, 306 = Grade 6 US sub-sample student

Calibration of the students for a sub-sample of US student essays using US raters, Lexile Analyzer and IEA (10 misfitting students)

Student			Observed	Measure	S.E.	Infit		Outfit	
Number		Group	score	(Logit)		MnSq	ZStd	MnSq	ZStd
25	US021	304	3.1	-1.11	0.52	1.41	1.0	1.31	0.8
257	US236	306	2.7	-1.24	0.51	1.48	1.1	1.51	1.1
259	US237	306	3.0	-1.56	0.56	1.49	1.0	1.55	1.1
236	US218	304	4.3	2.67	0.56	1.52	1.1	1.48	1.1
496	US454	306	4.4	0.34	0.55	1.57	1.3	1.56	1.3
60	US049	304	2.6	-1.79	0.48	1.60	1.5	1.46	1.2
502	US459	306	3.3	1.76	0.56	1.63	1.3	1.56	1.2
274	US248	306	4.8	3.76	0.56	1.79	1.7	1.81	1.7
461	US427	304	3.3	0.89	0.55	1.86	1.7	1.66	1.4
20	US017	304	1.4	-2.62	0.54	2.15	2.2	2.05	1.7

Separation: 2.97 Reliability: 0.90

Fixed (all same) chi-square: 475.5 significance *p*: .00

Note: 304 = Grade 4 US sub-sample student, 306 = Grade 6 US sub-sample student

Calibration of the prompts for a sub-sample of US student essays using US raters, Lexile Analyzer and IEA

Prompt	Observed score	Measure (Logit)	S.E.	Infit		Outfit	
				MnSq	ZStd	MnSq	ZStd
4BI	3.5	-0.32	0.13	1.02	0.1	1.03	0.3
4BN	3.4	0.13	0.13	0.72	-3.0	0.72	-2.8
4BP	3.2	0.19	0.13	1.25	2.2	1.24	2.1
Separation: 1.39			Reliability: 0.66				
Fixed (all same) chi-square: 8.8			significance p: .00				

Calibration of the US raters for a sub-sample of US student essays using US raters, Lexile Analyzer and IEA

US rater	Observed score	Measure (Logit)	S.E.	Infit		Outfit	
				MnSq	ZStd	MnSq	ZStd
USR101	3.6	-0.55	0.34	1.01	0.1	1.07	0.3
USR102	3.5	-0.35	0.30	0.84	-0.6	0.82	-0.7
USR103	3.4	0.69	0.35	0.61	-1.4	0.59	-1.3
USR104	3.1	1.08	0.33	0.57	-1.8	0.65	-1.3
USR105	3.4	0.19	0.30	0.54	-2.3	0.49	-2.4
USR106	3.7	-0.38	0.34	0.76	-0.9	0.73	-1.0
USR200	3.1	1.70	0.32	1.23	0.9	1.20	0.8
USR201	3.5	-0.72	0.32	1.29	1.1	1.28	1.0
USR202	3.2	-1.06	0.45	0.99	0.1	0.99	0.1
USR203	3.3	0.15	0.22	1.11	0.6	1.18	1.0
USR204	3.5	0.15	0.28	1.31	1.3	1.34	1.4
USR205	3.8	-1.99	0.38	1.12	0.4	1.05	0.2
USR206	3.3	-0.24	0.27	1.09	0.4	1.09	0.4
USR207	3.1	-0.18	0.34	0.70	-1.1	0.71	-1.0
USR208	3.1	0.03	0.45	1.28	0.8	1.24	0.7
USR209	3.0	0.89	0.35	1.57	1.8	1.54	1.5
USR210	3.7	-0.75	0.27	0.93	-0.2	0.92	-0.3
USR211	2.5	1.35	0.94	0.61	-0.4	0.60	-0.2
Separation: 2.00				Reliability: 0.80			
Fixed (all same) chi-square: 104.1				significance p: .00			

Calibration of the Lexile Analyzer for a sub-sample of US student essays using US raters, Lexile Analyzer and IEA

Lexile Analyzer	Observed score	Measure (Logit)	S.E.	Infit MnSq	Infit ZStd	Outfit MnSq	Outfit ZStd
Category 1	2.1	2.15	0.16	0.85	-1.1	0.88	-0.7
Category 2	2.9	1.61	0.17	1.22	1.5	1.18	1.2
Category 3	3.5	0.42	0.17	0.96	-0.2	0.97	-0.1
Category 4	3.9	-1.00	0.18	0.90	-0.7	0.90	-0.6
Category 5	4.2	-0.25	0.23	1.14	0.8	1.14	0.7
Category 6	4.8	-1.91	0.35	0.90	-0.3	0.89	-0.3
Category 7	/	/	/	/	/	/	/
Category 8	5.1	-0.51	0.55	0.76	-0.6	0.75	-0.7
Category 9	5.5	-0.51	0.99	0.87	0	0.89	0
Separation: 2.65			Reliability: 0.88				
Fixed (all same) chi-square: 278.9			significance *p*: .00				

Calibration of IEA for a sub-sample of US student essays using US raters, Lexile Analyzer and IEA

IEA	Observed score	Measure (Logit)	S.E.	Infit MnSq	Infit ZStd	Outfit MnSq	Outfit ZStd
Category 1	2.2	1.85	0.16	0.93	-0.4	0.96	-0.2
Category 2	2.9	1.54	0.19	0.87	-0.8	0.87	-0.7
Category 3	3.3	0.85	0.18	1.12	0.9	1.10	0.7
Category 4	3.6	1.10	0.18	1.04	0.3	1.01	0.1
Category 5	4.3	-0.75	0.20	1.06	0.4	1.06	0.4
Category 6	4.8	-2.15	0.29	1.01	0	0.99	0
Category 7	4.8	-0.62	0.68	0.68	-0.6	0.67	-0.6
Category 8	4.9	-1.83	0.69	0.64	-0.8	0.63	-0.8
Category 9	/	/	/	/	/	/	/
Separation: 3.61			Reliability: 0.93				
Fixed (all same) chi-square: 242.6			significance *p*: .00				

Appendix 17

```
+-------------------------------------+
|Measr|-Prompt|-HK Rater|+Student|R6SCA|
|-----+-------+---------+--------+-----|
|   3 +       +         +        + (6) |
|     |       |         |        |     |
|     |       |         |        |     |
|     |       |         |        |     |
|     |       |         |        |     |
|     |       |         | **     | --- |
|     |       |         |        |     |
|     |       |         |        |     |
|   2 +       +         +        +     |
|     |       |         | *      |     |
|     |       |         |        |     |
|     |       |         | ***    |  5  |
|     |       |         |        |     |
|     |       |         |        |     |
|     |       |         | *      |     |
|     |       |         |        |     |
|   1 +       +         + *      + --- |
|     |       |         | *****  |     |
|     |       |         | **     |     |
|     |       |         | **     |     |
|     |       |         | ****   |  4  |
|     |       | HK001   |        |     |
|     |       | HK004   | **     |     |
|     | 4-4BI |         | **     |     |
*   0 * 6-4BP *         * *      * --- *
|     | 5-4BN |         | **     |     |
|     |       | HK002   |        |     |
|     |       | HK003   | **     |     |
|     |       |         | *****  |  3  |
|     |       |         |        |     |
|     |       |         | **     |     |
|     |       |         | **     |     |
|  -1 +       +         + *      + --- |
|     |       |         |        |     |
|     |       |         | *      |     |
|     |       |         | ***    |     |
|     |       |         | *      |     |
|     |       |         |        |  2  |
|     |       |         |        |     |
|     |       |         | *      |     |
|  -2 +       +         +        +     |
|     |       |         | **     |     |
|     |       |         | *      |     |
|     |       |         |        | --- |
|     |       |         |        |     |
|     |       |         |        |     |
|     |       |         |        |     |
|     |       |         |        |     |
|  -3 +       +         + *      + (1) |
|-----+-------+---------+--------+-----|
|Measr|-Prompt|-HK Rater| * = 1  |R6SCA|
+-------------------------------------+
```

Vertical map for a sub-sample of US student essays with scoring of HK raters

Calibration of the students for sub-sample of US student essays using HK raters

Student			Observed	Measure	S.E.	Infit		Outfit	
	Number	Group	score	(Logit)		MnSq	ZStd	MnSq	ZStd
7	US006	304	4.0	0.52	0.41	0.57	-0.7	0.58	-0.7
15	US013	304	4.2	0.70	0.41	0.35	-1.5	0.35	-1.5
20	US017	304	1.7	-2.15	0.53	0.80	-0.1	0.76	-0.2

25	US021	304	3.2	-0.33	0.42	1.09	0.3	1.08	0.3
27	US022	306	3.3	-0.17	0.42	0.34	-1.3	0.34	-1.3
32	US026	306	3.7	0.19	0.41	0.36	-1.3	0.36	-1.3
34	US027	306	5.5	2.43	0.60	2.46	1.7	2.02	1.3
38	US030	306	3.5	0.02	0.42	1.77	1.2	1.77	1.2
43	US034	306	3.7	0.18	0.41	1.54	1.0	1.58	1.0
56	US046	304	3.0	-0.50	0.42	1.56	1.0	1.56	1.0
60	US049	304	3.0	-0.52	0.42	1.37	0.7	1.37	0.7
69	US057	304	3.5	0.19	0.51	2.65	1.8	2.67	1.8
73	US060	304	1.7	-2.15	0.53	0.94	0	0.89	0
78	US064	306	4.3	0.87	0.42	1.21	0.5	1.24	0.6
85	US070	306	1.8	-1.89	0.49	1.44	0.9	1.49	0.9
228	US212	304	2.2	-1.43	0.45	0.55	-0.8	0.58	-0.8
232	US215	304	3.2	-0.34	0.42	0.52	-0.8	0.53	-0.8
236	US218	304	4.0	0.53	0.41	1.16	0.4	1.17	0.4
246	US227	304	2.4	-1.24	0.48	0.59	-0.6	0.62	-0.5
255	US235	306	3.0	-0.50	0.42	1.65	1.1	1.65	1.1
257	US236	306	2.5	-1.06	0.43	0.91	0	0.93	0
259	US237	306	4.3	0.88	0.42	0.51	-0.9	0.52	-0.9
264	US241	306	4.3	0.86	0.42	0.79	-0.2	0.80	-0.2
267	US243	306	5.0	1.62	0.46	0.54	-0.8	0.55	-0.8
271	US246	306	3.7	0.18	0.41	0.69	-0.4	0.69	-0.4
274	US248	306	5.0	1.62	0.46	1.24	0.6	1.14	0.4
283	US256	304	1.7	-2.28	0.75	1.32	0.6	1.21	0.5
294	US266	304	2.2	-1.43	0.45	1.54	1.0	1.45	0.9
302	US273	304	4.0	0.52	0.41	1.04	0.2	1.04	0.2
315	US285	304	1.0	-3.90	1.80	Min			
435	US404	306	2.8	-0.70	0.42	0.33	-1.4	0.33	-1.4
451	US419	304	3.0	-0.50	0.42	0.43	-1.0	0.43	-1.0
453	US420	304	2.2	-1.45	0.45	0.72	-0.4	0.78	-0.3
461	US427	304	3.4	-0.15	0.46	1.52	0.9	1.54	0.9
464	US429	304	2.7	-0.88	0.43	1.06	0.2	1.05	0.2
472	US436	304	4.5	1.05	0.42	2.49	2.2	2.53	2.3
476	US439	306	3.0	-0.52	0.42	0.21	-1.9	0.21	-1.9
478	US440	306	5.2	1.85	0.49	0.38	-1.2	0.40	-1.1
482	US443	306	5.5	2.43	0.60	0.75	-0.1	0.82	0
488	US448	306	2.7	-0.86	0.43	1.29	0.6	1.28	0.6
491	US450	306	4.2	0.69	0.41	0.42	-1.2	0.42	-1.2

496	US454	306	4.2	0.70	0.41	0.86	-0.1	0.85	-0.1
502	US459	306	2.8	-0.70	0.42	1.57	1.0	1.57	1.0
504	US460	306	4.3	0.87	0.42	1.04	0.2	1.05	0.2
507	US462	306	4.2	0.69	0.41	0.51	-0.9	0.51	-0.9
509	US463	306	4.3	0.88	0.42	0.61	-0.7	0.61	-0.7
513	US466	306	4.7	1.22	0.43	1.66	1.2	1.77	1.3
518	US470	306	5.0	1.62	0.46	1.73	1.3	1.81	1.4
626	US577	304	4.0	0.54	0.51	0.44	-0.8	0.43	-0.8
636	US586	304	2.2	-1.42	0.45	0.38	-1.4	0.41	-1.3

Separation: 2.24 Reliability: 0.83

Fixed (all same) chi-square: 274.9 significance *p*: .00

Note: 304 = Grade 4 US sub-sample student, 306 = Grade 6 US sub-sample student

Calibration of the students for a sub-sample of US student essays using HK raters (13 misfitting students)

Student			Observed	Measure	S.E.	Infit		Outfit	
Number		Group	score	(Logit)		MnSq	ZStd	MnSq	ZStd
85	US070	306	1.8	-1.89	0.49	1.44	0.9	1.49	0.9
461	US427	304	3.4	-0.15	0.46	1.52	0.9	1.54	0.9
43	US034	306	3.7	0.18	0.41	1.54	1.0	1.58	1.0
294	US266	304	2.2	-1.43	0.45	1.54	1.0	1.45	0.9
56	US046	304	3.0	-0.50	0.42	1.56	1.0	1.56	1.0
502	US459	306	2.8	-0.70	0.42	1.57	1.0	1.57	1.0
255	US235	306	3.0	-0.50	0.42	1.65	1.1	1.65	1.1
513	US466	306	4.7	1.22	0.43	1.66	1.2	1.77	1.3
518	US470	306	5.0	1.62	0.46	1.73	1.3	1.81	1.4
38	US030	306	3.5	0.02	0.42	1.77	1.2	1.77	1.2
34	US027	306	5.5	2.43	0.60	2.46	1.7	2.02	1.3
472	US436	304	4.5	1.05	0.42	2.49	2.2	2.53	2.3
69	US057	304	3.5	0.19	0.51	2.65	1.8	2.67	1.8

Separation: 2.24 Reliability: 0.83

Fixed (all same) chi-square: 274.9 significance *p*: .00

Note: 304 = Grade 4 US sub-sample student, 306 = Grade 6 US sub-sample student

Calibration of the prompts for a sub-sample of US student essays using HK raters

Prompt	Observed score	Measure (Logit)	S.E.	Infit		Outfit	
				MnSq	ZStd	MnSq	ZStd
4BI	3.5	0.06	0.11	0.93	-0.4	0.93	-0.4
4BN	3.6	-0.08	0.11	0.89	-0.7	0.89	-0.7
4BP	3.5	0.02	0.11	1.18	1.2	1.20	1.3
Separation: 1.76			Reliability: 0.76				
Fixed (all same) chi-square: 12.1			significance *p*: .00				

Calibration of the HK raters for a sub-sample of US student essays using HK raters

HK rater	Observed score	Measure (Logit)	S.E.	Infit		Outfit	
				MnSq	ZStd	MnSq	ZStd
HKR001	3.1	0.38	0.13	0.84	-0.9	0.86	-0.8
HKR002	3.8	-0.26	0.12	1.16	0.9	1.17	1.0
HKR003	3.8	-0.43	0.13	0.80	-1.3	0.81	-1.2
HKR004	3.4	0.31	0.13	1.18	1.1	1.20	1.1
Separation: 2.59			Reliability: 0.87				
Fixed (all same) chi-square: 30.7			significance *p*: .00				

```
+------------------------------------------------------------------+
|Measr|-Prompt|-HK Rater|-Lexile Analyzer|-IEA |+Student  |R6SCA|
|-----+-------+---------+----------------+-----+----------+-----|
|   3 +       +         +                +     +          + (6) |
|     |       |         |                |     | *        |     |
|     |       |         |                |     |          | --- |
|     |       |         |                |     | *        |     |
|     |       |         |                |     |          |     |
|     |       |         |                |     |          |     |
|     |       |         |                |     | **       |     |
|     |       |         |                |     | *        |     |
|   2 +       +         +                +     +          +     |
|     |       |         |                |     |          |  5  |
|     |       |         |                |     |          |     |
|     |       |         |                |     | ***      |     |
|     |       |         | 1              |     | *        |     |
|     |       |         |                |     | ******   |     |
|     |       |         |                |     | *        | --- |
|     |       |         |                |     | **       |     |
|   1 +       +         + 2              +     +          +     |
|     |       |         |                | 2   | **       |     |
|     |       |         | 8              | 1 4 | **       |     |
|     |       |         |                |     | *******  |  4  |
|     |       | HK001   |                |     | *        |     |
|     |       | HK004   |                | 5   | ****     |     |
|     | 5-4BN |         |                | 3   | ***      |     |
|     |       |         |                |     | *        |     |
*   0 * 4-4BI *         * 3              *     * ***      * --- *
|     |       |         |                |     |          |     |
|     | 6-4BP | HK002   |                | 7   | **       |     |
|     |       |         | 4              | 6   | *        |     |
|     |       | HK003   |                |     |          |     |
|     |       |         |                |     | ***      |  3  |
|     |       |         |                |     |          |     |
|     |       |         | 9              |     |          |     |
|  -1 +       +         + 5              +     + **       +     |
|     |       |         | 6              |     |          |     |
|     |       |         |                |     |          | --- |
|     |       |         |                |     |          |     |
|     |       |         |                |     |          |     |
|     |       |         |                |     |          |     |
|     |       |         |                |     |          |     |
|     |       |         |                |     |          |     |
|  -2 +       +         +                +     +          +  2  |
|     |       |         |                |     |          |     |
|     |       |         |                | 8   |          |     |
|     |       |         |                |     |          |     |
|     |       |         |                |     |          |     |
|     |       |         |                |     |          |     |
|     |       |         |                |     |          | --- |
|     |       |         |                |     |          |     |
|  -3 +       +         +                +     + *        + (1) |
|-----+-------+---------+----------------+-----+----------+-----|
|Measr|-Prompt|-HK Rater|-Lexile Analyzer|-IEA | * = 1    |R6SCA|
+------------------------------------------------------------------+
```

Vertical map for a sub-sample of US student essays with scoring of HK raters, Lexile Analyzer and IEA

Calibration of the students for a sub-sample of US student essays using HK raters, Lexile Analyzer and IEA

Student			Observed	Measure	S.E.	Infit		Outfit	
Number		Group	score	(Logit)		MnSq	ZStd	MnSq	ZStd
7	US006	304	4.0	0.67	0.46	0.60	-0.6	0.6	-0.6
15	US013	304	4.2	0.61	0.46	0.56	-0.7	0.57	-0.7
20	US017	304	1.7	-0.29	0.56	1.00	0.1	0.96	0.1
25	US021	304	3.2	0.21	0.47	0.67	-0.4	0.66	-0.4
27	US022	306	3.3	-0.04	0.46	0.37	-1.2	0.37	-1.2
32	US026	306	3.7	0.04	0.46	0.31	-1.4	0.31	-1.4
34	US027	306	5.5	2.87	0.66	1.67	1.0	0.97	0.2
38	US030	306	3.5	0.60	0.46	2.40	1.9	2.43	1.9
43	US034	306	3.7	1.32	0.46	1.92	1.4	1.95	1.5
56	US046	304	3.0	0.72	0.47	0.72	-0.4	0.69	-0.4
60	US049	304	3.0	0.01	0.46	1.51	0.9	1.52	0.9
69	US057	304	3.5	1.33	0.57	2.55	1.7	2.53	1.7
73	US060	304	1.7	-1.00	0.57	0.34	-1.3	0.32	-1.1
78	US064	306	4.3	0.76	0.47	1.39	0.8	1.38	0.8
85	US070	306	1.8	-0.62	0.53	1.61	1.1	2.29	1.8
228	US212	304	2.2	-0.60	0.50	0.87	0	0.98	0.1
232	US215	304	3.2	0.47	0.47	0.71	-0.3	0.71	-0.3
236	US218	304	4.0	1.37	0.47	1.74	1.2	1.71	1.2
246	US227	304	2.4	0.33	0.53	0.48	-0.9	0.49	-0.8
255	US235	306	3.0	1.66	0.47	1.52	0.9	1.51	0.9
257	US236	306	2.5	-0.68	0.48	0.67	-0.4	0.67	-0.4
259	US237	306	4.3	1.61	0.48	0.56	-0.7	0.57	-0.7
264	US241	306	4.3	0.60	0.47	0.39	-1.2	0.42	-1.2
267	US243	306	5.0	1.61	0.50	0.40	-1.2	0.43	-1.1
271	US246	306	3.7	0.37	0.46	0.97	0.1	0.97	0.1
274	US248	306	5.0	2.19	0.50	0.98	0.1	0.92	0
283	US256	304	1.7	-0.36	0.80	2.01	1.2	1.80	1.0
294	US266	304	2.2	-0.21	0.50	1.12	0.3	1.06	0.2
302	US273	304	4.0	0.67	0.46	0.89	0	0.89	0
315	US285	304	1.0	-2.34	1.83	Minim	um		
435	US404	306	2.8	1.27	0.47	0.39	-1.2	0.38	-1.2
451	US419	304	3.0	1.49	0.47	0.36	-1.2	0.36	-1.2
453	US420	304	2.2	0.38	0.49	0.45	-1.1	0.48	-1.0
461	US427	304	3.4	0.86	0.51	1.04	0.2	1.04	0.2

464	US429	304	2.7	0.85	0.47	1.57	1.0	1.57	1.0
472	US436	304	4.5	0.33	0.46	2.06	1.7	2.04	1.7
476	US439	306	3.0	0.23	0.47	0.06	-3.0	0.06	-3.0
478	US440	306	5.2	1.08	0.53	0.35	-1.3	0.38	-1.2
482	US443	306	5.5	2.67	0.64	1.11	0.3	1.03	0.2
488	US448	306	2.7	0.19	0.47	1.30	0.6	1.31	0.6
491	US450	306	4.2	1.35	0.46	1.24	0.5	1.26	0.6
496	US454	306	4.2	-0.96	0.46	1.17	0.4	1.17	0.4
502	US459	306	2.8	0.63	0.47	1.57	1.0	1.54	1.0
504	US460	306	4.3	2.16	0.46	1.30	0.6	1.30	0.6
507	US462	306	4.2	0.68	0.46	0.75	-0.3	0.76	-0.2
509	US463	306	4.3	1.43	0.45	0.49	-1.0	0.49	-1.0
513	US466	306	4.7	1.36	0.51	1.32	0.6	1.32	0.6
518	US470	306	5.0	2.24	0.51	1.82	1.3	1.63	1.1
626	US577	304	4.0	1.15	0.56	0.69	-0.2	0.68	-0.3
636	US586	304	2.2	0.18	0.47	0.24	-2.2	0.24	-2.2
Separation: 1.44				Reliability: 0.67					
Fixed (all same) chi-square: 144.4				significance *p*: .00					

Note: 304 = Grade 4 US sub-sample student, 306 = Grade 6 US sub-sample student

Calibration of the students for a sub-sample of US student essays using HK raters, Lexile Analyzer and IEA (13 misfitting students)

Student			Observed	Measure	S.E.	Infit		Outfit	
Number		Group	score	(Logit)		MnSq	ZStd	MnSq	ZStd
60	US049	304	3.0	0.01	0.46	1.51	0.9	1.52	0.9
255	US235	306	3.0	1.66	0.47	1.52	0.9	1.51	0.9
464	US429	304	2.7	0.85	0.47	1.57	1.0	1.57	1.0
502	US459	306	2.8	0.63	0.47	1.57	1.0	1.54	1.0
85	US070	306	1.8	-0.62	0.53	1.61	1.1	2.29	1.8
34	US027	306	5.5	2.87	0.66	1.67	1.0	0.97	0.2
236	US218	304	4.0	1.37	0.47	1.74	1.2	1.71	1.2
518	US470	306	5.0	2.24	0.51	1.82	1.3	1.63	1.1
43	US034	306	3.7	1.32	0.46	1.92	1.4	1.95	1.5
283	US256	304	1.7	-0.36	0.80	2.01	1.2	1.80	1.0
472	US436	304	4.5	0.33	0.46	2.06	1.7	2.04	1.7
38	US030	306	3.5	0.60	0.46	2.40	1.9	2.43	1.9
69	US057	304	3.5	1.33	0.57	2.55	1.7	2.53	1.7

Separation: 1.44	Reliability: 0.67
Fixed (all same) chi-square: 144.4	significance *p*: .00

Note: 304 = Grade 4 US sub-sample student, 306 = Grade 6 US sub-sample student

Calibration of the prompts for a sub-sample of US student essays using HK raters, Lexile Analyzer and IEA

Prompt	Observed score	Measure (Logit)	S.E.	Infit		Outfit	
				MnSq	ZStd	MnSq	ZStd
4BI	3.5	0.03	0.12	0.97	-0.1	0.94	-0.3
4BN	3.6	0.23	0.12	0.82	-1.3	0.81	-1.3
4BP	3.5	-0.26	0.12	1.23	1.5	1.26	1.7
Separation: 1.35				Reliability: 0.65			
Fixed (all same) chi-square: 8.4				significance *p*: .00			

Calibration of the HK raters for a sub-sample of US student essays using HK raters, Lexile Analyzer and IEA

HK rater	Observed score	Measure (Logit)	S.E.	Infit		Outfit	
				MnSq	ZStd	MnSq	ZStd
HKR001	3.1	0.50	0.14	0.98	0	1.02	0.1
HKR002	3.8	-0.31	0.14	1.30	1.7	1.28	1.6
HKR003	3.8	-0.51	0.14	0.68	-2.1	0.68	-2.0
HKR004	3.4	0.32	0.14	1.04	0.2	1.02	0.1
Separation: 2.83				Reliability: 0.89			
Fixed (all same) chi-square: 36.0				significance *p*: .00			

Calibration of the Lexile Analyzer for a sub-sample of US student essays using HK raters, Lexile Analyzer and IEA

Lexile Analyzer	Observed score	Measure (Logit)	S.E.	Infit		Outfit	
				MnSq	ZStd	MnSq	ZStd
Category 1	2.2	1.56	0.17	0.99	0	1.06	0.3
Category 2	2.8	1.04	0.15	1.16	0.9	1.15	0.8
Category 3	3.8	-0.02	0.15	1.33	1.7	1.33	1.7
Category 4	4.0	-0.42	0.15	0.59	-2.6	0.59	-2.5
Category 5	4.6	-0.94	0.20	0.93	-0.2	0.90	-0.3
Category 6	4.7	-1.07	0.32	0.76	-0.5	0.71	-0.5
Category 7	/	/	/	/	/	/	/
Category 8	4.5	0.78	0.49	1.17	0.4	1.11	0.3
Category 9	5.5	-0.93	1.08	0.48	-0.2	0.48	-0.2

Separation: 1.84	Reliability: 0.77
Fixed (all same) chi-square: 158.8	significance *p*: .00

Calibration of IEA for a sub-sample of US student essays using HK raters, Lexile Analyzer and IEA

IEA	Observed score	Measure (Logit)	S.E.	Infit		Outfit	
				MnSq	ZStd	MnSq	ZStd
Category 1	2.3	0.69	0.17	1.00	0	0.98	0
Category 2	2.8	0.86	0.17	0.77	-1.1	0.89	-0.5
Category 3	3.5	0.24	0.16	1.16	0.8	1.15	0.8
Category 4	3.8	0.72	0.15	1.23	1.2	1.23	1.1
Category 5	4.5	0.42	0.17	0.76	-1.2	0.74	-1.3
Category 6	4.8	-0.40	0.26	1.17	0.6	1.08	0.3
Category 7	5.0	-0.29	0.69	0.62	-0.2	0.69	0
Category 8	5.0	-2.24	0.61	0.31	-1.2	0.33	-1.1
Category 9	/	/	/	/	/	/	/
Separation: 2.42			Reliability: 0.85				
Fixed (all same) chi-square: 43.6			significance *p*: .00				

Appendix 18

```
+----------------------------------------------------------------------------+
|Measr|-Prompts|-US Rater                           |-HK Rater|+Student|R6SCA|
|-----+--------+------------------------------------+---------+--------+-----|
|   4 +        +                                    +         +        + (6) |
|     |        |                                    |         |        |     |
|     |        |                                    |         |        |     |
|     |        |                                    |         |        | --- |
|     |        |                                    |         | *      |     |
|   3 +        +                                    +         +        +     |
|     |        |                                    |         |        |     |
|     |        |                                    |         |        |     |
|     |        |                                    |         |        |  5  |
|     |        |                                    |         | *      |     |
|   2 +        +                                    +         + *      +     |
|     |        |                                    |         | **     |     |
|     |        |                                    |         | *      | --- |
|     |        |                                    |         | *****  |     |
|     |        | US211                              |         | **     |     |
|   1 +        + US200                              +         + **     +     |
|     |        |                                    |         | **     |  4  |
|     |        | US209                              |         | *      |     |
|     |        | US104                              | HK001   | *      |     |
|     | 6-4BP  | US103   US203   US204              | HK004   |        |     |
*   0 * 4-4BI  * US105   US206   US207   US208      *         * ***    * --- *
|     | 5-4BN  | US102   US210                      |         | ****   |     |
|     |        | US101   US106   US201   US202      |         | **     |     |
|     |        |                                    | HK002   | **     |     |
|     |        | US205                              | HK003   | **     |  3  |
|  -1 +        +                                    +         + *      +     |
|     |        |                                    |         | *      |     |
|     |        |                                    |         | *****  |     |
|     |        |                                    |         | ****   | --- |
|     |        |                                    |         |        |     |
|  -2 +        +                                    +         + *      +     |
|     |        |                                    |         | *      |     |
|     |        |                                    |         | *      |  2  |
|     |        |                                    |         | *      |     |
|     |        |                                    |         |        |     |
|  -3 +        +                                    +         + *      +     |
|     |        |                                    |         |        |     |
|     |        |                                    |         | *      | --- |
|     |        |                                    |         |        |     |
|     |        |                                    |         |        |     |
|  -4 +        +                                    +         +        +     |
|     |        |                                    |         |        |     |
|     |        |                                    |         |        |     |
|     |        |                                    |         |        |     |
|     |        |                                    |         |        |     |
|  -5 +        +                                    +         +        +     |
|     |        |                                    |         | *      |     |
|     |        |                                    |         |        |     |
|     |        |                                    |         |        |     |
|     |        |                                    |         |        |     |
|  -6 +        +                                    +         +        + (1) |
|-----+--------+------------------------------------+---------+--------+-----|
|Measr|-Prompts|-US Rater                           |-HK Rater| * = 1  |R6SCA|
+----------------------------------------------------------------------------+
```

Vertical map for a sub-sample of US student essays with scoring of US and HK raters

Calibration of the students for a sub-sample of US student essays using US and HK raters

Student			Observed	Measure	S.E.	Infit		Outfit	
Number		Group	score	(Logit)		MnSq	ZStd	MnSq	ZStd
7	US006	304	3.9	0.74	0.21	0.45	-2.9	0.44	-2.9
15	US013	304	4.3	1.37	0.21	0.70	-1.4	0.71	-1.3
20	US017	304	1.5	-3.41	0.26	1.24	0.9	1.10	0.4

25	US021	304	3.1	-0.76	0.22	1.42	1.5	1.44	1.6
27	US022	306	3.2	-0.75	0.22	0.53	-2.1	0.52	-2.2
32	US026	306	3.9	0.59	0.21	0.42	-3.0	0.42	-3.1
34	US027	306	4.8	1.89	0.21	1.34	1.5	1.31	1.4
38	US030	306	3.5	-0.16	0.22	1.26	1.0	1.25	1.0
43	US034	306	3.3	-0.55	0.22	1.21	0.8	1.21	0.8
56	US046	304	2.7	-1.41	0.21	1.19	0.8	1.25	1.0
60	US049	304	2.7	-1.33	0.21	1.64	2.3	1.62	2.3
69	US057	304	3.5	-0.14	0.27	1.82	2.2	1.84	2.2
73	US060	304	2.0	-2.55	0.21	1.42	2.0	1.41	1.9
78	US064	306	4.2	1.09	0.21	0.93	-0.2	0.93	-0.2
85	US070	306	2.4	-1.95	0.21	1.03	0.2	1.07	0.3
228	US212	304	3.0	-1.04	0.22	1.11	0.5	1.11	0.5
232	US215	304	3.5	-0.18	0.22	0.66	-1.4	0.66	-1.4
236	US218	304	4.2	1.02	0.21	1.14	0.6	1.15	0.6
246	US227	304	2.6	-1.62	0.25	0.54	-2.0	0.52	-2.1
255	US235	306	2.6	-1.48	0.21	1.14	0.6	1.12	0.5
257	US236	306	2.6	-1.49	0.21	0.80	-0.8	0.79	-0.9
259	US237	306	3.4	-0.09	0.22	1.27	1.0	1.25	1.0
264	US241	306	4.3	1.11	0.21	0.92	-0.3	0.92	-0.3
267	US243	306	4.4	1.37	0.21	0.76	-1.1	0.76	-1.1
271	US246	306	3.6	-0.08	0.22	1.00	0	1.01	0.1
274	US248	306	4.9	2.07	0.21	1.67	2.7	1.62	2.6
283	US256	304	1.8	-2.91	0.28	0.86	-0.4	0.82	-0.6
294	US266	304	2.1	-2.50	0.21	1.18	0.9	1.15	0.7
302	US273	304	4.1	0.86	0.21	0.87	-0.5	0.86	-0.5
315	US285	304	1.1	-5.13	0.71	0.95	0.1	1.20	0.4
435	US404	306	2.7	-1.60	0.21	0.42	-3.1	0.43	-3.1
451	US419	304	2.8	-1.27	0.21	0.35	-3.6	0.35	-3.6
453	US420	304	2.2	-2.28	0.21	0.50	-3.0	0.52	-2.8
461	US427	304	3.3	-0.48	0.23	1.17	0.6	1.15	0.6
464	US429	304	2.6	-1.59	0.21	0.63	-1.8	0.64	-1.7
472	US436	304	4.4	1.36	0.21	1.75	2.8	1.76	2.9
476	US439	306	3.4	-0.14	0.22	0.89	-0.3	0.89	-0.3
478	US440	306	5.4	3.25	0.25	0.62	-1.8	0.63	-1.7
482	US443	306	4.8	2.20	0.21	0.88	-0.5	0.86	-0.6
488	US448	306	2.6	-1.52	0.21	0.80	-0.8	0.83	-0.7
491	US450	306	3.7	0.33	0.22	0.96	0	0.95	-0.1

496	US454	306	4.3	1.42	0.21	1.12	0.6	1.13	0.6
502	US459	306	3.1	-0.60	0.22	1.71	2.4	1.73	2.4
504	US460	306	3.4	-0.02	0.22	1.46	1.6	1.44	1.6
507	US462	306	4.2	1.17	0.21	0.71	-1.3	0.72	-1.3
509	US463	306	4.3	1.36	0.21	0.52	-2.5	0.51	-2.5
513	US466	306	4.6	1.87	0.21	1.27	1.2	1.34	1.5
518	US470	306	4.4	1.62	0.21	1.05	0.3	1.06	0.3
626	US577	304	3.3	-0.41	0.23	1.05	0.2	1.05	0.2
636	US586	304	2.7	-1.39	0.21	0.87	-0.4	0.89	-0.3

Separation: 6.81 Reliability: 0.98

Fixed (all same) chi-square: 2313.6 significance *p*: .00

Note: 304 = Grade 4 US sub-sample student, 306 = Grade 6 US sub-sample student

Calibration of the students for a sub-sample of US student essays using US and HK raters (8 misfitting students)

Student			Observed	Measure	S.E.	Infit		Outfit	
Number		Group	score	(Logit)		MnSq	ZStd	MnSq	ZStd
25	US021	304	3.1	-0.76	0.22	1.42	1.5	1.44	1.6
73	US060	304	2.0	-2.55	0.21	1.42	2.0	1.41	1.9
504	US460	306	3.4	-0.02	0.22	1.46	1.6	1.44	1.6
60	US049	304	2.7	-1.33	0.21	1.64	2.3	1.62	2.3
274	US248	306	4.9	2.07	0.21	1.67	2.7	1.62	2.6
502	US459	306	3.1	-0.60	0.22	1.71	2.4	1.73	2.4
472	US436	304	4.4	1.36	0.21	1.75	2.8	1.76	2.9
69	US057	304	3.5	-0.14	0.27	1.82	2.2	1.84	2.2

Separation: 6.81 Reliability: 0.98

Fixed (all same) chi-square: 2313.6 significance *p*: .00

Note: 304 = Grade 4 US sub-sample student, 306 = Grade 6 US sub-sample student

Calibration of the prompts for a sub-sample of US student essays using US and HK raters

Prompt	Observed	Measure	S.E.	Infit		Outfit	
	score	(Logit)		MnSq	ZStd	MnSq	ZStd
4BI	3.5	-0.08	0.05	1.01	0.2	1.00	0
4BN	3.5	-0.13	0.05	0.84	-3.0	0.86	-2.4
4BP	3.3	0.22	0.05	1.16	2.6	1.15	2.5

Separation: 2.70 Reliability: 0.88

Fixed (all same) chi-square: 24.6 significance *p*: .00

Calibration of the US raters for a sub-sample of US student essays using US and HK raters

US rater	Observed score	Measure (Logit)	S.E.	Infit MnSq	Infit ZStd	Outfit MnSq	Outfit ZStd
USR101	3.6	-0.34	0.24	0.63	-1.5	0.75	-0.9
USR102	3.5	-0.17	0.21	0.54	-2.4	0.55	-2.2
USR103	3.4	0.27	0.24	0.35	-3.3	0.34	-3.1
USR104	3.1	0.46	0.23	0.32	-3.8	0.33	-3.6
USR105	3.4	0.01	0.20	0.35	-3.9	0.34	-3.7
USR106	3.7	-0.31	0.24	0.62	-1.6	0.62	-1.6
USR200	3.1	0.92	0.22	0.94	-0.1	0.93	-0.2
USR201	3.5	-0.40	0.22	0.65	-1.5	0.65	-1.5
USR202	3.2	-0.45	0.32	1.03	0.1	1.00	0.1
USR203	3.3	0.12	0.16	0.84	-0.8	0.85	-0.8
USR204	3.5	0.17	0.20	0.69	-1.6	0.68	-1.6
USR205	3.8	-0.82	0.26	0.61	-1.5	0.61	-1.5
USR206	3.3	-0.05	0.19	0.92	-0.3	0.92	-0.3
USR207	3.1	0.03	0.24	0.73	-1.1	0.72	-1.1
USR208	3.1	-0.08	0.32	0.77	-0.5	0.74	-0.6
USR209	3.0	0.52	0.25	0.77	-0.8	0.76	-0.7
USR210	3.7	-0.29	0.19	0.66	-1.8	0.66	-1.8
USR211	2.5	1.14	0.72	0.53	-0.4	0.48	-0.3
Separation: 1.55				Reliability: 0.71			
Fixed (all same) chi-square: 100.1				significance p: .00			

Calibration of the HK raters for a sub-sample of US student essays using US and HK raters

HK rater	Observed score	Measure (Logit)	S.E.	Infit MnSq	Infit ZStd	Outfit MnSq	Outfit ZStd
HKR001	3.1	0.35	0.15	1.53	2.8	1.64	3.2
HKR002	3.7	-0.56	0.15	1.84	4.2	1.80	3.9
HKR003	3.8	-0.78	0.15	1.57	3.0	1.54	2.8
HKR004	3.4	0.25	0.15	1.75	3.7	1.76	3.7
Separation: 1.54				Reliability: 0.70			
Fixed (all same) chi-square: 98.4				significance p: .00			

Appendix 19

```
+-----------------------------------------------------------------------------------------------------------------------------------------+
|Measr|-Prompt                  |-US Rater                                                    |-HK Rater|-Lexile Analyzer|+Student   |R6SCA|
|-----+------------------------+-------------------------------------------------------------+---------+----------------+-----------+-----|
|   5 +                        +                                                             +         +                +           + (6) |
|     |                        |                                                             |         |                |           |     |
|     |                        |                                                             |         |                | .         | --- |
|     |                        |                                                             |         |                |           |     |
|   4 +                        +                                                             +         +                +           +     |
|     |                        |                                                             |         |                | .         |     |
|     |                        |                                                             |         |                | .         |  5  |
|     |                        |                                                             |         |                | .         |     |
|   3 +                        +                                                             +         +                + .         +     |
|     |                        |                                                             |         |                | *.        |     |
|     |                        |                                                             |         | 1              | *.        |     |
|     |                        |                                                             |         |                | *.        | --- |
|   2 + 15-12AP                +                                                             +         +                + **.       +     |
|     | 16-12BI                |                                                             |         | 2              | ***.      |     |
|     | 13-12AI   9-8AP        |                                                             |         |                | ******.   |     |
|     | 17-12BN                |                                                             |         | 3              | *****.    |  4  |
|   1 + 11-8BN    18-12BP      + US200    US211                                              +         +                + *******.  +     |
|     | 14-12AN   12-8BP       |                                                             |         | 4              | ****.     |     |
|     | 10-8BI                 | US103    US104    US209                                      | HK004   |                | ******.   |     |
|     |                        | US206                                                       | HK001   |                | *******.  |     |
*   0 * 7-8AI     8-8AN        * US102    US105    US201    US203    US204    US205    US210 *         * 5              * ********. * --- *
|     |                        | US101    US202    US207    US208                             |         |                | *******.  |     |
|     |                        | US106                                                       | HK002   | 6              | *******.  |     |
|     |                        |                                                             |         |                | ******    |     |
|  -1 +                        +                                                             +         +                + *****.    +     |
|     |                        |                                                             |         | 7              | ****.     |  3  |
|     |                        | US108                                                       | HK003   |                | ****.     |     |
|     | 4-4BI     5-4BN        |                                                             |         | 8              | ****.     |     |
|  -2 + 6-4BP                  +                                                             +         +                + **.       +     |
|     | 1-4AI     2-4AN   3-4AP|                                                             |         |                | *.        | --- |
|     |                        |                                                             |         | 9              | **.       |     |
|     |                        |                                                             |         |                | *.        |     |
|  -3 +                        +                                                             +         +                + .         +     |
|     |                        |                                                             |         |                | *         |  2  |
|     |                        |                                                             |         |                | .         |     |
|     |                        |                                                             |         |                | .         |     |
|  -4 +                        +                                                             +         +                + *.        +     |
|     |                        |                                                             |         |                | .         |     |
|     |                        |                                                             |         |                | .         | --- |
|     |                        |                                                             |         |                | .         |     |
|  -5 +                        +                                                             +         +                +           +     |
|     |                        |                                                             |         |                | .         |     |
|     |                        |                                                             |         |                |           |     |
|     |                        |                                                             |         |                | .         |     |
|  -6 +                        +                                                             +         +                + .         +     |
|     |                        |                                                             |         |                | .         |     |
|     |                        |                                                             |         |                |           |     |
|     |                        |                                                             |         |                |           |     |
|  -7 +                        +                                                             +         +                + .         + (1) |
|-----+------------------------+-------------------------------------------------------------+---------+----------------+-----------+-----|
|Measr|-Prompt                  |-US Rater                                                    |-HK Rater|-Lexile Analyzer| * = 7     |R6SCA|
+-----------------------------------------------------------------------------------------------------------------------------------------+
```

Vertical map for all student essays with scoring of US and HK raters and Lexile Analyzer

Calibration of the students for all student essays using US and HK raters and Lexile Analyzer

Student			Observed	Measure	S.E.	Infit		Outfit	
Number		Group	score	(Logit)		MnSq	ZStd	MnSq	ZStd
1	US001	104	3.7	-0.44	0.31	0.59	-1.5	0.58	-1.5
2	US002	104	3.8	-0.19	0.34	0.77	-0.6	0.77	-0.6
3	US003	104	4.1	-0.04	0.31	0.82	-0.5	0.82	-0.5
4	US004	104	4.2	-0.06	0.31	0.61	-1.4	0.61	-1.4
5	US005	104	3.1	-1.29	0.31	0.87	-0.3	0.87	-0.3
6	US006	104	4.1	-0.18	0.28	0.57	-1.8	0.56	-1.8
7	US006	304	3.9	-0.29	0.26	0.57	-2.0	0.55	-2.1
8	US007	104	3.3	-1.20	0.31	0.78	-0.7	0.77	-0.7
9	US008	104	3.2	-0.91	0.31	0.42	-2.4	0.42	-2.4
10	US009	104	4.1	-0.06	0.31	1.22	0.8	1.23	0.8
11	US010	104	3.5	-0.83	0.31	0.46	-2.2	0.46	-2.2

12	US011	104	5.3	2.19	0.33	1.12	0.5	1.09	0.3
13	US012	104	1.7	-4.06	0.41	0.57	-1.4	0.66	-1.0
14	US013	104	4.3	0.55	0.27	1.14	0.6	1.16	0.6
15	US013	304	4.3	0.98	0.25	1.10	0.5	1.13	0.6
16	US014	104	4.6	0.56	0.30	0.86	-0.4	0.85	-0.4
17	US015	104	4.1	-0.67	0.34	0.47	-1.9	0.47	-1.9
18	US016	104	3.4	-1.72	0.31	0.66	-1.2	0.66	-1.2
19	US017	104	1.6	-3.97	0.33	1.59	2.0	1.57	1.9
20	US017	304	1.5	-4.03	0.30	1.88	3.1	1.78	2.7
21	US018	104	3.8	-0.74	0.31	1.42	1.3	1.42	1.3
22	US019	104	3.7	-0.45	0.31	0.53	-1.8	0.52	-1.8
23	US020	104	4.2	-0.17	0.31	0.94	-0.1	0.95	0
24	US021	104	3.3	-1.18	0.28	1.55	1.9	1.55	1.9
25	US021	304	3.1	-1.70	0.25	1.46	1.7	1.46	1.7
26	US022	106	2.7	-1.75	0.28	0.64	-1.5	0.64	-1.5
27	US022	306	3.2	-1.9	0.25	0.72	-1.2	0.72	-1.2
28	US023	106	4.0	0.35	0.31	0.84	-0.4	0.85	-0.4
29	US024	106	3.1	-0.75	0.31	0.42	-2.5	0.43	-2.4
30	US025	106	2.8	-1.74	0.32	1.03	0.1	1.02	0.1
31	US026	106	3.5	-1.01	0.28	0.67	-1.3	0.66	-1.3
32	US026	306	3.9	-0.83	0.26	0.68	-1.3	0.68	-1.3
33	US027	106	4.3	0.96	0.27	1.43	1.5	1.44	1.6
34	US027	306	4.8	1.38	0.25	1.19	0.8	1.17	0.7
35	US028	106	3.3	-0.68	0.31	0.60	-1.4	0.60	-1.4
36	US029	106	3.3	0.51	0.31	0.82	-0.5	0.82	-0.5
37	US030	106	3.3	-0.57	0.28	1.94	2.9	1.96	2.9
38	US030	306	3.5	-0.90	0.26	2.18	3.7	2.17	3.7
39	US031	106	3.5	0.30	0.31	1.36	1.1	1.36	1.1
40	US032	106	3.1	-0.74	0.31	0.87	-0.3	0.86	-0.4
41	US033	106	3.1	-0.54	0.31	0.47	-2.2	0.47	-2.2
42	US034	106	3.1	-0.36	0.28	1.36	1.3	1.37	1.3
43	US034	306	3.3	-0.87	0.25	1.80	2.8	1.79	2.7
44	US035	106	4.3	-0.12	0.30	1.19	0.7	1.19	0.7
45	US036	106	3.7	-0.34	0.31	0.81	-0.6	0.80	-0.6
46	US037	106	3.8	-0.03	0.31	0.66	-1.2	0.67	-1.1
47	US038	106	3.3	-0.18	0.31	0.65	-1.2	0.64	-1.3
48	US039	106	2.5	-1.68	0.31	1.12	0.4	1.12	0.4
49	US040	106	3.8	-0.13	0.31	0.83	-0.5	0.82	-0.5

50	US041	104	3.8	-0.90	0.31	1.01	0.1	1.01	0.1
51	US042	104	2.6	-1.87	0.30	0.35	-3.0	0.34	-3.0
52	US043	104	3.6	-0.09	0.31	0.90	-0.2	0.90	-0.2
53	US044	104	3.8	-0.60	0.31	1.42	1.3	1.41	1.3
54	US045	104	4.0	-0.04	0.31	0.70	-1.0	0.70	-1.0
55	US046	104	2.8	-1.80	0.27	0.82	-0.6	0.82	-0.6
56	US046	304	2.7	-1.98	0.25	1.12	0.5	1.14	0.6
57	US047	104	3.6	-0.69	0.31	0.40	-2.5	0.40	-2.5
58	US048	104	3.4	-1.59	0.38	0.66	-0.9	0.65	-0.9
59	US049	104	2.7	-2.36	0.27	1.87	2.8	1.86	2.8
60	US049	304	2.7	-2.32	0.25	2.07	3.6	2.04	3.5
61	US050	104	3.0	-1.83	0.34	0.70	-0.9	0.70	-0.9
62	US051	104	4.5	0.90	0.30	1.05	0.2	1.03	0.2
63	US052	104	3.2	-1.46	0.31	1.28	0.9	1.27	0.9
64	US053	104	3.8	-0.74	0.34	0.42	-2.2	0.41	-2.2
65	US054	104	3.6	-1.00	0.31	1.56	1.7	1.56	1.7
66	US055	104	3.0	-2.15	0.31	1.18	0.6	1.19	0.7
67	US056	104	2.9	-2.73	0.31	1.39	1.2	1.39	1.3
68	US057	104	3.7	0.23	0.34	1.99	2.4	1.99	2.4
69	US057	304	3.5	-0.13	0.31	2.08	2.9	2.09	2.9
70	US058	104	2.5	-2.12	0.30	2.43	3.8	2.44	3.8
71	US059	104	3.3	-1.31	0.34	0.34	-2.6	0.35	-2.6
72	US060	104	2.3	-2.76	0.27	1.17	0.7	1.18	0.7
73	US060	304	2.0	-3.30	0.26	1.36	1.5	1.36	1.5
74	US061	104	2.4	-2.70	0.37	0.69	-0.9	0.69	-0.9
75	US062	106	3.7	-0.22	0.31	1.36	1.1	1.36	1.1
76	US063	106	3.9	0.70	0.38	0.65	-0.9	0.63	-1.0
77	US064	106	3.8	0.33	0.28	0.94	-0.1	0.92	-0.2
78	US064	306	4.2	0.43	0.25	1.34	1.3	1.30	1.2
79	US065	106	3.7	-0.64	0.31	0.90	-0.2	0.91	-0.2
80	US066	106	4.5	1.45	0.30	0.72	-1.0	0.71	-1.1
81	US067	106	3.9	1.15	0.31	0.76	-0.8	0.75	-0.8
82	US068	106	3.8	0.07	0.31	0.78	-0.6	0.79	-0.6
83	US069	106	3.5	-0.43	0.31	0.56	-1.6	0.56	-1.6
84	US070	106	2.3	-1.89	0.29	1.61	2.1	1.58	1.9
85	US070	306	2.4	-2.61	0.25	1.64	2.4	1.66	2.5
86	US071	106	3.8	0.58	0.31	1.88	2.5	1.91	2.5
87	US072	106	3.0	-0.58	0.31	0.78	-0.7	0.78	-0.7

88	US073	106	2.7	-1.42	0.31	1.24	0.8	1.24	0.8
89	US074	106	3.6	0.88	0.31	1.13	0.5	1.15	0.5
90	US075	106	4.3	1.58	0.30	1.58	1.8	1.58	1.8
91	US076	106	4.0	0.76	0.31	0.65	-1.3	0.64	-1.3
92	US077	106	2.5	-1.34	0.34	0.53	-1.7	0.53	-1.7
93	US078	106	4.2	0.45	0.31	1.28	0.9	1.25	0.9
94	US079	106	4.1	0.98	0.31	0.69	-1.1	0.69	-1.1
95	US080	106	2.5	-2.09	0.31	0.83	-0.5	0.83	-0.5
96	US081	106	3.5	-0.50	0.31	0.94	-0.1	0.94	-0.1
97	US082	108	3.4	0.54	0.31	0.41	-2.5	0.41	-2.5
98	US083	108	3.1	0.38	0.31	0.47	-2.1	0.47	-2.1
99	US084	108	1.5	-2.43	0.40	0.94	0	0.90	-0.2
100	US085	108	3.0	0.02	0.31	0.68	-1.1	0.69	-1.1
101	US086	108	2.8	-0.50	0.31	0.64	-1.3	0.65	-1.2
102	US087	108	3.0	-0.73	0.31	0.93	-0.1	0.93	-0.1
103	US088	108	3.0	-0.05	0.31	0.52	-1.9	0.52	-1.9
104	US089	108	3.0	0.07	0.31	0.55	-1.8	0.55	-1.7
105	US090	108	3.3	-0.26	0.31	0.68	-1.1	0.68	-1.1
106	US091	108	2.2	-1.19	0.31	1.18	0.7	1.19	0.7
107	US092	108	3.5	0.32	0.31	1.14	0.5	1.12	0.4
108	US093	108	2.8	-1.09	0.31	1.37	1.2	1.40	1.3
109	US094	108	3.7	1.04	0.31	0.55	-1.7	0.55	-1.7
110	US095	108	3.3	0.16	0.31	1.33	1.1	1.34	1.1
111	US096	108	3.1	-0.23	0.31	1.53	1.6	1.52	1.6
112	US097	108	3.6	0.19	0.31	1.20	0.7	1.22	0.8
113	US098	108	2.8	-0.49	0.31	0.52	-1.9	0.52	-1.9
114	US099	108	2.8	-1.08	0.31	0.73	-0.9	0.73	-0.9
115	US100	108	3.0	0	0.31	0.55	-1.7	0.55	-1.7
116	US101	108	2.9	0.01	0.31	1.17	0.6	1.18	0.6
117	US102	108	3.0	0.58	0.31	0.66	-1.2	0.64	-1.3
118	US103	108	2.6	-1.06	0.31	0.84	-0.5	0.83	-0.5
119	US104	108	3.6	0.30	0.31	0.94	-0.1	0.94	-0.1
120	US105	108	2.9	0.04	0.31	0.42	-2.5	0.42	-2.5
121	US106	108	3.6	0.79	0.31	0.75	-0.8	0.75	-0.8
122	US107	108	3.3	0.33	0.31	1.02	0.1	1.03	0.1
123	US108	108	3.7	0.81	0.31	0.68	-1.1	0.68	-1.0
124	US109	108	3.2	-0.46	0.31	0.86	-0.4	0.85	-0.4
125	US110	108	3.2	0.24	0.34	1.10	0.3	1.10	0.3

126	US111	108	3.4	0.15	0.31	0.47	-2.1	0.47	-2.1
127	US112	108	3.3	0.83	0.31	0.89	-0.2	0.89	-0.2
128	US113	108	3.4	-0.07	0.31	0.65	-1.2	0.65	-1.2
129	US114	108	3.1	0.05	0.31	0.62	-1.4	0.62	-1.4
130	US115	108	3.2	-0.26	0.31	0.44	-2.3	0.44	-2.3
131	US116	108	3.1	0.45	0.31	0.55	-1.7	0.55	-1.7
132	US117	108	2.3	-0.01	0.30	0.32	-3.3	0.32	-3.4
133	US118	108	2.3	-1.22	0.30	0.61	-1.6	0.61	-1.5
134	US119	108	3.0	-0.04	0.31	0.41	-2.5	0.41	-2.5
135	US120	108	2.8	-0.06	0.31	0.58	-1.6	0.58	-1.6
136	US121	108	2.1	-1.85	0.31	0.64	-1.4	0.66	-1.3
137	US122	108	2.6	-0.90	0.30	0.60	-1.5	0.60	-1.5
138	US123	108	2.5	-1.53	0.33	0.49	-1.9	0.49	-1.9
139	US124	108	3.2	0.36	0.31	0.28	-3.4	0.27	-3.5
140	US125	108	2.3	-1.01	0.31	0.80	-0.6	0.81	-0.6
141	US126	108	2.5	-0.67	0.31	0.58	-1.6	0.58	-1.6
142	US127	108	2.7	-0.09	0.31	0.65	-1.3	0.65	-1.3
143	US128	108	3.1	0.14	0.31	0.36	-2.8	0.36	-2.8
144	US129	108	3.5	0.82	0.38	1.60	1.5	1.61	1.5
145	US130	108	2.4	-0.35	0.33	0.33	-2.9	0.33	-3.0
146	US131	108	4.6	1.92	0.39	0.46	-1.8	0.48	-1.7
147	US132	108	4.1	1.09	0.31	0.65	-1.2	0.66	-1.2
148	US133	108	4.8	1.82	0.30	1.47	1.6	1.47	1.6
149	US134	108	4.0	0.99	0.31	0.61	-1.4	0.61	-1.4
150	US135	108	5.3	2.64	0.33	2.06	3.0	2.63	4.1
151	US136	108	5.5	3.69	0.37	1.29	0.9	1.22	0.7
152	US137	108	4.9	2.51	0.31	1.49	1.6	1.68	2.1
153	US138	108	3.8	1.10	0.31	0.72	-0.9	0.73	-0.9
154	US139	108	4.5	2.08	0.30	1.05	0.2	1.04	0.2
155	US140	108	4.3	2.03	0.30	0.51	-2.0	0.51	-2.0
156	US141	108	4.1	1.35	0.31	0.47	-2.2	0.46	-2.2
157	US142	108	4.5	1.86	0.30	0.47	-2.4	0.47	-2.4
158	US143	108	4.6	1.44	0.31	0.76	-0.8	0.76	-0.8
159	US144	108	5.4	2.79	0.34	1.29	1.0	1.45	1.4
160	US145	108	4.6	1.99	0.30	1.13	0.5	1.13	0.5
161	US146	108	4.5	1.89	0.30	0.45	-2.4	0.45	-2.4
162	US147	108	4.4	1.18	0.34	1.33	1.0	1.31	1.0
163	US148	108	5.0	2.62	0.32	0.90	-0.2	1.04	0.2

164	US149	108	5.5	3.54	0.40	1.11	0.4	1.01	0.1
165	US150	108	5.4	3.41	0.35	0.75	-0.9	0.92	-0.1
166	US151	108	4.6	2.27	0.30	0.91	-0.2	0.91	-0.2
167	US152	108	5.0	2.56	0.35	1.01	0.1	0.98	0
168	US153	108	2.3	-0.39	0.39	0.34	-2.5	0.37	-2.4
169	US154	108	3.3	-0.10	0.34	1.33	1.0	1.30	0.9
170	US155	108	3.0	-0.63	0.31	0.77	-0.8	0.76	-0.8
171	US156	108	3.3	1.72	0.31	0.40	-2.6	0.40	-2.6
172	US157	108	3.4	0.72	0.31	0.59	-1.5	0.58	-1.6
173	US158	108	3.5	-0.01	0.34	0.75	-0.7	0.75	-0.7
174	US159	108	3.4	0.34	0.31	0.55	-1.7	0.55	-1.7
175	US160	108	2.4	-0.40	0.30	0.35	-3.0	0.35	-3.1
176	US161	108	3.4	-0.03	0.34	2.13	2.7	2.14	2.8
177	US162	108	4.0	1.28	0.31	0.93	-0.1	0.95	0
178	US163	108	4.1	1.51	0.34	0.68	-0.9	0.69	-0.9
179	US164	108	2.1	-1.72	0.32	0.77	-0.8	0.78	-0.7
180	US165	108	2.5	-0.49	0.31	1.13	0.5	1.11	0.4
181	US166	108	1.8	-0.91	0.39	0.86	-0.3	0.84	-0.4
182	US167	108	3.1	0.85	0.31	0.65	-1.2	0.66	-1.2
183	US168	108	3.5	0.30	0.31	0.62	-1.4	0.62	-1.4
184	US169	108	2.4	-1.30	0.34	0.66	-1.1	0.66	-1.1
185	US170	108	2.5	-0.77	0.30	0.73	-0.9	0.73	-0.9
186	US171	108	3.1	-0.05	0.31	1.32	1.0	1.32	1.0
187	US172	108	3.2	0.14	0.31	1.33	1.1	1.34	1.1
188	US173	108	3.5	0.60	0.31	0.72	-0.9	0.73	-0.9
189	US174	108	3.2	-0.46	0.31	0.50	-2.0	0.50	-2.0
190	US175	108	4.0	0.89	0.34	0.57	-1.4	0.56	-1.4
191	US176	108	4.1	0.71	0.31	0.56	-1.7	0.57	-1.6
192	US177	108	3.8	1.18	0.31	0.75	-0.8	0.76	-0.7
193	US178	108	4.2	1.13	0.31	1.10	0.4	1.12	0.4
194	US179	108	4.0	1.54	0.34	1.04	0.2	1.04	0.2
195	US180	108	3.5	0.92	0.31	0.82	-0.5	0.83	-0.5
196	US181	108	3.3	-0.07	0.31	1.01	0.1	1.01	0.1
197	US182	108	3.3	0.28	0.34	0.75	-0.7	0.75	-0.7
198	US183	108	5.0	2.19	0.31	1.39	1.4	1.46	1.6
199	US184	108	3.3	0.71	0.31	0.72	-0.9	0.72	-0.9
200	US185	108	3.6	1.02	0.34	0.88	-0.2	0.88	-0.2
201	US186	108	4.2	1.69	0.31	1.51	1.5	1.49	1.5

202	US187	108	3.4	0.10	0.31	0.51	-1.9	0.51	-1.9
203	US188	108	3.3	0.21	0.31	1.47	1.5	1.48	1.5
204	US189	108	3.9	1.06	0.31	1.85	2.4	1.87	2.4
205	US190	108	3.9	1.41	0.34	0.47	-1.9	0.47	-1.9
206	US191	108	3.6	1.41	0.31	0.72	-0.9	0.72	-0.9
207	US192	108	2.8	-0.22	0.31	0.35	-3.0	0.35	-3.0
208	US193	108	3.0	-0.02	0.31	0.51	-1.9	0.52	-1.9
209	US194	108	3.6	0.33	0.31	0.52	-1.9	0.52	-1.9
210	US195	108	3.6	0.45	0.34	1.33	1.0	1.33	1.0
211	US196	108	4.2	1.10	0.31	0.67	-1.1	0.68	-1.1
212	US197	108	3.6	0.47	0.31	0.67	-1.1	0.67	-1.1
213	US198	108	3.7	0.59	0.31	0.37	-2.6	0.37	-2.6
214	US199	108	3.5	0.18	0.31	0.69	-1.0	0.69	-1.0
215	US200	108	3.1	0	0.31	0.45	-2.2	0.46	-2.2
216	US201	108	3.9	1.42	0.34	0.91	-0.1	0.90	-0.2
217	US202	108	3.4	-0.19	0.31	0.63	-1.3	0.63	-1.3
218	US203	108	3.9	1.26	0.31	1.54	1.6	1.55	1.7
219	US204	108	3.4	1.08	0.31	0.93	-0.1	0.93	-0.1
220	US205	108	4.0	1.16	0.31	0.67	-1.2	0.66	-1.2
221	US206	108	4.3	1.64	0.31	1.74	2.2	1.70	2.0
222	US207	108	4.1	2.56	0.31	0.81	-0.5	0.82	-0.5
223	US208	108	4.8	2.36	0.30	1.19	0.7	1.19	0.7
224	US209	108	4.0	1.15	0.31	0.88	-0.3	0.89	-0.3
225	US210	108	3.8	1.62	0.34	1.02	0.1	1.02	0.1
226	US211	104	4.0	-0.66	0.31	0.46	-2.2	0.46	-2.2
227	US212	104	3.3	-1.46	0.28	1.10	0.4	1.09	0.4
228	US212	304	3.0	-1.65	0.25	1.50	1.9	1.51	1.9
229	US213	104	4.1	-0.33	0.38	0.56	-1.3	0.57	-1.2
230	US214	104	2.9	-2.13	0.31	2.44	3.7	2.49	3.8
231	US215	104	3.7	-0.22	0.28	0.66	-1.3	0.65	-1.4
232	US215	304	3.5	-0.58	0.26	0.87	-0.5	0.86	-0.5
233	US216	104	3.7	-0.14	0.31	0.61	-1.4	0.61	-1.4
234	US217	104	3.5	-0.72	0.31	1.16	0.6	1.15	0.6
235	US218	104	4.2	0.57	0.28	1.34	1.2	1.32	1.2
236	US218	304	4.2	0.97	0.25	1.31	1.2	1.31	1.2
237	US219	104	4.3	0.73	0.31	0.75	-0.8	0.72	-0.9
238	US220	104	3.6	-0.81	0.31	0.83	-0.5	0.83	-0.4
239	US221	104	2.5	-2.53	0.30	1.24	0.8	1.23	0.8

240	US222	104	4.1	-0.19	0.31	0.52	-2.0	0.50	-2.0
241	US223	104	3.8	-0.13	0.31	0.74	-0.8	0.73	-0.8
242	US224	104	2.4	-1.87	0.37	2.08	2.5	2.08	2.5
243	US225	104	3.3	-1.57	0.31	0.87	-0.3	0.87	-0.3
244	US226	104	4.3	0.24	0.31	0.59	-1.6	0.60	-1.5
245	US227	104	3.1	-1.09	0.30	0.93	-0.1	0.92	-0.1
246	US227	304	2.6	-2.00	0.29	0.67	-1.3	0.67	-1.3
247	US228	104	3.0	-1.86	0.31	1.07	0.3	1.07	0.3
248	US229	104	3.5	-0.91	0.38	0.49	-1.6	0.49	-1.6
249	US230	106	3.0	-1.05	0.31	1.17	0.6	1.20	0.7
250	US231	106	3.1	-0.84	0.31	0.74	-0.8	0.74	-0.9
251	US232	106	3.3	-0.67	0.31	0.41	-2.5	0.40	-2.5
252	US233	106	2.7	-1.48	0.31	0.64	-1.3	0.64	-1.3
253	US234	106	3.0	-0.31	0.31	0.86	-0.4	0.87	-0.3
254	US235	106	2.8	-0.31	0.27	1.62	2.1	1.63	2.1
255	US235	306	2.6	-1.32	0.25	1.30	1.2	1.30	1.2
256	US236	106	2.4	-2.42	0.28	1.20	0.8	1.18	0.7
257	US236	306	2.6	-2.98	0.25	1.17	0.7	1.17	0.7
258	US237	106	3.3	-0.41	0.28	1.26	1.0	1.27	1.0
259	US237	306	3.4	-0.86	0.25	1.56	2.0	1.56	2.0
260	US238	106	3.3	-0.50	0.31	0.89	-0.3	0.87	-0.3
261	US239	106	3.5	-0.60	0.31	1.18	0.7	1.17	0.6
262	US240	106	3.1	-0.36	0.31	0.42	-2.4	0.42	-2.4
263	US241	106	3.7	-0.16	0.28	1.18	0.7	1.17	0.7
264	US241	306	4.3	0.26	0.25	0.87	-0.5	0.88	-0.4
265	US242	106	3.7	-0.04	0.31	1.10	0.4	1.11	0.4
266	US243	106	4.0	0.45	0.27	1.10	0.4	1.10	0.4
267	US243	306	4.4	0.30	0.25	1.31	1.2	1.32	1.3
268	US244	106	3.8	0.77	0.31	0.67	-1.1	0.68	-1.1
269	US245	106	2.4	-1.70	0.31	0.85	-0.4	0.85	-0.4
270	US246	106	3.0	-1.18	0.28	0.82	-0.6	0.83	-0.6
271	US246	306	3.6	-1.23	0.26	1.10	0.4	1.10	0.4
272	US247	106	3.7	0.89	0.31	0.62	-1.4	0.62	-1.4
273	US248	106	4.0	0.23	0.28	2.26	3.7	2.22	3.6
274	US248	306	4.9	1.07	0.25	1.89	3.2	1.84	3.1
275	US249	106	3.7	-0.12	0.31	0.81	-0.6	0.81	-0.6
276	US250	104	3.4	-0.46	0.31	0.48	-2.1	0.48	-2.1
277	US251	104	3.2	-1.23	0.31	0.65	-1.2	0.65	-1.2

278	US252	104	3.4	-1.35	0.31	0.77	-0.7	0.77	-0.7
279	US253	104	3.1	-1.43	0.31	0.78	-0.7	0.77	-0.7
280	US254	104	3.9	0.41	0.31	0.48	-2.0	0.48	-2.0
281	US255	104	2.7	-2.17	0.30	0.38	-2.7	0.38	-2.8
282	US256	104	1.8	-3.34	0.32	1.23	0.9	1.20	0.8
283	US256	304	1.8	-3.30	0.33	1.78	2.4	1.70	2.2
284	US257	104	3.4	-0.76	0.31	0.66	-1.2	0.66	-1.2
285	US258	104	3.5	-0.74	0.31	0.64	-1.2	0.64	-1.2
286	US259	104	4.4	-0.35	0.30	0.62	-1.4	0.62	-1.4
287	US260	104	3.0	-1.62	0.31	1.30	1.0	1.29	1.0
288	US261	104	4.1	-0.37	0.31	0.82	-0.5	0.83	-0.5
289	US262	104	2.2	-3.25	0.34	1.23	0.8	1.29	1.0
290	US263	104	2.1	-3.04	0.31	0.46	-2.4	0.46	-2.4
291	US264	104	2.9	-1.36	0.31	0.95	0	0.96	0
292	US265	104	3.9	-0.59	0.31	1.45	1.4	1.44	1.4
293	US266	104	2.4	-2.71	0.27	1.52	1.9	1.53	1.9
294	US266	304	2.1	-3.36	0.25	1.52	2.1	1.50	2.0
295	US267	104	2.5	-2.11	0.30	0.96	0	0.95	0
296	US268	104	2.2	-2.84	0.31	0.93	-0.1	0.92	-0.2
297	US269	104	5.0	1.05	0.30	1.64	2.1	1.63	2.1
298	US270	104	3.9	-0.36	0.34	1.44	1.3	1.44	1.3
299	US271	104	3.2	-1.91	0.34	3.63	5.1	3.65	5.1
300	US272	104	4.5	-0.26	0.30	1.58	1.8	1.56	1.8
301	US273	104	4.1	0	0.28	0.67	-1.2	0.66	-1.3
302	US273	304	4.1	-0.28	0.25	0.89	-0.3	0.88	-0.4
303	US274	104	3.6	-0.51	0.31	0.65	-1.2	0.65	-1.2
304	US275	104	4.7	0.88	0.30	0.48	-2.3	0.47	-2.3
305	US276	104	3.9	-0.43	0.31	0.80	-0.6	0.80	-0.6
306	US277	104	4.4	0.79	0.30	0.55	-1.8	0.56	-1.7
307	US278	104	4.8	0.91	0.30	1.36	1.2	1.35	1.2
308	US279	104	3.5	-0.53	0.31	0.52	-1.8	0.52	-1.8
309	US280	104	4.4	0.90	0.37	1.06	0.2	1.04	0.2
310	US281	104	4.2	0.35	0.31	0.97	0	0.98	0
311	US282	104	4.1	0.35	0.31	0.35	-2.8	0.35	-2.9
312	US283	104	3.8	-0.44	0.31	2.13	2.9	2.14	2.9
313	US284	104	4.1	0.43	0.31	0.97	0	0.99	0
314	US285	104	1.2	-6.26	0.56	1.26	0.6	1.60	1.0
315	US285	304	1.1	-5.90	0.75	1.11	0.3	1.46	0.7

316	US286	104	4.3	0.20	0.30	1.35	1.1	1.36	1.2
317	US287	104	4.3	-0.20	0.30	0.67	-1.2	0.67	-1.2
318	US288	104	4.5	0.47	0.30	0.88	-0.3	0.87	-0.3
319	US289	104	3.0	-1.39	0.31	0.64	-1.3	0.63	-1.3
320	US290	104	3.4	-1.59	0.38	0.58	-1.2	0.58	-1.2
321	US291	106	3.3	-0.56	0.31	1.08	0.3	1.08	0.3
322	US292	106	5.2	1.51	0.33	1.73	2.2	1.52	1.6
323	US293	106	4.3	0.86	0.30	0.85	-0.4	0.84	-0.5
324	US294	106	4.4	1.46	0.33	0.81	-0.5	0.81	-0.5
325	US295	106	3.1	-1.01	0.31	0.69	-1.1	0.69	-1.0
326	US296	106	4.2	0.68	0.31	1.11	0.4	1.15	0.6
327	US297	106	3.9	1.06	0.31	1.17	0.6	1.16	0.6
328	US298	106	4.7	1.40	0.31	1.03	0.2	1.03	0.2
329	US299	106	3.8	-0.12	0.34	1.65	1.8	1.71	1.9
330	US300	106	4.4	1.46	0.30	0.86	-0.4	0.87	-0.3
331	US301	106	3.5	-0.23	0.31	0.85	-0.4	0.84	-0.4
332	US302	106	5.2	2.43	0.35	0.74	-0.8	0.67	-0.9
333	US303	106	4.1	0.91	0.30	0.73	-0.9	0.72	-1.0
334	US304	106	5.0	2.38	0.32	1.11	0.4	1.11	0.4
335	US305	106	4.5	1.24	0.33	0.74	-0.8	0.75	-0.7
336	US306	106	3.0	-0.33	0.38	1.26	0.8	1.29	0.8
337	US307	110	4.8	2.67	0.30	0.58	-1.8	0.57	-1.8
338	US308	110	3.9	1.51	0.31	1.10	0.4	1.10	0.4
339	US309	110	4.1	1.65	0.31	0.46	-2.2	0.45	-2.2
340	US310	110	3.8	1.02	0.31	0.82	-0.5	0.81	-0.5
341	US311	110	4.9	2.33	0.31	0.59	-1.7	0.59	-1.7
342	US312	110	5.0	3.07	0.31	1.02	0.1	0.98	0
343	US313	110	4.7	2.60	0.30	0.71	-1.1	0.70	-1.1
344	US314	110	4.5	1.48	0.30	0.60	-1.6	0.59	-1.6
345	US315	110	5.0	2.40	0.31	0.96	0	0.93	-0.1
346	US316	110	4.0	1.72	0.31	0.78	-0.7	0.77	-0.7
347	US317	110	4.4	1.81	0.31	1.13	0.5	1.12	0.5
348	US318	110	4.1	1.52	0.31	1.09	0.3	1.11	0.4
349	US319	110	4.8	2.35	0.30	0.63	-1.5	0.63	-1.5
350	US320	110	5.3	2.57	0.32	0.83	-0.6	0.82	-0.6
351	US321	110	3.1	-0.08	0.31	0.51	-2.0	0.51	-2.0
352	US322	110	3.0	0.43	0.31	1.25	0.8	1.24	0.8
353	US323	110	2.8	-0.35	0.31	1.15	0.6	1.16	0.6

354	US324	110	4.3	1.83	0.31	0.61	-1.5	0.61	-1.4
355	US325	110	3.3	0.19	0.31	0.51	-1.9	0.51	-1.9
356	US326	110	3.1	0.14	0.34	1.72	1.9	1.71	1.9
357	US327	110	3.1	-0.11	0.31	1.14	0.5	1.13	0.5
358	US328	110	3.9	1.55	0.31	0.61	-1.4	0.61	-1.4
359	US329	110	3.0	-0.10	0.31	0.91	-0.2	0.91	-0.2
360	US330	110	3.2	0.47	0.34	0.28	-3.1	0.28	-3.1
361	US331	110	4.0	1.53	0.31	1.00	0	0.98	0
362	US332	110	2.6	-1.13	0.37	0.93	0	0.93	-0.1
363	US333	110	3.3	0.57	0.31	0.75	-0.8	0.74	-0.8
364	US334	110	4.0	1.24	0.31	0.97	0	0.95	0
365	US335	110	4.9	2.80	0.30	0.70	-1.1	0.70	-1.1
366	US336	110	4.3	1.60	0.31	0.61	-1.4	0.62	-1.4
367	US337	110	3.4	0.66	0.31	0.48	-2.1	0.47	-2.1
368	US338	110	3.5	0.27	0.31	0.50	-2.0	0.50	-2.0
369	US339	110	4.5	2.04	0.30	0.8	-0.7	0.79	-0.7
370	US340	110	3.5	0.24	0.31	1.12	0.4	1.13	0.5
371	US341	110	3.3	0.39	0.31	1.79	2.3	1.78	2.2
372	US342	110	4.8	2.54	0.31	1.00	0.1	0.96	0
373	US343	110	4.5	2.15	0.30	1.70	2.1	1.68	2.0
374	US344	110	4.8	2.11	0.30	1.48	1.6	1.49	1.7
375	US345	110	4.7	2.81	0.30	0.53	-2.0	0.55	-1.9
376	US346	110	4.1	1.47	0.31	0.82	-0.5	0.82	-0.5
377	US347	110	4.1	1.84	0.31	0.53	-1.8	0.53	-1.8
378	US348	110	4.1	1.49	0.31	0.93	-0.1	0.93	-0.1
379	US349	110	3.8	1.15	0.31	1.15	0.5	1.15	0.5
380	US350	110	4.0	1.54	0.31	0.72	-0.9	0.72	-0.9
381	US351	110	4.6	2.14	0.30	1.31	1.1	1.33	1.2
382	US352	110	3.4	0.28	0.31	1.00	0.1	1.01	0.1
383	US353	110	4.8	2.65	0.30	1.41	1.4	1.41	1.4
384	US354	110	4.1	1.47	0.31	1.24	0.8	1.25	0.8
385	US355	110	3.9	0.67	0.31	1.35	1.1	1.35	1.1
386	US356	110	4.5	1.31	0.30	1.47	1.5	1.50	1.6
387	US357	110	4.1	1.30	0.31	1.72	2.1	1.70	2.0
388	US358	110	3.9	1.15	0.31	0.64	-1.2	0.63	-1.3
389	US359	110	4.1	1.45	0.31	1.07	0.3	1.08	0.3
390	US360	110	3.4	-0.46	0.31	2.02	2.8	2.04	2.8
391	US361	110	3.9	1.46	0.31	1.12	0.4	1.12	0.4

392	US362	110	3.3	0.28	0.31	1.39	1.3	1.36	1.2
393	US363	110	3.6	0.96	0.38	0.93	0	0.93	0
394	US364	110	3.0	-0.42	0.31	0.82	-0.5	0.82	-0.5
395	US365	110	2.2	-0.94	0.38	1.29	0.9	1.28	0.8
396	US366	110	3.9	-0.15	0.31	3.19	4.9	3.18	4.8
397	US367	110	3.5	1.03	0.31	0.80	-0.6	0.80	-0.6
398	US368	110	4.3	2.08	0.31	0.94	-0.1	0.94	-0.1
399	US369	110	3.1	0.48	0.31	0.81	-0.5	0.82	-0.5
400	US370	110	3.6	1.03	0.31	2.32	3.3	2.29	3.2
401	US371	110	4.2	1.84	0.31	0.65	-1.2	0.64	-1.2
402	US372	110	3.2	0.35	0.31	1.06	0.3	1.07	0.3
403	US373	110	3.3	-0.26	0.31	0.73	-0.9	0.74	-0.9
404	US374	110	2.8	0.08	0.31	0.85	-0.4	0.85	-0.4
405	US375	110	3.0	-0.54	0.31	0.48	-2.1	0.49	-2.0
406	US376	110	2.6	-0.96	0.34	0.99	0	0.99	0
407	US377	110	3.3	1.31	0.31	1.08	0.3	1.08	0.3
408	US378	110	3.3	0.22	0.31	0.95	0	0.95	0
409	US379	110	3.3	0.51	0.31	1.05	0.2	1.06	0.2
410	US380	110	4.3	2.24	0.31	0.97	0	0.95	0
411	US381	110	3.3	0.11	0.31	0.70	-1.0	0.69	-1.0
412	US382	110	3.1	0.62	0.31	0.48	-2.1	0.48	-2.1
413	US383	110	3.4	0.20	0.34	0.86	-0.3	0.84	-0.4
414	US384	110	3.7	1.59	0.31	0.41	-2.5	0.41	-2.5
415	US385	110	4.4	1.92	0.30	1.00	0	0.98	0
416	US386	110	4.0	1.86	0.31	0.75	-0.8	0.76	-0.7
417	US387	110	4.6	2.17	0.30	1.25	0.9	1.24	0.9
418	US388	110	3.8	1.18	0.31	0.96	0	0.96	0
419	US389	110	4.6	2.98	0.30	1.10	0.4	1.09	0.4
420	US390	110	4.1	1.66	0.31	0.92	-0.1	0.92	-0.1
421	US391	110	4.1	1.40	0.31	0.66	-1.2	0.66	-1.2
422	US392	110	4.2	1.16	0.31	1.04	0.2	1.03	0.1
423	US393	110	4.7	2.26	0.30	1.13	0.5	1.14	0.6
424	US394	110	3.6	1.00	0.31	0.93	-0.1	0.92	-0.1
425	US395	110	3.6	0.51	0.31	0.98	0	0.97	0
426	US396	110	5.3	2.83	0.34	0.73	-0.9	0.72	-1.0
427	US397	110	4.5	2.48	0.30	0.82	-0.6	0.82	-0.6
428	US398	110	4.0	1.38	0.31	0.94	-0.1	0.94	0
429	US399	106	3.5	-0.62	0.31	0.69	-1.0	0.69	-1.0

430	US400	106	3.7	0.04	0.31	0.74	-0.8	0.76	-0.8
431	US401	106	4.3	1.18	0.30	1.81	2.4	1.81	2.4
432	US402	106	3.8	0.56	0.31	1.34	1.1	1.34	1.1
433	US403	106	3.3	-0.96	0.31	0.64	-1.3	0.65	-1.2
434	US404	106	2.3	-1.38	0.28	0.47	-2.5	0.47	-2.5
435	US404	306	2.7	-1.62	0.25	0.65	-1.6	0.66	-1.6
436	US405	106	3.7	0.11	0.31	1.07	0.3	1.09	0.3
437	US406	106	3.4	0.55	0.31	1.04	0.2	1.04	0.2
438	US407	106	2.9	-0.80	0.31	0.43	-2.4	0.43	-2.4
439	US408	106	4.1	0.38	0.31	0.38	-2.8	0.38	-2.7
440	US409	106	4.5	1.33	0.32	0.77	-0.7	0.76	-0.8
441	US410	106	5.2	1.95	0.33	1.24	0.9	1.16	0.6
442	US411	106	4.1	0.94	0.31	0.88	-0.3	0.90	-0.2
443	US412	106	3.4	-0.25	0.31	0.44	-2.3	0.44	-2.3
444	US413	106	2.8	-0.97	0.31	0.76	-0.8	0.77	-0.8
445	US414	106	2.9	-1.50	0.31	0.91	-0.2	0.91	-0.2
446	US415	106	3.5	0.24	0.31	0.78	-0.7	0.79	-0.7
447	US416	106	4.2	0.85	0.31	0.56	-1.7	0.57	-1.6
448	US417	106	3.2	-0.52	0.31	0.42	-2.5	0.42	-2.5
449	US418	104	3.1	-1.00	0.31	0.87	-0.3	0.88	-0.3
450	US419	104	2.9	-1.25	0.27	0.94	-0.1	0.95	0
451	US419	304	2.8	-1.21	0.25	0.45	-2.9	0.46	-2.8
452	US420	104	2.2	-2.64	0.27	0.40	-3.2	0.40	-3.2
453	US420	304	2.2	-2.67	0.25	0.57	-2.2	0.57	-2.2
454	US421	104	3.6	-0.49	0.31	1.29	1.0	1.30	1.0
455	US422	104	3.8	-0.59	0.31	0.77	-0.7	0.76	-0.7
456	US423	104	3.1	-1.03	0.31	1.13	0.5	1.12	0.5
457	US424	104	3.5	-0.39	0.31	0.62	-1.4	0.62	-1.3
458	US425	104	3.9	-0.19	0.31	0.43	-2.3	0.43	-2.3
459	US426	104	3.6	-0.77	0.31	0.71	-0.9	0.72	-0.9
460	US427	104	3.3	-0.99	0.28	1.17	0.6	1.16	0.6
461	US427	304	3.3	-0.77	0.26	1.59	2.1	1.58	2.1
462	US428	104	4.3	-0.03	0.31	1.24	0.8	1.25	0.8
463	US429	104	2.8	-1.62	0.27	0.76	-0.9	0.76	-0.9
464	US429	304	2.6	-1.69	0.25	0.99	0	0.99	0
465	US430	104	3.0	-1.75	0.31	0.75	-0.8	0.75	-0.8
466	US431	104	2.6	-1.75	0.30	0.48	-2.2	0.48	-2.2
467	US432	104	2.2	-2.45	0.37	0.74	-0.7	0.75	-0.7

468	US433	104	3.7	-0.51	0.31	1.35	1.1	1.35	1.1
469	US434	104	3.4	-0.45	0.34	0.66	-1.1	0.66	-1.1
470	US435	104	3.0	-1.62	0.31	1.04	0.2	1.03	0.2
471	US436	104	4.4	0.04	0.27	2.05	3.2	2.04	3.2
472	US436	304	4.4	0.21	0.25	2.84	5.4	2.84	5.4
473	US437	104	2.3	-2.69	0.30	0.73	-1.0	0.72	-1.0
474	US438	106	3.5	-0.02	0.31	0.89	-0.2	0.89	-0.2
475	US439	106	3.2	-0.35	0.28	0.99	0	0.99	0
476	US439	306	3.4	-0.78	0.25	1.11	0.5	1.11	0.5
477	US440	106	5.0	1.38	0.30	0.80	-0.7	1.04	0.2
478	US440	306	5.4	1.12	0.29	0.85	-0.5	1.06	0.3
479	US441	106	3.6	0.52	0.31	0.62	-1.4	0.62	-1.3
480	US442	106	4.4	0.90	0.31	1.80	2.4	1.75	2.2
481	US443	106	4.5	1.73	0.27	1.02	0.1	1.02	0.1
482	US443	306	4.8	1.71	0.25	1.18	0.8	1.17	0.8
483	US444	106	4.8	1.51	0.30	1.11	0.4	1.09	0.4
484	US445	106	3.4	-0.02	0.31	0.60	-1.4	0.61	-1.4
485	US446	106	4.3	0.89	0.30	1.46	1.5	1.44	1.4
486	US447	106	4.5	0.65	0.30	1.47	1.6	1.47	1.6
487	US448	106	2.5	-1.21	0.28	1.50	1.7	1.49	1.8
488	US448	306	2.6	-2.03	0.25	1.12	0.5	1.13	0.6
489	US449	106	4.7	1.53	0.31	1.09	0.3	1.09	0.4
490	US450	106	3.4	0.10	0.28	0.95	-0.1	0.95	0
491	US450	306	3.7	-0.13	0.26	1.10	0.4	1.10	0.4
492	US451	106	4.0	0.46	0.31	2.06	2.8	2.07	2.8
493	US452	106	4.5	1.27	0.30	1.12	0.5	1.12	0.5
494	US453	106	4.0	1.25	0.31	0.78	-0.7	0.79	-0.6
495	US454	106	4.1	0.21	0.28	1.39	1.4	1.36	1.3
496	US454	306	4.3	0.27	0.25	1.79	2.8	1.75	2.6
497	US455	106	2.9	-0.75	0.31	0.79	-0.6	0.80	-0.6
498	US456	106	3.4	-0.32	0.31	1.54	1.6	1.57	1.7
499	US457	106	3.7	0.08	0.34	0.99	0	0.99	0
500	US458	106	3.9	0.22	0.31	0.78	-0.6	0.79	-0.6
501	US459	106	2.8	-0.54	0.27	1.55	1.9	1.56	1.9
502	US459	306	3.1	-0.74	0.25	1.99	3.3	2.01	3.3
503	US460	106	3.1	-0.62	0.28	1.57	1.9	1.57	1.9
504	US460	306	3.4	-0.38	0.26	2.14	3.7	2.15	3.7
505	US461	106	3.3	0.36	0.31	1.13	0.5	1.15	0.5

506	US462	106	3.5	-0.07	0.28	1.20	0.8	1.20	0.8
507	US462	306	4.2	0.52	0.25	1.13	0.5	1.14	0.6
508	US463	106	3.7	0.68	0.28	1.01	0.1	1.01	0.1
509	US463	306	4.3	1.17	0.25	0.56	-2.1	0.55	-2.1
510	US464	106	3.9	0.69	0.31	1.18	0.6	1.18	0.6
511	US465	106	2.3	-1.81	0.32	2.70	4.2	2.64	4.0
512	US466	106	4.2	0.55	0.28	1.84	2.7	2.08	3.3
513	US466	306	4.6	0.29	0.25	2.37	4.4	2.85	5.5
514	US467	106	3.6	0.12	0.31	0.96	0	0.96	0
515	US468	106	3.8	0.41	0.31	0.68	-1.1	0.68	-1.0
516	US469	106	3.5	-0.28	0.31	0.72	-0.9	0.72	-0.9
517	US470	106	4.0	1.26	0.27	1.03	0.1	1.01	0.1
518	US470	306	4.4	1.37	0.25	1.26	1.1	1.26	1.1
519	US471	106	4.2	0.94	0.30	0.76	-0.8	0.76	-0.8
520	US472	106	3.5	0.10	0.31	0.69	-1.0	0.70	-1.0
521	US473	112	4.0	1.01	0.31	0.50	-2.0	0.50	-2.0
522	US474	112	3.9	0.67	0.31	0.50	-2.0	0.50	-2.0
523	US475	112	3.8	1.17	0.31	0.71	-0.9	0.71	-0.9
524	US476	112	3.9	1.28	0.31	0.83	-0.5	0.82	-0.5
525	US477	112	4.1	1.76	0.31	0.79	-0.6	0.77	-0.7
526	US478	112	4.1	1.40	0.31	1.17	0.6	1.19	0.7
527	US479	112	4.2	1.83	0.31	0.49	-2.0	0.49	-2.0
528	US480	112	3.6	0.59	0.31	1.02	0.1	1.01	0.1
529	US481	112	4.3	1.11	0.30	0.65	-1.3	0.65	-1.3
530	US482	112	4.1	1.61	0.31	0.95	0	0.94	0
531	US483	112	4.5	1.57	0.30	1.03	0.1	1.02	0.1
532	US484	112	3.7	0.99	0.31	1.11	0.4	1.10	0.4
533	US485	112	3.8	0.97	0.31	0.81	-0.5	0.82	-0.5
534	US486	112	4.5	2.08	0.31	0.55	-1.8	0.55	-1.8
535	US487	112	3.8	0.82	0.34	0.65	-1.1	0.65	-1.1
536	US488	112	4.2	1.60	0.31	0.81	-0.5	0.79	-0.6
537	US489	112	4.7	2.95	0.30	1.04	0.2	1.02	0.1
538	US490	112	3.8	0.78	0.31	0.87	-0.3	0.88	-0.3
539	US491	112	4.0	1.01	0.31	1.40	1.3	1.42	1.3
540	US492	112	2.8	0.45	0.31	0.73	-0.9	0.73	-0.9
541	US493	112	4.1	1.41	0.31	1.13	0.5	1.10	0.4
542	US494	112	3.4	0.27	0.31	0.71	-1.0	0.71	-0.9
543	US495	112	3.5	0.61	0.31	0.78	-0.7	0.77	-0.7

544	US496	112	3.3	1.11	0.31	0.84	-0.4	0.84	-0.4
545	US497	112	3.5	0.02	0.31	0.92	-0.1	0.92	-0.1
546	US498	112	3.2	0.43	0.31	0.65	-1.2	0.66	-1.2
547	US499	112	2.4	-0.96	0.31	1.24	0.8	1.25	0.9
548	US500	112	3.3	0.31	0.31	0.87	-0.3	0.88	-0.3
549	US501	112	3.9	0.85	0.31	1.30	1.0	1.31	1.0
550	US502	112	2.6	-0.22	0.31	0.80	-0.6	0.79	-0.7
551	US503	112	3.6	1.30	0.31	0.53	-1.8	0.52	-1.8
552	US504	112	2.1	-1.41	0.31	0.66	-1.3	0.66	-1.3
553	US505	112	2.3	-0.65	0.33	0.98	0	0.98	0
554	US506	112	2.8	0.07	0.31	1.13	0.5	1.14	0.5
555	US507	112	3.3	0.41	0.31	0.24	-3.7	0.24	-3.7
556	US508	112	3.4	1.03	0.31	0.90	-0.2	0.89	-0.2
557	US509	112	3.4	0.61	0.31	0.75	-0.8	0.74	-0.8
558	US510	112	3.8	1.07	0.31	0.91	-0.2	0.91	-0.2
559	US511	112	2.9	0.34	0.34	1.23	0.7	1.23	0.7
560	US512	112	3.5	0.58	0.31	0.78	-0.7	0.78	-0.6
561	US513	112	3.7	0.19	0.31	2.38	3.4	2.36	3.4
562	US514	112	3.4	0.81	0.31	0.55	-1.7	0.54	-1.7
563	US515	112	2.9	0.07	0.31	0.60	-1.5	0.60	-1.5
564	US516	112	3.1	0.04	0.31	1.39	1.2	1.39	1.2
565	US517	112	5.1	3.01	0.31	0.78	-0.8	0.79	-0.7
566	US518	112	3.8	0.93	0.31	1.37	1.2	1.38	1.2
567	US519	112	3.8	1.33	0.31	0.90	-0.2	0.90	-0.2
568	US520	112	3.3	0.14	0.31	0.58	-1.6	0.57	-1.6
569	US521	112	4.4	1.82	0.31	1.14	0.5	1.12	0.4
570	US522	112	4.1	1.00	0.31	0.70	-1.0	0.70	-1.0
571	US523	112	5.6	4.47	0.40	0.68	-1.0	0.71	-0.8
572	US524	112	4.3	1.58	0.31	0.79	-0.6	0.80	-0.6
573	US525	112	4.5	2.27	0.30	0.50	-2.1	0.50	-2.0
574	US526	112	4.3	1.24	0.30	0.71	-1.0	0.69	-1.1
575	US527	112	4.2	1.21	0.31	0.85	-0.4	0.85	-0.4
576	US528	112	4.3	1.80	0.30	0.95	0	0.95	0
577	US529	112	3.8	0.98	0.31	0.71	-1.0	0.71	-0.9
578	US530	112	3.9	0.45	0.31	0.93	-0.1	0.93	-0.1
579	US531	112	4.6	2.07	0.30	1.17	0.6	1.16	0.6
580	US532	112	4.1	1.20	0.31	0.94	-0.1	0.94	-0.1
581	US533	112	5.2	3.30	0.32	1.18	0.7	1.19	0.7

582	US534	112	4.3	1.29	0.31	0.80	-0.6	0.79	-0.6
583	US535	112	3.8	0.90	0.31	0.82	-0.5	0.82	-0.5
584	US536	112	4.2	1.43	0.31	1.23	0.8	1.22	0.8
585	US537	112	4.0	1.31	0.31	1.24	0.8	1.24	0.8
586	US538	112	3.5	0.47	0.31	0.90	-0.2	0.90	-0.2
587	US539	112	4.9	2.11	0.30	1.32	1.1	1.29	1.0
588	US540	112	5.2	2.75	0.31	1.44	1.6	1.46	1.6
589	US541	112	4.0	1.20	0.31	0.79	-0.6	0.81	-0.5
590	US542	112	3.9	1.31	0.31	1.07	0.3	1.07	0.3
591	US543	112	4.4	2.49	0.30	0.52	-1.9	0.53	-1.9
592	US544	112	4.6	2.52	0.31	0.75	-0.8	0.75	-0.9
593	US545	112	4.1	1.19	0.31	0.82	-0.5	0.83	-0.5
594	US546	112	4.1	1.56	0.31	0.42	-2.4	0.42	-2.4
595	US547	112	3.8	0.97	0.31	0.41	-2.5	0.41	-2.5
596	US548	112	4.3	0.74	0.31	0.43	-2.4	0.42	-2.4
597	US549	112	3.3	0.37	0.31	0.48	-2.1	0.49	-2.0
598	US550	112	3.9	1.34	0.34	0.88	-0.2	0.87	-0.2
599	US551	112	3.5	-0.06	0.31	0.71	-1.0	0.71	-0.9
600	US552	112	4.5	2.22	0.30	0.79	-0.7	0.79	-0.7
601	US553	112	4.4	1.47	0.30	0.62	-1.4	0.61	-1.5
602	US554	112	3.6	0.49	0.31	0.64	-1.2	0.65	-1.2
603	US555	112	4.9	2.69	0.33	0.62	-1.4	0.62	-1.4
604	US556	112	3.9	1.07	0.31	0.50	-1.9	0.50	-1.9
605	US557	112	3.5	1.18	0.31	0.55	-1.7	0.55	-1.7
606	US558	112	2.5	-1.23	0.31	1.17	0.6	1.20	0.7
607	US559	112	3.1	1.20	0.31	0.80	-0.6	0.80	-0.6
608	US560	112	3.3	-0.44	0.31	0.37	-2.7	0.37	-2.7
609	US561	112	2.2	-1.03	0.30	0.74	-0.9	0.73	-1.0
610	US562	112	3.3	0.33	0.31	0.71	-0.9	0.72	-0.9
611	US563	112	2.5	-1.32	0.30	0.59	-1.6	0.59	-1.6
612	US564	112	3.5	0.88	0.31	0.90	-0.2	0.91	-0.2
613	US565	112	3.9	1.12	0.31	0.72	-0.9	0.73	-0.9
614	US566	112	3.2	-0.44	0.31	0.51	-1.9	0.50	-1.9
615	US567	112	3.8	0.80	0.31	1.25	0.8	1.26	0.8
616	US568	112	2.3	-1.65	0.31	1.10	0.4	1.11	0.4
617	US569	112	2.9	-0.22	0.31	0.74	-0.8	0.74	-0.9
618	US570	112	2.5	-1.45	0.30	1.59	1.9	1.58	1.8
619	US571	104	4.7	1.02	0.30	0.90	-0.2	0.90	-0.3

620	US572	104	3.2	-1.44	0.31	2.18	3.2	2.19	3.2
621	US573	104	4.5	0.82	0.30	1.55	1.8	1.59	1.8
622	US574	104	3.5	-1.06	0.31	0.79	-0.6	0.79	-0.6
623	US575	104	3.8	-0.28	0.31	0.75	-0.8	0.75	-0.8
624	US576	104	4.0	0.27	0.31	0.66	-1.2	0.65	-1.2
625	US577	104	3.1	-1.63	0.29	3.52	5.8	3.61	5.9
626	US577	304	3.3	-1.10	0.27	1.28	1.0	1.28	1.1
627	US578	104	5.0	1.60	0.35	0.98	0	1.04	0.2
628	US579	104	3.8	-0.78	0.31	6.12	8.6	6.10	8.5
629	US580	104	3.5	-0.87	0.31	1.04	0.2	1.05	0.2
630	US581	104	3.1	-1.83	0.31	1.92	2.5	1.92	2.6
631	US582	104	3.9	0.02	0.34	0.77	-0.6	0.77	-0.6
632	US583	104	4.0	0.20	0.31	0.70	-1.0	0.69	-1.1
633	US584	104	3.8	-0.52	0.31	0.45	-2.2	0.45	-2.2
634	US585	104	3.5	-0.96	0.31	1.47	1.4	1.48	1.5
635	US586	104	2.7	-1.77	0.27	1.02	0.1	1.02	0.1
636	US586	304	2.7	-1.55	0.25	0.89	-0.4	0.90	-0.3
637	US587	104	2.7	-1.87	0.30	0.99	0	0.98	0
638	US588	104	3.5	-0.03	0.34	0.51	-1.7	0.51	-1.7
639	US589	104	3.5	-1.13	0.31	0.84	-0.4	0.84	-0.4
640	HK001	206	1.3	-4.56	1.16	0.52	-0.3	0.46	-0.4
641	HK002	206	3.2	-1.32	0.62	0.94	0	0.92	0
642	HK003	206	3.8	-1.22	0.62	0.93	0	0.95	0.1
643	HK004	206	3.3	-0.76	0.62	0.67	-0.4	0.66	-0.4
644	HK005	206	1.3	-5.71	1.15	0.79	0	0.54	0
645	HK006	206	3.5	-1.34	0.63	0.58	-0.6	0.58	-0.6
646	HK007	206	5.3	1.84	0.68	0.43	-1.1	0.45	-1.0
647	HK008	206	3.2	-0.92	0.62	0.68	-0.4	0.67	-0.4
648	HK009	206	3.7	-0.99	0.63	1.60	1.0	1.61	1.0
649	HK010	206	4.8	1.06	0.60	0.37	-1.4	0.39	-1.3
650	HK011	206	2.3	-2.41	0.61	2.46	2.0	2.51	2.1
651	HK012	206	4.7	1.11	0.6	0.47	-1.0	0.47	-1.0
652	HK013	206	3.7	-0.40	0.62	0.83	-0.1	0.83	-0.1
653	HK014	206	4.2	0.55	0.62	0.95	0.1	0.93	0
654	HK015	206	5.3	1.24	0.71	0.63	-0.5	0.72	-0.2
655	HK016	206	4.2	-0.22	0.62	0.23	-1.8	0.22	-1.8
656	HK017	206	3.0	-2.38	0.62	0.38	-1.2	0.38	-1.2
657	HK018	206	3.3	-1.55	0.62	0.18	-2.0	0.18	-2.0

658	HK019	206	3.8	-0.67	0.62	2.54	2.0	2.60	2.1
659	HK020	206	3.0	-1.70	0.62	0.34	-1.3	0.35	-1.3
660	HK021	206	3.8	-1.05	0.62	8.64	5.5	8.34	5.4
661	HK022	206	1.6	-3.92	0.76	0.47	-0.9	0.56	-0.7
662	HK023	206	4.0	-1.24	0.62	1.15	0.4	1.11	0.3
663	HK024	206	4.8	1.81	0.60	0.84	-0.1	0.88	0
664	HK025	206	4.7	0.90	0.61	0.61	-0.6	0.61	-0.6
665	HK026	206	3.7	-0.17	0.63	0.41	-1.0	0.41	-1.0
666	HK027	206	2.8	-1.91	0.61	2.10	1.6	2.14	1.6
667	HK028	206	3.2	-1.60	0.62	0.30	-1.5	0.30	-1.5
668	HK029	206	3.0	-1.53	0.62	0.80	-0.1	0.79	-0.1
669	HK030	206	4.2	-0.65	0.62	0.52	-0.8	0.53	-0.7
670	HK031	206	3.2	-2.18	0.62	0.59	-0.6	0.60	-0.6
671	HK032	206	4.8	0.01	0.61	0.64	-0.5	0.65	-0.5
672	HK033	206	3.0	-1.10	0.62	0.20	-1.9	0.19	-1.9
673	HK034	206	4.2	0.32	0.62	1.43	0.8	1.48	0.9
674	HK035	206	3.3	-1.56	0.62	0.15	-2.1	0.16	-2.1
675	HK036	206	3.8	-1.60	0.68	5.47	3.7	5.54	3.7
676	HK037	206	3.0	-2.71	0.76	0.91	0.1	0.91	0.1
677	HK038	206	3.7	-1.18	0.62	0.63	-0.5	0.63	-0.5
678	HK039	206	3.0	-1.10	0.62	1.07	0.3	1.10	0.3
679	HK040	206	4.0	-0.06	0.62	0.58	-0.6	0.58	-0.6
680	HK041	206	2.8	-2.49	0.61	0.65	-0.5	0.65	-0.5
681	HK042	206	4.2	-0.86	0.61	0.94	0	0.96	0.1
682	HK043	206	3.0	-1.98	0.62	0.25	-1.7	0.24	-1.7
683	HK044	206	3.8	0.19	0.62	2.77	2.2	2.77	2.2
684	HK045	206	3.2	-1.54	0.62	0.48	-0.9	0.47	-0.9
685	HK046	206	2.8	-2.10	0.61	0.18	-2.0	0.18	-2.0
686	HK047	206	3.5	-0.77	0.62	0.33	-1.3	0.33	-1.3
687	HK048	206	5.2	2.85	0.65	2.67	2.3	3.03	2.5
688	HK049	206	3.7	-0.38	0.63	2.34	1.8	2.34	1.8
689	HK050	206	3.5	-1.41	0.77	0.74	-0.1	0.74	-0.1
690	HK051	206	2.8	-1.68	0.61	0.70	-0.3	0.70	-0.3
691	HK052	206	3.2	-1.83	0.62	0.31	-1.4	0.31	-1.4
692	HK053	206	2.8	-1.90	0.61	0.83	-0.1	0.84	-0.1
693	HK054	206	3.8	0.18	0.62	2.09	1.6	2.03	1.5
694	HK055	206	3.8	-0.77	0.63	2.28	1.7	2.31	1.8
695	HK056	206	2.5	-2.24	0.61	1.45	0.8	1.46	0.9

696	HK057	206	2.0	-3.35	0.63	0.33	-1.5	0.34	-1.5
697	HK058	206	3.3	-0.99	0.62	2.59	2.1	2.62	2.1
698	HK059	206	3.3	-0.76	0.62	2.20	1.7	2.19	1.7
699	HK060	206	2.7	-2.08	0.61	0.54	-0.8	0.54	-0.7
700	HK061	206	2.8	-1.91	0.61	0.45	-1.0	0.45	-1.0
701	HK062	206	2.7	-2.91	0.61	2.49	2.0	2.55	2.1
702	HK063	206	4.2	-0.62	0.62	2.74	2.2	2.74	2.2
703	HK064	206	2.5	-3.10	0.61	1.14	0.4	1.16	0.4
704	HK065	206	4.0	0	0.62	1.99	1.5	1.90	1.4
705	HK066	206	3.2	-0.74	0.62	1.00	0.1	1.02	0.2
706	HK067	206	4.0	-0.38	0.62	0.87	0	0.88	0
707	HK068	206	2.8	-2.36	0.61	0.74	-0.3	0.74	-0.3
708	HK069	206	2.8	-2.48	0.61	1.99	1.5	1.94	1.4
709	HK070	206	4.0	-0.01	0.62	1.45	0.8	1.47	0.9
710	HK071	206	2.3	-3.82	0.61	0.57	-0.7	0.57	-0.7
711	HK072	206	2.3	-2.56	0.87	2.41	1.5	2.37	1.5
712	HK073	206	2.8	-1.69	0.61	2.31	1.8	2.28	1.8
713	HK074	206	2.0	-3.59	0.64	1.84	1.4	1.74	1.3
714	HK075	206	4.2	-0.22	0.62	0.90	0	0.92	0
715	HK076	206	1.8	-3.99	0.65	1.30	0.6	1.22	0.5
716	HK077	206	2.2	-3.65	0.69	1.55	0.9	1.56	1.0
717	HK078	206	2.4	-4.13	0.67	2.99	2.3	2.98	2.3
718	HK079	206	3.8	-0.39	0.62	0.88	0	0.87	0
719	HK080	206	3.0	-1.75	0.62	2.52	2.0	2.47	2.0
720	HK081	206	1.0	-5.65	1.95	Minim	um		
721	HK082	206	1.0	-5.77	1.94	Minim	um		
722	HK083	206	3.5	-1.21	0.62	0.90	0	0.91	0
723	HK084	206	4.3	0.05	0.75	3.60	2.5	3.79	2.6
724	HK085	206	2.5	-2.22	0.61	0.24	-1.8	0.23	-1.8
725	HK086	206	2.2	-3.25	0.68	6.62	4.6	6.44	4.5
726	HK087	206	3.7	-0.20	0.63	1.59	1.0	1.57	1.0
727	HK088	206	1.5	-4.73	0.74	0.92	0	0.97	0.1
728	HK089	206	1.8	-3.94	0.72	0.38	-1.2	0.39	-1.1
729	HK090	206	3.3	-1.16	0.62	2.83	2.3	2.84	2.3
730	HK091	206	2.2	-3.81	0.63	1.48	0.9	1.45	0.9
731	HK092	206	2.3	-2.62	0.61	1.05	0.2	1.05	0.2
732	HK093	206	2.7	-2.74	0.61	0.34	-1.4	0.33	-1.4
733	HK094	206	4.0	-1.05	0.62	2.17	1.7	2.22	1.7

734	HK095	206	5.0	1.78	0.60	1.58	1.1	1.52	1.0
735	HK096	206	3.5	-1.40	0.63	0.12	-2.3	0.12	-2.3
736	HK097	206	3.3	-0.76	0.62	0.94	0	0.93	0
737	HK098	206	1.0	-7.25	1.91	Minim	um		
738	HK099	206	2.8	-2.32	0.61	0.81	-0.1	0.82	-0.1
739	HK100	206	2.3	-2.62	0.61	0.78	-0.2	0.78	-0.2
740	HK101	206	3.7	-0.78	0.63	0.72	-0.3	0.72	-0.3
741	HK102	206	2.3	-3.43	0.62	0.29	-1.6	0.29	-1.6
742	HK103	206	3.8	-0.56	0.62	0.98	0.1	0.99	0.1
743	HK104	206	2.6	-2.34	0.67	1.03	0.2	1.04	0.2
744	HK105	206	2.7	-2.51	0.61	0.57	-0.7	0.56	-0.7
745	HK106	206	2.0	-4.05	0.63	0.81	-0.2	0.82	-0.1
746	HK107	206	2.5	-2.45	0.61	0.93	0	0.94	0
747	HK108	206	1.4	-5.21	0.88	0.97	0.1	0.83	0
748	HK109	206	4.0	0.80	0.62	0.44	-1.0	0.45	-0.9
749	HK110	206	2.8	-2.58	0.61	1.51	0.9	1.55	1.0
750	HK111	206	2.5	-2.88	0.61	0.51	-0.8	0.52	-0.8
751	HK112	206	3.3	-0.95	0.62	3.03	2.4	2.98	2.4
752	HK113	206	4.8	1.47	0.61	1.42	0.8	1.41	0.8
753	HK114	206	3.8	-0.27	0.62	4.48	3.5	4.49	3.5
754	HK115	206	3.2	-1.60	0.62	1.80	1.2	1.80	1.3
755	HK116	206	4.2	0.73	0.62	0.44	-1.0	0.42	-1.1
756	HK117	206	3.0	-1.75	0.62	2.21	1.7	2.19	1.7
757	HK118	206	3.8	-0.44	0.62	0.73	-0.3	0.74	-0.3
758	HK119	206	3.7	-0.55	0.62	0.79	-0.1	0.78	-0.2
759	HK120	206	3.0	-1.94	0.62	1.26	0.6	1.23	0.5
760	HK121	206	3.0	-1.30	0.62	2.41	1.9	2.45	1.9
761	HK122	206	2.0	-3.83	0.64	0.67	-0.5	0.70	-0.4
762	HK123	206	3.3	-0.98	0.62	0.26	-1.6	0.26	-1.6
763	HK124	206	1.3	-5.14	0.84	0.35	-1.1	0.30	-1.0
764	HK125	206	3.2	-1.54	0.62	1.05	0.2	1.05	0.2
765	HK126	206	4.7	1.10	0.63	1.50	0.9	1.86	1.3
766	HK127	206	5.0	2.10	0.66	2.28	1.8	2.24	1.7
767	HK128	206	2.7	-1.87	0.61	2.30	1.8	2.36	1.9
768	HK129	206	5.2	1.88	0.64	0.43	-1.1	0.73	-0.3
769	HK130	206	4.2	-0.24	0.62	1.11	0.3	1.16	0.4
770	HK131	206	1.8	-3.93	0.81	0.20	-1.7	0.23	-1.5
771	HK132	206	4.0	-0.06	0.62	0.62	-0.5	0.61	-0.5

772	HK133	206	3.0	-1.33	0.62	1.39	0.7	1.38	0.7
773	HK134	206	3.0	-1.75	0.62	0.34	-1.3	0.34	-1.3
774	HK135	206	5.8	3.17	1.13	1.48	0.7	1.79	1.0
775	HK136	206	3.0	-2.15	0.62	0.68	-0.4	0.68	-0.4
776	HK137	206	2.3	-2.79	0.61	0.89	0	0.91	0

Separation: 3.75 Reliability: 0.93

Fixed (all same) chi-square: 13749.9 significance *p*: .00

Note: 104 = Grade 4 US student, 106 = Grade 6 US student, 108 = Grade 8 US student, 110 = Grade 10 US student, 112 = Grade 12 US student, 206 = Grade 6 HK student, 304 = Grade 4 US sub-sample student, 306 = Grade 6 US sub-sample student

Calibration of the students for all student essays using US and HK raters and Lexile Analyzer (130 misfitting students)

Student		Observed	Measure	S.E.	Infit		Outfit	
Number	Group	score	(Logit)		MnSq	ZStd	MnSq	ZStd
383	110	4.8	2.65	0.30	1.41	1.4	1.41	1.4
21	104	3.8	-0.74	0.31	1.42	1.3	1.42	1.3
53	104	3.8	-0.60	0.31	1.42	1.3	1.41	1.3
752	206	4.8	1.47	0.61	1.42	0.8	1.41	0.8
33	106	4.3	0.96	0.27	1.43	1.5	1.44	1.6
673	206	4.2	0.32	0.62	1.43	0.8	1.48	0.9
298	104	3.9	-0.36	0.34	1.44	1.3	1.44	1.3
588	112	5.2	2.75	0.31	1.44	1.6	1.46	1.6
292	104	3.9	-0.59	0.31	1.45	1.4	1.44	1.4
695	206	2.5	-2.24	0.61	1.45	0.8	1.46	0.9
709	206	4.0	-0.01	0.62	1.45	0.8	1.47	0.9
25	304	3.1	-1.70	0.25	1.46	1.7	1.46	1.7
485	106	4.3	0.89	0.30	1.46	1.5	1.44	1.4
148	108	4.8	1.82	0.30	1.47	1.6	1.47	1.6
203	108	3.3	0.21	0.31	1.47	1.5	1.48	1.5
386	110	4.5	1.31	0.30	1.47	1.5	1.50	1.6
486	106	4.5	0.65	0.30	1.47	1.6	1.47	1.6
634	104	3.5	-0.96	0.31	1.47	1.4	1.48	1.5
374	110	4.8	2.11	0.30	1.48	1.6	1.49	1.7
730	206	2.2	-3.81	0.63	1.48	0.9	1.45	0.9
774	206	5.8	3.17	1.13	1.48	0.7	1.79	1.0
152	108	4.9	2.51	0.31	1.49	1.6	1.68	2.1
228	304	3.0	-1.65	0.25	1.50	1.9	1.51	1.9

487	106	2.5	-1.21	0.28	1.50	1.7	1.49	1.8
765	206	4.7	1.10	0.63	1.50	0.9	1.86	1.3
201	108	4.2	1.69	0.31	1.51	1.5	1.49	1.5
749	206	2.8	-2.58	0.61	1.51	0.9	1.55	1.0
293	104	2.4	-2.71	0.27	1.52	1.9	1.53	1.9
294	304	2.1	-3.36	0.25	1.52	2.1	1.50	2.0
111	108	3.1	-0.23	0.31	1.53	1.6	1.52	1.6
218	108	3.9	1.26	0.31	1.54	1.6	1.55	1.7
498	106	3.4	-0.32	0.31	1.54	1.6	1.57	1.7
24	104	3.3	-1.18	0.28	1.55	1.9	1.55	1.9
501	106	2.8	-0.54	0.27	1.55	1.9	1.56	1.9
621	104	4.5	0.82	0.30	1.55	1.8	1.59	1.8
716	206	2.2	-3.65	0.69	1.55	0.9	1.56	1.0
65	104	3.6	-1.00	0.31	1.56	1.7	1.56	1.7
259	306	3.4	-0.86	0.25	1.56	2.0	1.56	2.0
503	106	3.1	-0.62	0.28	1.57	1.9	1.57	1.9
90	106	4.3	1.58	0.30	1.58	1.8	1.58	1.8
300	104	4.5	-0.26	0.30	1.58	1.8	1.56	1.8
734	206	5.0	1.78	0.60	1.58	1.1	1.52	1.0
19	104	1.6	-3.97	0.33	1.59	2.0	1.57	1.9
461	304	3.3	-0.77	0.26	1.59	2.1	1.58	2.1
618	112	2.5	-1.45	0.30	1.59	1.9	1.58	1.8
726	206	3.7	-0.20	0.63	1.59	1.0	1.57	1.0
144	108	3.5	0.82	0.38	1.60	1.5	1.61	1.5
648	206	3.7	-0.99	0.63	1.60	1.0	1.61	1.0
84	106	2.3	-1.89	0.29	1.61	2.1	1.58	1.9
254	106	2.8	-0.31	0.27	1.62	2.1	1.63	2.1
85	306	2.4	-2.61	0.25	1.64	2.4	1.66	2.5
297	104	5.0	1.05	0.30	1.64	2.1	1.63	2.1
329	106	3.8	-0.12	0.34	1.65	1.8	1.71	1.9
373	110	4.5	2.15	0.30	1.70	2.1	1.68	2.0
356	110	3.1	0.14	0.34	1.72	1.9	1.71	1.9
387	110	4.1	1.30	0.31	1.72	2.1	1.70	2.0
322	106	5.2	1.51	0.33	1.73	2.2	1.52	1.6
221	108	4.3	1.64	0.31	1.74	2.2	1.70	2.0
283	304	1.8	-3.30	0.33	1.78	2.4	1.70	2.2
371	110	3.3	0.39	0.31	1.79	2.3	1.78	2.2
496	306	4.3	0.27	0.25	1.79	2.8	1.75	2.6

43	306	3.3	-0.87	0.25	1.80	2.8	1.79	2.7
480	106	4.4	0.90	0.31	1.80	2.4	1.75	2.2
754	206	3.2	-1.60	0.62	1.80	1.2	1.80	1.3
431	106	4.3	1.18	0.30	1.81	2.4	1.81	2.4
512	106	4.2	0.55	0.28	1.84	2.7	2.08	3.3
713	206	2.0	-3.59	0.64	1.84	1.4	1.74	1.3
204	108	3.9	1.06	0.31	1.85	2.4	1.87	2.4
59	104	2.7	-2.36	0.27	1.87	2.8	1.86	2.8
20	304	1.5	-4.03	0.30	1.88	3.1	1.78	2.7
86	106	3.8	0.58	0.31	1.88	2.5	1.91	2.5
274	306	4.9	1.07	0.25	1.89	3.2	1.84	3.1
630	104	3.1	-1.83	0.31	1.92	2.5	1.92	2.6
37	106	3.3	-0.57	0.28	1.94	2.9	1.96	2.9
68	104	3.7	0.23	0.34	1.99	2.4	1.99	2.4
502	306	3.1	-0.74	0.25	1.99	3.3	2.01	3.3
704	206	4.0	0	0.62	1.99	1.5	1.90	1.4
708	206	2.8	-2.48	0.61	1.99	1.5	1.94	1.4
390	110	3.4	-0.46	0.31	2.02	2.8	2.04	2.8
471	104	4.4	0.04	0.27	2.05	3.2	2.04	3.2
150	108	5.3	2.64	0.33	2.06	3.0	2.63	4.1
492	106	4.0	0.46	0.31	2.06	2.8	2.07	2.8
60	304	2.7	-2.32	0.25	2.07	3.6	2.04	3.5
69	304	3.5	-0.13	0.31	2.08	2.9	2.09	2.9
242	104	2.4	-1.87	0.37	2.08	2.5	2.08	2.5
693	206	3.8	0.18	0.62	2.09	1.6	2.03	1.5
666	206	2.8	-1.91	0.61	2.10	1.6	2.14	1.6
176	108	3.4	-0.03	0.34	2.13	2.7	2.14	2.8
312	104	3.8	-0.44	0.31	2.13	2.9	2.14	2.9
504	306	3.4	-0.38	0.26	2.14	3.7	2.15	3.7
733	206	4.0	-1.05	0.62	2.17	1.7	2.22	1.7
38	306	3.5	-0.90	0.26	2.18	3.7	2.17	3.7
620	104	3.2	-1.44	0.31	2.18	3.2	2.19	3.2
698	206	3.3	-0.76	0.62	2.20	1.7	2.19	1.7
756	206	3.0	-1.75	0.62	2.21	1.7	2.19	1.7
273	106	4.0	0.23	0.28	2.26	3.7	2.22	3.6
694	206	3.8	-0.77	0.63	2.28	1.7	2.31	1.8
766	206	5.0	2.10	0.66	2.28	1.8	2.24	1.7
767	206	2.7	-1.87	0.61	2.30	1.8	2.36	1.9

712	206	2.8	-1.69	0.61	2.31	1.8	2.28	1.8
400	110	3.6	1.03	0.31	2.32	3.3	2.29	3.2
688	206	3.7	-0.38	0.63	2.34	1.8	2.34	1.8
513	306	4.6	0.29	0.25	2.37	4.4	2.85	5.5
561	112	3.7	0.19	0.31	2.38	3.4	2.36	3.4
711	206	2.3	-2.56	0.87	2.41	1.5	2.37	1.5
760	206	3.0	-1.30	0.62	2.41	1.9	2.45	1.9
70	104	2.5	-2.12	0.30	2.43	3.8	2.44	3.8
230	104	2.9	-2.13	0.31	2.44	3.7	2.49	3.8
650	206	2.3	-2.41	0.61	2.46	2.0	2.51	2.1
701	206	2.7	-2.91	0.61	2.49	2.0	2.55	2.1
719	206	3.0	-1.75	0.62	2.52	2.0	2.47	2.0
658	206	3.8	-0.67	0.62	2.54	2.0	2.60	2.1
697	206	3.3	-0.99	0.62	2.59	2.1	2.62	2.1
687	206	5.2	2.85	0.65	2.67	2.3	3.03	2.5
511	106	2.3	-1.81	0.32	2.70	4.2	2.64	4.0
702	206	4.2	-0.62	0.62	2.74	2.2	2.74	2.2
683	206	3.8	0.19	0.62	2.77	2.2	2.77	2.2
729	206	3.3	-1.16	0.62	2.83	2.3	2.84	2.3
472	304	4.4	0.21	0.25	2.84	5.4	2.84	5.4
717	206	2.4	-4.13	0.67	2.99	2.3	2.98	2.3
751	206	3.3	-0.95	0.62	3.03	2.4	2.98	2.4
396	110	3.9	-0.15	0.31	3.19	4.9	3.18	4.8
625	104	3.1	-1.63	0.29	3.52	5.8	3.61	5.9
723	206	4.3	0.05	0.75	3.60	2.5	3.79	2.6
299	104	3.2	-1.91	0.34	3.63	5.1	3.65	5.1
753	206	3.8	-0.27	0.62	4.48	3.5	4.49	3.5
675	206	3.8	-1.60	0.68	5.47	3.7	5.54	3.7
628	104	3.8	-0.78	0.31	6.12	8.6	6.10	8.5
725	206	2.2	-3.25	0.68	6.62	4.6	6.44	4.5
660	206	3.8	-1.05	0.62	8.64	5.5	8.34	5.4

Separation: 3.75 Reliability: 0.93

Fixed (all same) chi-square: 13749.9 significance *p*: .00

Note: 104 = Grade 4 US student, 106 = Grade 6 US student, 108 = Grade 8 US student, 110 = Grade 10 US student, 112 = Grade 12 US student, 206 = Grade 6 HK student, 304 = Grade 4 US sub-sample student, 306 = Grade 6 US sub-sample student

Calibration of the prompts for all student essays using US and HK raters and Lexile Analyzer

Prompt	Observed score	Measure (Logit)	S.E.	Infit		Outfit	
				MnSq	ZStd	MnSq	ZStd
4AI	3.6	-2.17	0.06	0.72	-5.0	0.72	-5.1
4AN	3.7	-2.23	0.07	0.80	-3.4	0.82	-3.1
4AP	3.6	-2.35	0.07	1.52	7.3	1.53	7.4
4BI	3.6	-1.86	0.03	1.16	4.8	1.16	4.7
4BN	3.6	-1.63	0.03	0.93	-2.2	0.96	-1.2
4BP	3.5	-1.89	0.03	1.44	9.0	1.43	9.0
8AI	3.5	0.11	0.05	0.98	-0.3	1.00	0
8AN	3.7	0.12	0.05	0.73	-6.5	0.75	-6.1
8AP	2.9	1.45	0.05	0.89	-2.5	0.89	-2.4
8BI	3.8	0.39	0.05	0.91	-1.8	0.91	-1.9
8BN	3.7	1.07	0.05	0.66	-8.0	0.67	-7.7
8BP	3.6	0.73	0.05	1.09	1.7	1.09	1.7
12AI	3.8	1.48	0.06	0.88	-2.3	0.88	-2.4
12AN	3.9	0.82	0.06	0.80	-4.2	0.79	-4.3
12AP	3.4	1.96	0.06	0.90	-2.0	0.90	-2.0
12BI	3.5	1.84	0.08	0.88	-1.6	0.88	-1.6
12BN	4.0	1.18	0.08	0.94	-0.8	0.93	-0.9
12BP	4.0	0.98	0.08	1.02	0.3	1.03	0.3

Separation: 26.59 Reliability: 1.00

Fixed (all same) chi-square: 15712.8 significance p: .00

Calibration of the US raters for all student essays using US and HK raters and Lexile Analyzer

US rater	Observed score	Measure (Logit)	S.E.	Infit		Outfit	
				MnSq	ZStd	MnSq	ZStd
USR101	3.8	-0.17	0.05	0.76	-5.4	0.79	-4.6
USR102	3.8	0.02	0.05	0.75	-6.1	0.76	-5.9
USR103	3.6	0.52	0.09	0.70	-4.1	0.71	-3.9
USR104	3.6	0.54	0.05	0.69	-7.6	0.71	-7.2
USR105	3.8	0.04	0.05	0.79	-5.0	0.78	-5.1
USR106	4.0	-0.52	0.06	0.72	-5.5	0.73	-5.2
USR108	3.3	-1.39	0.35	0.57	-1.4	0.57	-1.4
USR200	3.2	0.95	0.06	1.15	2.5	1.15	2.5
USR201	3.6	-0.10	0.05	1.07	1.4	1.07	1.4

USR202	3.6	-0.32	0.06	1.03	0.5	1.02	0.3
USR203	3.5	0.02	0.04	1.05	1.2	1.05	1.1
USR204	3.6	0.03	0.05	0.87	-3.0	0.87	-2.9
USR205	3.5	-0.02	0.05	0.95	-0.9	0.96	-0.6
USR206	3.4	0.30	0.05	0.86	-3.3	0.86	-3.4
USR207	3.6	-0.34	0.05	1.03	0.5	1.03	0.5
USR208	3.7	-0.25	0.06	0.92	-1.6	0.92	-1.5
USR209	3.4	0.56	0.05	0.99	-0.2	0.98	-0.3
USR210	3.7	-0.11	0.05	1.23	4.7	1.23	4.6
USR211	3.2	0.93	0.33	0.64	-1.2	0.61	-1.3

Separation: 4.44 Reliability: 0.95

Fixed (all same) chi-square: 869.6 significance *p*: .00

Calibration of the HK raters for all student essays using US and HK raters and Lexile Analyzer

HK rater	Observed score	Measure (Logit)	S.E.	Infit		Outfit	
				MnSq	ZStd	MnSq	ZStd
HKR001	3.1	0.20	0.09	1.45	4.5	1.49	4.8
HKR002	3.5	-0.52	0.09	1.32	3.4	1.32	3.2
HKR003	3.8	-1.54	0.09	1.78	7.4	1.81	6.9
HKR004	3.0	0.55	0.09	1.55	5.5	1.52	5.2

Separation: 2.05 Reliability: 0.81

Fixed (all same) chi-square: 382.4 significance *p*: .00

Calibration of the Lexile Analyzer for all student essays using US and HK raters and Lexile Analyzer

Lexile Analyzer	Observed score	Measure (Logit)	S.E.	Infit		Outfit	
				MnSq	ZStd	MnSq	ZStd
Category 1	2.3	2.46	0.05	1.10	2.3	1.11	2.3
Category 2	2.9	1.85	0.04	1.15	4.2	1.16	4.2
Category 3	3.2	1.17	0.03	1.01	0.2	1.01	0.4
Category 4	3.5	0.65	0.03	0.92	-3.0	0.92	-2.9
Category 5	3.6	0.01	0.03	1.04	1.6	1.04	1.4
Category 6	3.9	-0.58	0.03	0.90	-3.5	0.90	-3.5
Category 7	4.1	-1.22	0.04	0.95	-1.4	0.95	-1.5
Category 8	4.5	-1.74	0.04	0.97	-0.6	0.99	-0.2
Category 9	4.9	-2.62	0.06	1.12	2.3	1.18	2.9

Separation: 40.50 Reliability: 1.00

Fixed (all same) chi-square: 11779.2 significance *p*: .00

Appendix 20

```
+--------------------------------------------------------------------------------------------------------------------+
|Measr|-Prompt                    |-US Rater                                      |-HK Rater |-IEA|+Student    |R6SCA|
|-----+---------------------------+-----------------------------------------------+----------+----+------------+-----|
|   5 +                           +                                               +          +    +            + (6) |
|     |                           |                                               |          |    | .          |     |
|     |                           |                                               |          |    |            |     |
|   4 +                           +                                               +          +    + .          +     |
|     |                           |                                               |          |    |            |  5  |
|     |                           |                                               |          | 1  | .          |     |
|   3 +                           +                                               +          +    + .          +     |
|     |                           |                                               |          |    | .          | --- |
|     |                           |                                               |          | 2  | **.        |     |
|   2 + 13-12AI  16-12BI  15-12AP +                                               +          +    + **.        +     |
|     | 9-8AP                     |                                               |          | 3  | *****.     |     |
|     |                           |                                               |          |    | *******    |  4  |
|   1 + 17-12BN  18-12BP  12-8BP  + US104  US200                                  +          + 4  + *******.   +     |
|     | 14-12AN  11-8BN           | US103  US209  US211                           | HK004    |    | *********. |     |
|     | 10-8BI                    | US102  US105  US206                           | HK001    |    | *********. |     |
*   0 * 7-8AI    8-8AN            * US101  US203  US204  US205                    *          * 5  * *********. * --- *
|     |                           | US106  US201  US202  US207  US208  US210      |          |    | *********. |     |
|     |                           |                                               | HK002    |    | ******.    |     |
|  -1 +                           +                                               +          + 6  + *****.     +     |
|     |                           |                                               |          | 7  | ***.       |  3  |
|     | 4-4BI    5-4BN    6-4BP   | US108                                         | HK003    |    | ***        |     |
|  -2 +                           +                                               +          +    + *.         +     |
|     | 1-4AI    2-4AN    3-4AP   |                                               |          | 8  | *.         |     |
|     |                           |                                               |          |    | *.         | --- |
|  -3 +                           +                                               +          +    + .          +     |
|     |                           |                                               |          | 9  | *.         |     |
|     |                           |                                               |          |    | *.         |  2  |
|  -4 +                           +                                               +          +    + .          +     |
|     |                           |                                               |          |    | .          |     |
|     |                           |                                               |          |    | .          |     |
|  -5 +                           +                                               +          +    + *.         + --- |
|     |                           |                                               |          |    | .          |     |
|     |                           |                                               |          |    | .          |     |
|  -6 +                           +                                               +          +    + .          +     |
|     |                           |                                               |          |    | .          |     |
|     |                           |                                               |          |    | .          |     |
|  -7 +                           +                                               +          +    +            +     |
|     |                           |                                               |          |    | .          |     |
|     |                           |                                               |          |    |            |     |
|  -8 +                           +                                               +          +    + .          +     |
|     |                           |                                               |          |    |            |     |
|     |                           |                                               |          |    | .          |     |
|  -9 +                           +                                               +          +    + .          + (1) |
|-----+---------------------------+-----------------------------------------------+----------+----+------------+-----|
|Measr|-Prompt                    |-US Rater                                      |-HK Rater |-IEA| * = 8      |R6SCA|
+--------------------------------------------------------------------------------------------------------------------+
```

Vertical map for all student essays with scoring of US and HK raters and IEA

Calibration of the students for all student essays using US and HK raters and IEA

Student			Observed	Measure	S.E.	Infit		Outfit	
Number		Group	score	(Logit)		MnSq	ZStd	MnSq	ZStd
1	US001	104	3.7	-0.14	0.33	0.90	-0.2	0.90	-0.2
2	US002	104	3.8	-0.02	0.36	1.29	0.9	1.30	0.9
3	US003	104	4.1	0.49	0.32	0.91	-0.2	0.91	-0.2
4	US004	104	4.2	0.09	0.32	0.47	-2.2	0.48	-2.1
5	US005	104	3.1	-2.29	0.32	1.48	1.5	1.50	1.6
6	US006	104	4.1	0.12	0.29	0.52	-2.0	0.51	-2.1
7	US006	304	3.9	0.06	0.27	0.60	-1.8	0.58	-1.9
8	US007	104	3.3	-1.45	0.33	1.03	0.1	1.03	0.1
9	US008	104	3.2	-0.89	0.33	0.43	-2.3	0.43	-2.3

10	US009	104	4.1	0.68	0.32	1.17	0.6	1.17	0.6
11	US010	104	3.5	-0.74	0.33	0.75	-0.8	0.75	-0.8
12	US011	104	5.3	1.54	0.32	0.61	-1.7	0.60	-1.7
13	US012	104	1.7	-3.59	0.43	0.80	-0.5	0.93	-0.1
14	US013	104	4.3	0.85	0.29	1.08	0.3	1.10	0.4
15	US013	304	4.3	0.63	0.26	1.15	0.7	1.20	0.8
16	US014	104	4.6	0.65	0.31	0.59	-1.6	0.57	-1.7
17	US015	104	4.1	0.05	0.36	0.34	-2.5	0.34	-2.5
18	US016	104	3.4	-0.93	0.33	0.68	-1.1	0.68	-1.1
19	US017	104	1.6	-3.46	0.34	1.89	2.9	1.85	2.7
20	US017	304	1.5	-3.63	0.30	1.94	3.4	1.84	3.0
21	US018	104	3.8	-0.67	0.33	1.29	1.0	1.29	1.0
22	US019	104	3.7	-0.77	0.33	0.51	-1.9	0.51	-1.9
23	US020	104	4.2	-0.56	0.32	0.84	-0.4	0.85	-0.4
24	US021	104	3.3	-1.49	0.29	2.16	3.4	2.16	3.4
25	US021	304	3.1	-1.65	0.27	2.75	5.0	2.76	5.1
26	US022	106	2.7	-0.84	0.29	0.58	-1.9	0.59	-1.8
27	US022	306	3.2	-0.94	0.27	0.63	-1.6	0.63	-1.6
28	US023	106	4.0	0.61	0.32	0.74	-0.9	0.73	-0.9
29	US024	106	3.1	-0.11	0.32	0.62	-1.4	0.64	-1.3
30	US025	106	2.8	-0.67	0.33	1.21	0.7	1.20	0.7
31	US026	106	3.5	-0.05	0.29	0.76	-0.9	0.76	-0.9
32	US026	306	3.9	-0.06	0.27	0.91	-0.2	0.91	-0.3
33	US027	106	4.3	1.53	0.29	1.34	1.2	1.34	1.2
34	US027	306	4.8	1.49	0.26	1.30	1.3	1.29	1.2
35	US028	106	3.3	0.13	0.33	0.60	-1.4	0.60	-1.4
36	US029	106	3.3	-0.21	0.33	0.86	-0.4	0.86	-0.4
37	US030	106	3.3	-0.33	0.29	1.95	2.9	1.96	2.9
38	US030	306	3.5	-0.69	0.27	1.97	3.2	1.96	3.2
39	US031	106	3.5	-0.75	0.32	1.63	1.9	1.63	1.9
40	US032	106	3.1	-1.29	0.32	0.65	-1.3	0.64	-1.3
41	US033	106	3.1	-0.06	0.32	0.46	-2.3	0.47	-2.3
42	US034	106	3.1	0.52	0.29	1.40	1.4	1.41	1.4
43	US034	306	3.3	-0.11	0.27	1.83	2.8	1.83	2.8
44	US035	106	4.3	1.29	0.31	1.30	1.0	1.29	1.0
45	US036	106	3.7	0.10	0.32	0.77	-0.8	0.75	-0.8
46	US037	106	3.8	0.35	0.32	0.94	-0.1	0.94	-0.1
47	US038	106	3.3	0.40	0.33	0.51	-1.9	0.51	-1.9

48	US039	106	2.5	-0.82	0.32	0.82	-0.5	0.85	-0.4
49	US040	106	3.8	0.09	0.32	0.89	-0.3	0.88	-0.3
50	US041	104	3.8	-0.29	0.32	0.66	-1.2	0.67	-1.1
51	US042	104	2.6	-1.62	0.32	0.42	-2.5	0.42	-2.5
52	US043	104	3.6	-0.27	0.32	0.82	-0.5	0.83	-0.5
53	US044	104	3.8	0.22	0.32	0.81	-0.6	0.82	-0.5
54	US045	104	4.0	0.16	0.32	0.47	-2.1	0.47	-2.1
55	US046	104	2.8	-1.38	0.29	1.13	0.5	1.14	0.5
56	US046	304	2.7	-1.67	0.26	1.47	1.8	1.50	1.9
57	US047	104	3.6	-0.14	0.32	0.65	-1.2	0.64	-1.2
58	US048	104	3.4	-0.60	0.40	1.02	0.1	1.01	0.1
59	US049	104	2.7	-1.81	0.29	2.81	4.8	2.85	4.9
60	US049	304	2.7	-1.85	0.26	3.13	6	3.16	6.0
61	US050	104	3.0	-0.85	0.36	1.17	0.5	1.17	0.6
62	US051	104	4.5	0.68	0.31	0.96	0	0.98	0
63	US052	104	3.2	-0.93	0.33	1.73	2.1	1.74	2.2
64	US053	104	3.8	0.14	0.36	0.65	-1.1	0.64	-1.1
65	US054	104	3.6	-0.41	0.32	1.12	0.4	1.14	0.5
66	US055	104	3.0	-0.80	0.32	0.78	-0.7	0.78	-0.7
67	US056	104	2.9	-1.66	0.32	1.06	0.3	1.07	0.3
68	US057	104	3.7	-0.84	0.35	2.14	2.7	2.18	2.8
69	US057	304	3.5	-0.36	0.33	2.41	3.5	2.40	3.5
70	US058	104	2.5	-2.00	0.31	1.72	2.2	1.73	2.2
71	US059	104	3.3	-1.27	0.36	0.46	-2.0	0.46	-2.0
72	US060	104	2.3	-2.48	0.28	1.39	1.5	1.39	1.5
73	US060	304	2.0	-3.22	0.26	1.67	2.6	1.61	2.4
74	US061	104	2.4	-2.51	0.38	0.80	-0.5	0.80	-0.5
75	US062	106	3.7	-0.30	0.33	1.25	0.8	1.25	0.8
76	US063	106	3.9	0.55	0.40	0.58	-1.2	0.56	-1.3
77	US064	106	3.8	-0.56	0.29	1.02	0.1	0.98	0
78	US064	306	4.2	-0.25	0.26	1.42	1.5	1.37	1.4
79	US065	106	3.7	0.15	0.32	1.03	0.1	1.05	0.2
80	US066	106	4.5	1.38	0.32	0.90	-0.2	0.88	-0.3
81	US067	106	3.9	0.80	0.32	0.68	-1.1	0.69	-1.1
82	US068	106	3.8	0.38	0.33	1.12	0.4	1.11	0.4
83	US069	106	3.5	-0.15	0.33	0.87	-0.3	0.87	-0.3
84	US070	106	2.3	-1.35	0.29	1.63	2.1	1.63	2.1
85	US070	306	2.4	-2.00	0.26	1.82	3.0	1.89	3.2

86	US071	106	3.8	-0.09	0.32	1.35	1.1	1.34	1.1
87	US072	106	3.0	-0.51	0.32	1.26	0.9	1.26	0.9
88	US073	106	2.7	-0.91	0.32	1.33	1.1	1.31	1.0
89	US074	106	3.6	-0.07	0.33	1.03	0.1	1.03	0.2
90	US075	106	4.3	1.22	0.32	1.16	0.6	1.16	0.6
91	US076	106	4.0	0.77	0.32	0.43	-2.4	0.42	-2.4
92	US077	106	2.5	-0.20	0.36	0.75	-0.7	0.74	-0.7
93	US078	106	4.2	0.01	0.32	0.93	-0.1	0.93	-0.1
94	US079	106	4.1	-0.17	0.32	0.46	-2.2	0.45	-2.3
95	US080	106	2.5	-1.28	0.32	0.56	-1.7	0.57	-1.7
96	US081	106	3.5	0.06	0.33	0.88	-0.3	0.88	-0.3
97	US082	108	3.4	0.55	0.33	0.47	-2.1	0.47	-2.1
98	US083	108	3.1	-0.69	0.33	0.62	-1.4	0.62	-1.4
99	US084	108	1.5	-1.73	0.41	0.82	-0.5	0.87	-0.3
100	US085	108	3.0	-0.18	0.32	0.79	-0.6	0.80	-0.6
101	US086	108	2.8	-0.77	0.32	0.75	-0.8	0.76	-0.8
102	US087	108	3.0	-0.11	0.32	1.11	0.4	1.12	0.4
103	US088	108	3.0	-0.07	0.32	0.54	-1.8	0.54	-1.8
104	US089	108	3.0	0.36	0.32	0.52	-1.9	0.52	-1.9
105	US090	108	3.3	-0.18	0.33	0.51	-2.0	0.51	-2.0
106	US091	108	2.2	-0.49	0.32	0.79	-0.7	0.78	-0.7
107	US092	108	3.5	-0.21	0.32	0.85	-0.4	0.86	-0.4
108	US093	108	2.8	-0.34	0.32	1.29	1.0	1.29	1.0
109	US094	108	3.7	0.89	0.33	0.53	-1.7	0.53	-1.7
110	US095	108	3.3	0.52	0.33	1.20	0.7	1.17	0.6
111	US096	108	3.1	-0.34	0.32	0.97	0	0.96	0
112	US097	108	3.6	-0.03	0.32	0.65	-1.2	0.66	-1.2
113	US098	108	2.8	-0.37	0.32	0.60	-1.5	0.60	-1.6
114	US099	108	2.8	-0.90	0.32	0.83	-0.5	0.84	-0.4
115	US100	108	3.0	-0.27	0.32	0.65	-1.3	0.65	-1.2
116	US101	108	2.9	-0.56	0.32	1.77	2.2	1.77	2.2
117	US102	108	3.0	0.25	0.32	0.82	-0.5	0.81	-0.6
118	US103	108	2.6	-0.30	0.32	0.77	-0.7	0.78	-0.7
119	US104	108	3.6	0.40	0.33	0.77	-0.7	0.77	-0.7
120	US105	108	2.9	0.14	0.32	0.45	-2.3	0.45	-2.3
121	US106	108	3.6	0.94	0.33	0.59	-1.5	0.59	-1.5
122	US107	108	3.3	0.17	0.32	0.72	-0.9	0.72	-0.9
123	US108	108	3.7	0.81	0.33	0.67	-1.1	0.67	-1.1

124	US109	108	3.2	-0.13	0.33	0.68	-1.1	0.68	-1.1
125	US110	108	3.2	-0.21	0.36	1.10	0.4	1.11	0.4
126	US111	108	3.4	0.44	0.32	0.47	-2.1	0.47	-2.1
127	US112	108	3.3	0.43	0.33	1.03	0.1	1.02	0.1
128	US113	108	3.4	0.40	0.33	0.82	-0.5	0.83	-0.5
129	US114	108	3.1	0.32	0.32	0.75	-0.8	0.75	-0.8
130	US115	108	3.2	0.44	0.32	0.58	-1.6	0.58	-1.6
131	US116	108	3.1	-0.24	0.33	0.76	-0.8	0.76	-0.8
132	US117	108	2.3	-0.49	0.32	0.39	-2.8	0.40	-2.8
133	US118	108	2.3	-2.15	0.32	0.57	-1.7	0.58	-1.7
134	US119	108	3.0	0.15	0.32	0.45	-2.3	0.45	-2.3
135	US120	108	2.8	0.19	0.32	0.55	-1.7	0.55	-1.7
136	US121	108	2.1	-0.69	0.32	0.90	-0.3	0.89	-0.3
137	US122	108	2.6	0.31	0.32	0.30	-3.4	0.29	-3.4
138	US123	108	2.5	-0.74	0.35	0.29	-3.2	0.29	-3.2
139	US124	108	3.2	0.82	0.32	0.35	-2.8	0.35	-2.8
140	US125	108	2.3	-0.41	0.32	0.72	-1.0	0.73	-0.9
141	US126	108	2.5	-0.45	0.32	0.59	-1.6	0.59	-1.6
142	US127	108	2.7	0.51	0.32	0.59	-1.6	0.58	-1.6
143	US128	108	3.1	-0.26	0.32	0.33	-3.0	0.33	-3.1
144	US129	108	3.5	0.52	0.40	0.69	-0.8	0.69	-0.8
145	US130	108	2.4	-0.38	0.35	0.20	-4.0	0.20	-4.0
146	US131	108	4.6	1.46	0.40	0.45	-1.9	0.47	-1.8
147	US132	108	4.1	1.65	0.32	0.96	0	0.96	0
148	US133	108	4.8	1.80	0.31	1.12	0.5	1.11	0.4
149	US134	108	4.0	1.18	0.33	0.68	-1.1	0.68	-1.1
150	US135	108	5.3	3.18	0.33	1.53	1.8	1.59	2.0
151	US136	108	5.5	2.29	0.37	1.08	0.3	0.95	0
152	US137	108	4.9	2.03	0.32	1.32	1.1	1.31	1.1
153	US138	108	3.8	0.73	0.33	0.53	-1.8	0.53	-1.8
154	US139	108	4.5	1.50	0.32	1.16	0.6	1.13	0.5
155	US140	108	4.3	1.48	0.31	0.53	-1.9	0.52	-1.9
156	US141	108	4.1	1.46	0.32	0.51	-1.9	0.50	-2.0
157	US142	108	4.5	1.65	0.31	0.47	-2.3	0.47	-2.3
158	US143	108	4.6	1.51	0.32	0.73	-0.9	0.73	-0.9
159	US144	108	5.4	3.10	0.35	0.75	-0.9	0.82	-0.6
160	US145	108	4.6	1.60	0.32	0.79	-0.7	0.90	-0.2
161	US146	108	4.5	1.37	0.31	0.58	-1.7	0.59	-1.6

162	US147	108	4.4	1.68	0.35	1.23	0.7	1.23	0.7
163	US148	108	5.0	2.05	0.31	0.68	-1.3	0.67	-1.3
164	US149	108	5.5	3.13	0.42	1.14	0.5	1.04	0.2
165	US150	108	5.4	2.69	0.34	0.45	-2.5	0.49	-2.2
166	US151	108	4.6	1.36	0.32	0.99	0	0.98	0
167	US152	108	5.0	1.79	0.35	1.01	0.1	1.02	0.1
168	US153	108	2.3	-0.35	0.42	0.46	-1.8	0.51	-1.6
169	US154	108	3.3	0.03	0.36	0.94	0	0.94	0
170	US155	108	3.0	0.28	0.32	0.71	-0.9	0.72	-0.9
171	US156	108	3.3	-0.03	0.32	0.65	-1.2	0.63	-1.3
172	US157	108	3.4	0.55	0.33	0.63	-1.3	0.63	-1.3
173	US158	108	3.5	0.30	0.36	0.36	-2.5	0.37	-2.5
174	US159	108	3.4	-0.15	0.33	0.71	-1.0	0.70	-1.0
175	US160	108	2.4	-0.56	0.31	0.30	-3.5	0.29	-3.5
176	US161	108	3.4	0.32	0.36	1.56	1.5	1.58	1.6
177	US162	108	4.0	0.85	0.33	0.95	0	0.95	0
178	US163	108	4.1	1.37	0.36	0.71	-0.8	0.72	-0.8
179	US164	108	2.1	-0.59	0.32	0.61	-1.5	0.62	-1.5
180	US165	108	2.5	-0.48	0.33	0.78	-0.7	0.77	-0.7
181	US166	108	1.8	-0.84	0.40	0.66	-1.0	0.69	-1.0
182	US167	108	3.1	0.35	0.32	0.51	-2.0	0.52	-1.9
183	US168	108	3.5	-0.53	0.33	0.68	-1.1	0.68	-1.1
184	US169	108	2.4	-0.44	0.35	0.87	-0.3	0.87	-0.3
185	US170	108	2.5	-0.81	0.32	0.52	-2.0	0.51	-2.0
186	US171	108	3.1	-0.52	0.32	0.82	-0.5	0.82	-0.5
187	US172	108	3.2	0.61	0.33	0.76	-0.7	0.77	-0.7
188	US173	108	3.5	0.98	0.33	0.76	-0.8	0.76	-0.8
189	US174	108	3.2	0.57	0.32	0.53	-1.9	0.53	-1.9
190	US175	108	4.0	1.18	0.36	0.50	-1.7	0.49	-1.8
191	US176	108	4.1	0.83	0.33	0.47	-2.1	0.47	-2.0
192	US177	108	3.8	1.06	0.33	0.85	-0.4	0.85	-0.4
193	US178	108	4.2	1.61	0.32	0.79	-0.6	0.78	-0.6
194	US179	108	4.0	2.00	0.35	0.80	-0.5	0.79	-0.5
195	US180	108	3.5	0.93	0.33	1.26	0.8	1.28	0.9
196	US181	108	3.3	0.46	0.33	0.87	-0.3	0.87	-0.3
197	US182	108	3.3	0.63	0.36	1.41	1.2	1.39	1.1
198	US183	108	5.0	2.23	0.32	1.17	0.6	1.11	0.5
199	US184	108	3.3	0.23	0.33	0.90	-0.2	0.89	-0.2

200	US185	108	3.6	0.58	0.36	0.81	-0.5	0.81	-0.5
201	US186	108	4.2	1.51	0.32	1.28	0.9	1.32	1.0
202	US187	108	3.4	0.25	0.32	0.59	-1.5	0.59	-1.5
203	US188	108	3.3	0.89	0.33	1.73	2.1	1.74	2.2
204	US189	108	3.9	0.88	0.32	1.23	0.8	1.21	0.7
205	US190	108	3.9	1.21	0.36	0.41	-2.2	0.41	-2.2
206	US191	108	3.6	0.33	0.33	0.59	-1.5	0.59	-1.5
207	US192	108	2.8	-0.09	0.32	0.63	-1.3	0.63	-1.4
208	US193	108	3.0	0.38	0.32	0.63	-1.3	0.63	-1.3
209	US194	108	3.6	0.72	0.33	0.46	-2.1	0.46	-2.1
210	US195	108	3.6	-0.33	0.36	1.77	2.1	1.77	2.0
211	US196	108	4.2	1.12	0.32	1.02	0.1	1.03	0.1
212	US197	108	3.6	-0.40	0.33	0.90	-0.2	0.91	-0.2
213	US198	108	3.7	0.46	0.33	0.37	-2.7	0.36	-2.7
214	US199	108	3.5	0.36	0.33	0.75	-0.8	0.75	-0.8
215	US200	108	3.1	-0.38	0.33	0.38	-2.7	0.38	-2.6
216	US201	108	3.9	1.03	0.36	0.93	-0.1	0.92	-0.1
217	US202	108	3.4	0.03	0.33	0.61	-1.4	0.61	-1.4
218	US203	108	3.9	0.67	0.32	0.83	-0.5	0.84	-0.4
219	US204	108	3.4	0.28	0.32	0.53	-1.9	0.53	-1.9
220	US205	108	4.0	0.42	0.32	0.62	-1.4	0.61	-1.4
221	US206	108	4.3	1.36	0.32	1.06	0.3	1.06	0.2
222	US207	108	4.1	1.80	0.33	1.03	0.1	1.02	0.1
223	US208	108	4.8	1.69	0.31	1.11	0.4	1.10	0.4
224	US209	108	4.0	0.13	0.33	1.00	0.1	1.01	0.1
225	US210	108	3.8	0.70	0.36	0.87	-0.3	0.87	-0.2
226	US211	104	4.0	-0.17	0.33	0.98	0	0.98	0
227	US212	104	3.3	-1.01	0.29	1.49	1.7	1.50	1.7
228	US212	304	3.0	-1.54	0.26	1.83	2.8	1.84	2.8
229	US213	104	4.1	0.46	0.40	0.53	-1.4	0.55	-1.3
230	US214	104	2.9	-2.16	0.32	2.29	3.5	2.32	3.5
231	US215	104	3.7	-0.12	0.29	0.87	-0.4	0.86	-0.4
232	US215	304	3.5	-0.77	0.27	0.95	-0.1	0.95	-0.1
233	US216	104	3.7	0.07	0.33	0.56	-1.6	0.56	-1.6
234	US217	104	3.5	-0.85	0.33	0.96	0	0.96	0
235	US218	104	4.2	0.34	0.29	1.30	1.0	1.28	1.0
236	US218	304	4.2	0.27	0.26	1.57	2.0	1.57	2.0
237	US219	104	4.3	0.61	0.32	0.40	-2.4	0.40	-2.5

238	US220	104	3.6	-0.42	0.33	1.00	0	1.01	0.1
239	US221	104	2.5	-1.96	0.31	1.15	0.6	1.15	0.6
240	US222	104	4.1	-0.09	0.32	0.59	-1.5	0.58	-1.5
241	US223	104	3.8	-0.23	0.33	0.80	-0.6	0.79	-0.6
242	US224	104	2.4	-1.60	0.39	1.88	2.1	1.88	2.1
243	US225	104	3.3	-0.96	0.33	0.84	-0.4	0.84	-0.4
244	US226	104	4.3	0.17	0.32	0.83	-0.5	0.86	-0.4
245	US227	104	3.1	-1.44	0.32	0.97	0	0.96	0
246	US227	304	2.6	-1.71	0.30	1.07	0.3	1.06	0.3
247	US228	104	3.0	-1.24	0.32	0.73	-0.9	0.73	-0.9
248	US229	104	3.5	-0.64	0.40	1.01	0.1	1.02	0.1
249	US230	106	3.0	-0.51	0.32	0.96	0	0.97	0
250	US231	106	3.1	-0.54	0.32	0.60	-1.5	0.59	-1.5
251	US232	106	3.3	0.32	0.33	0.52	-1.8	0.52	-1.8
252	US233	106	2.7	-0.59	0.32	0.64	-1.3	0.64	-1.3
253	US234	106	3.0	-1.30	0.32	0.41	-2.5	0.41	-2.5
254	US235	106	2.8	-0.45	0.29	1.29	1.1	1.28	1.0
255	US235	306	2.6	-1.00	0.26	1.57	2.1	1.55	2.1
256	US236	106	2.4	-0.78	0.29	0.85	-0.5	0.83	-0.6
257	US236	306	2.6	-1.05	0.26	1.01	0.1	1.00	0
258	US237	106	3.3	-0.94	0.29	1.93	2.8	1.97	3.0
259	US237	306	3.4	-1.51	0.27	2.43	4.4	2.44	4.4
260	US238	106	3.3	-0.26	0.32	0.79	-0.7	0.78	-0.7
261	US239	106	3.5	0.25	0.32	1.06	0.2	1.05	0.2
262	US240	106	3.1	0.06	0.32	0.41	-2.5	0.41	-2.5
263	US241	106	3.7	0.14	0.29	1.30	1.1	1.29	1.1
264	US241	306	4.3	0.80	0.26	0.75	-1.0	0.78	-0.9
265	US242	106	3.7	0.88	0.32	1.12	0.4	1.17	0.6
266	US243	106	4.0	0.92	0.29	1.41	1.4	1.42	1.4
267	US243	306	4.4	1.35	0.26	1.56	2.1	1.55	2.0
268	US244	106	3.8	0.51	0.32	0.69	-1.1	0.69	-1.0
269	US245	106	2.4	-0.93	0.33	0.49	-2.1	0.51	-2.1
270	US246	106	3.0	-1.19	0.29	0.84	-0.5	0.86	-0.4
271	US246	306	3.6	-1.13	0.27	1.23	0.9	1.23	0.9
272	US247	106	3.7	0.11	0.33	1.02	0.1	1.04	0.2
273	US248	106	4.0	0.99	0.28	2.12	3.4	2.07	3.3
274	US248	306	4.9	2.06	0.26	1.65	2.5	1.59	2.3
275	US249	106	3.7	0.26	0.32	0.98	0	0.97	0

276	US250	104	3.4	-0.49	0.33	0.42	-2.4	0.42	-2.4
277	US251	104	3.2	-1.66	0.33	0.83	-0.5	0.83	-0.5
278	US252	104	3.4	-0.95	0.33	1.16	0.6	1.16	0.6
279	US253	104	3.1	-1.10	0.32	0.77	-0.7	0.77	-0.7
280	US254	104	3.9	0.40	0.33	0.46	-2.2	0.45	-2.2
281	US255	104	2.7	-1.37	0.32	0.55	-1.8	0.54	-1.8
282	US256	104	1.8	-2.98	0.33	1.28	1.0	1.26	0.9
283	US256	304	1.8	-2.85	0.34	1.77	2.4	1.71	2.2
284	US257	104	3.4	-2.30	0.33	0.88	-0.3	0.87	-0.3
285	US258	104	3.5	-0.15	0.33	1.25	0.8	1.25	0.8
286	US259	104	4.4	0.84	0.32	0.47	-2.3	0.45	-2.3
287	US260	104	3.0	-1.72	0.32	1.47	1.5	1.47	1.5
288	US261	104	4.1	0.22	0.32	0.76	-0.7	0.76	-0.7
289	US262	104	2.2	-2.64	0.35	0.30	-3.2	0.30	-3.2
290	US263	104	2.1	-2.46	0.32	0.93	-0.1	0.94	-0.1
291	US264	104	2.9	-1.34	0.32	0.55	-1.8	0.56	-1.7
292	US265	104	3.9	-0.24	0.32	1.22	0.7	1.20	0.7
293	US266	104	2.4	-1.83	0.28	1.53	1.9	1.52	1.8
294	US266	304	2.1	-2.26	0.26	2.05	3.8	2.02	3.7
295	US267	104	2.5	-1.48	0.32	1.12	0.4	1.10	0.4
296	US268	104	2.2	-2.79	0.32	0.84	-0.5	0.84	-0.5
297	US269	104	5.0	1.30	0.31	1.67	2.2	1.66	2.2
298	US270	104	3.9	0.04	0.36	1.51	1.4	1.51	1.4
299	US271	104	3.2	-1.36	0.36	4.22	5.8	4.27	5.9
300	US272	104	4.5	-0.76	0.32	1.42	1.4	1.44	1.4
301	US273	104	4.1	0.59	0.29	0.67	-1.3	0.65	-1.4
302	US273	304	4.1	0.75	0.26	0.82	-0.7	0.81	-0.7
303	US274	104	3.6	-0.22	0.33	0.64	-1.2	0.65	-1.2
304	US275	104	4.7	1.10	0.31	0.46	-2.4	0.46	-2.4
305	US276	104	3.9	0.22	0.33	1.00	0	0.97	0
306	US277	104	4.4	0.68	0.31	0.68	-1.1	0.71	-1.0
307	US278	104	4.8	1.28	0.31	1.47	1.6	1.45	1.5
308	US279	104	3.5	0.54	0.33	0.60	-1.4	0.60	-1.4
309	US280	104	4.4	-0.22	0.39	1.42	1.1	1.40	1.1
310	US281	104	4.2	-0.63	0.32	2.27	3.2	2.33	3.3
311	US282	104	4.1	0.29	0.33	0.43	-2.3	0.43	-2.3
312	US283	104	3.8	-0.76	0.33	1.34	1.1	1.34	1.1
313	US284	104	4.1	-0.82	0.32	1.16	0.6	1.16	0.5

314	US285	104	1.2	-5.69	0.57	1.52	1.1	2.18	1.6
315	US285	304	1.1	-5.63	0.76	1.17	0.4	1.56	0.8
316	US286	104	4.3	-0.07	0.32	0.93	-0.1	0.92	-0.2
317	US287	104	4.3	0.60	0.32	0.76	-0.8	0.76	-0.7
318	US288	104	4.5	0.74	0.31	1.57	1.8	1.58	1.8
319	US289	104	3.0	-0.64	0.32	0.58	-1.6	0.58	-1.5
320	US290	104	3.4	-1.13	0.40	0.42	-2.0	0.42	-2.0
321	US291	106	3.3	0.25	0.33	1.27	0.9	1.26	0.9
322	US292	106	5.2	2.27	0.34	1.58	1.8	1.32	0.9
323	US293	106	4.3	1.26	0.31	0.70	-1.1	0.69	-1.1
324	US294	106	4.4	0.80	0.34	1.32	1.0	1.30	1.0
325	US295	106	3.1	-0.13	0.32	0.57	-1.6	0.57	-1.6
326	US296	106	4.2	0.72	0.32	1.35	1.2	1.35	1.2
327	US297	106	3.9	0.21	0.33	1.11	0.4	1.12	0.4
328	US298	106	4.7	0.95	0.32	1.10	0.4	1.09	0.4
329	US299	106	3.8	0.43	0.35	2.39	3.3	2.50	3.4
330	US300	106	4.4	1.72	0.32	0.82	-0.5	0.82	-0.6
331	US301	106	3.5	-0.55	0.33	1.08	0.3	1.08	0.3
332	US302	106	5.2	1.74	0.36	0.79	-0.6	0.72	-0.8
333	US303	106	4.1	-0.16	0.32	0.70	-1.0	0.69	-1.1
334	US304	106	5.0	1.97	0.35	0.87	-0.3	0.87	-0.3
335	US305	106	4.5	0.81	0.35	0.72	-0.8	0.75	-0.7
336	US306	106	3.0	-1.55	0.39	1.95	2.2	2.01	2.3
337	US307	110	4.8	2.29	0.31	0.50	-2.3	0.49	-2.3
338	US308	110	3.9	0.66	0.32	0.52	-1.9	0.51	-1.9
339	US309	110	4.1	1.76	0.32	0.50	-2.0	0.49	-2.0
340	US310	110	3.8	1.21	0.33	0.80	-0.6	0.79	-0.6
341	US311	110	4.9	1.80	0.32	0.99	0	1.00	0.1
342	US312	110	5.0	2.78	0.31	1.07	0.3	1.07	0.3
343	US313	110	4.7	2.29	0.31	0.68	-1.2	0.66	-1.3
344	US314	110	4.5	2.00	0.31	0.66	-1.3	0.64	-1.3
345	US315	110	5.0	2.28	0.32	0.99	0	0.97	0
346	US316	110	4.0	1.34	0.32	0.88	-0.3	0.86	-0.3
347	US317	110	4.4	1.51	0.32	0.99	0	0.97	0
348	US318	110	4.1	1.17	0.32	1.37	1.2	1.37	1.2
349	US319	110	4.8	2.01	0.31	0.85	-0.5	0.83	-0.5
350	US320	110	5.3	2.41	0.33	0.95	-0.1	0.97	0
351	US321	110	3.1	0.50	0.32	0.56	-1.7	0.56	-1.7

352	US322	110	3.0	-0.05	0.32	1.13	0.5	1.11	0.4
353	US323	110	2.8	0.24	0.33	0.96	0	0.98	0
354	US324	110	4.3	1.19	0.32	0.60	-1.5	0.61	-1.4
355	US325	110	3.3	0.29	0.33	0.45	-2.3	0.45	-2.3
356	US326	110	3.1	-0.26	0.35	1.14	0.5	1.15	0.5
357	US327	110	3.1	-0.50	0.32	1.18	0.6	1.19	0.7
358	US328	110	3.9	1.11	0.33	0.59	-1.5	0.59	-1.5
359	US329	110	3.0	-0.34	0.32	0.80	-0.6	0.80	-0.6
360	US330	110	3.2	0.73	0.36	0.05	-5.5	0.05	-5.5
361	US331	110	4.0	0.73	0.32	1.11	0.4	1.09	0.3
362	US332	110	2.6	-0.92	0.39	0.76	-0.6	0.76	-0.6
363	US333	110	3.3	0.78	0.33	0.67	-1.1	0.68	-1.1
364	US334	110	4.0	0.62	0.32	1.00	0	0.99	0
365	US335	110	4.9	1.54	0.31	1.30	1.1	1.29	1.1
366	US336	110	4.3	1.30	0.32	0.61	-1.4	0.61	-1.4
367	US337	110	3.4	0.76	0.33	0.64	-1.2	0.64	-1.2
368	US338	110	3.5	0.28	0.33	0.62	-1.4	0.62	-1.4
369	US339	110	4.5	1.42	0.31	0.91	-0.2	0.92	-0.1
370	US340	110	3.5	0.25	0.32	0.99	0	0.99	0
371	US341	110	3.3	0.04	0.33	1.16	0.6	1.16	0.6
372	US342	110	4.8	2.48	0.33	0.72	-1.0	0.72	-1.0
373	US343	110	4.5	1.39	0.31	1.49	1.5	1.47	1.5
374	US344	110	4.8	2.09	0.31	1.28	1.0	1.33	1.2
375	US345	110	4.7	1.66	0.31	0.61	-1.6	0.63	-1.5
376	US346	110	4.1	0.87	0.32	0.93	-0.1	0.92	-0.1
377	US347	110	4.1	1.44	0.32	0.65	-1.2	0.66	-1.2
378	US348	110	4.1	1.50	0.32	0.64	-1.2	0.64	-1.2
379	US349	110	3.8	0.79	0.33	1.29	1.0	1.28	0.9
380	US350	110	4.0	1.27	0.32	0.77	-0.7	0.76	-0.7
381	US351	110	4.6	2.23	0.31	1.14	0.5	1.16	0.6
382	US352	110	3.4	-0.13	0.33	1.18	0.6	1.18	0.6
383	US353	110	4.8	2.27	0.31	1.22	0.8	1.23	0.8
384	US354	110	4.1	1.09	0.33	1.06	0.2	1.06	0.3
385	US355	110	3.9	0.91	0.33	0.99	0	0.99	0
386	US356	110	4.5	1.80	0.31	1.30	1.0	1.30	1.0
387	US357	110	4.1	1.08	0.32	1.65	1.9	1.67	2.0
388	US358	110	3.9	0.88	0.32	0.73	-0.8	0.74	-0.8
389	US359	110	4.1	0.60	0.32	1.12	0.4	1.13	0.5

390	US360	110	3.4	-0.38	0.33	1.49	1.5	1.51	1.5
391	US361	110	3.9	0.91	0.33	1.07	0.3	1.09	0.3
392	US362	110	3.3	-0.64	0.33	1.98	2.7	1.96	2.6
393	US363	110	3.6	1.03	0.40	0.73	-0.7	0.74	-0.6
394	US364	110	3	-0.15	0.32	0.74	-0.9	0.75	-0.9
395	US365	110	2.2	0.07	0.41	0.87	-0.3	0.84	-0.3
396	US366	110	3.9	1.14	0.32	2.50	3.6	2.47	3.5
397	US367	110	3.5	0.50	0.33	0.62	-1.4	0.62	-1.4
398	US368	110	4.3	1.50	0.32	0.84	-0.4	0.84	-0.4
399	US369	110	3.1	0.14	0.32	0.57	-1.6	0.58	-1.6
400	US370	110	3.6	0.53	0.33	1.26	0.9	1.25	0.8
401	US371	110	4.2	1.70	0.32	0.60	-1.5	0.58	-1.5
402	US372	110	3.2	0.76	0.32	0.88	-0.3	0.87	-0.3
403	US373	110	3.3	0.13	0.33	0.75	-0.8	0.75	-0.8
404	US374	110	2.8	0.13	0.32	0.46	-2.3	0.47	-2.3
405	US375	110	3.0	0.18	0.32	0.39	-2.6	0.39	-2.6
406	US376	110	2.6	-0.95	0.37	0.84	-0.4	0.78	-0.5
407	US377	110	3.3	0.39	0.33	0.56	-1.7	0.56	-1.6
408	US378	110	3.3	0.12	0.33	0.77	-0.7	0.78	-0.7
409	US379	110	3.3	0.47	0.33	0.65	-1.2	0.65	-1.2
410	US380	110	4.3	1.73	0.32	0.93	-0.1	0.90	-0.2
411	US381	110	3.3	-0.07	0.33	0.55	-1.7	0.55	-1.7
412	US382	110	3.1	-0.39	0.32	0.52	-1.9	0.52	-1.9
413	US383	110	3.4	0.56	0.36	1.23	0.7	1.23	0.7
414	US384	110	3.7	0.67	0.33	0.50	-2.0	0.50	-1.9
415	US385	110	4.4	1.83	0.32	0.89	-0.3	0.86	-0.4
416	US386	110	4.0	1.15	0.32	0.94	-0.1	0.96	0
417	US387	110	4.6	1.90	0.31	1.33	1.1	1.33	1.1
418	US388	110	3.8	1.13	0.33	1.03	0.2	1.03	0.2
419	US389	110	4.6	1.95	0.31	0.92	-0.2	0.92	-0.2
420	US390	110	4.1	0.73	0.32	1.11	0.4	1.11	0.4
421	US391	110	4.1	0.70	0.32	0.82	-0.5	0.84	-0.5
422	US392	110	4.2	1.45	0.32	1.15	0.5	1.12	0.4
423	US393	110	4.7	2.43	0.31	1.35	1.2	1.38	1.3
424	US394	110	3.6	0.55	0.33	1.03	0.2	1.03	0.1
425	US395	110	3.6	1.10	0.33	0.75	-0.8	0.77	-0.7
426	US396	110	5.3	2.81	0.34	0.78	-0.7	0.78	-0.7
427	US397	110	4.5	1.71	0.31	0.79	-0.7	0.78	-0.7

428	US398	110	4.0	0.20	0.33	0.58	-1.5	0.59	-1.5
429	US399	106	3.5	0.50	0.32	0.55	-1.7	0.55	-1.7
430	US400	106	3.7	0.83	0.32	0.78	-0.6	0.80	-0.6
431	US401	106	4.3	1.06	0.31	1.39	1.3	1.40	1.3
432	US402	106	3.8	1.02	0.32	0.81	-0.6	0.82	-0.5
433	US403	106	3.3	0.18	0.32	0.41	-2.6	0.41	-2.6
434	US404	106	2.3	-0.93	0.29	0.56	-1.9	0.55	-2.0
435	US404	306	2.7	-1.12	0.26	0.83	-0.7	0.83	-0.6
436	US405	106	3.7	0.60	0.32	0.96	0	0.98	0
437	US406	106	3.4	-0.42	0.32	0.83	-0.5	0.84	-0.4
438	US407	106	2.9	-0.49	0.32	0.85	-0.4	0.85	-0.4
439	US408	106	4.1	0.68	0.32	0.47	-2.2	0.47	-2.1
440	US409	106	4.5	1.24	0.32	0.64	-1.3	0.65	-1.3
441	US410	106	5.2	2.60	0.35	1.01	0.1	0.88	-0.2
442	US411	106	4.1	1.17	0.32	1.04	0.2	1.07	0.3
443	US412	106	3.4	-0.15	0.32	0.36	-2.8	0.36	-2.8
444	US413	106	2.8	-0.37	0.32	0.58	-1.6	0.58	-1.6
445	US414	106	2.9	-0.97	0.32	1.00	0.1	1.00	0
446	US415	106	3.5	-0.45	0.32	0.70	-1.0	0.70	-1.0
447	US416	106	4.2	1.13	0.32	0.55	-1.9	0.55	-1.8
448	US417	106	3.2	0.09	0.32	0.64	-1.3	0.63	-1.3
449	US418	104	3.1	-0.83	0.32	1.01	0.1	1.01	0.1
450	US419	104	2.9	-0.97	0.29	0.64	-1.5	0.64	-1.5
451	US419	304	2.8	-0.68	0.26	0.50	-2.5	0.51	-2.5
452	US420	104	2.2	-2.25	0.28	0.45	-2.8	0.45	-2.8
453	US420	304	2.2	-2.25	0.26	0.59	-2.0	0.60	-2.0
454	US421	104	3.6	-0.46	0.33	1.23	0.8	1.23	0.8
455	US422	104	3.8	-0.41	0.33	1.12	0.4	1.10	0.4
456	US423	104	3.1	0.20	0.33	1.02	0.1	1.00	0.1
457	US424	104	3.5	-0.02	0.33	0.43	-2.4	0.43	-2.4
458	US425	104	3.9	-0.06	0.33	0.37	-2.6	0.36	-2.6
459	US426	104	3.6	-0.06	0.33	0.93	-0.1	0.92	-0.1
460	US427	104	3.3	-0.83	0.30	1.26	0.9	1.24	0.9
461	US427	304	3.3	-0.59	0.27	1.65	2.3	1.63	2.2
462	US428	104	4.3	0.39	0.32	0.69	-1.0	0.67	-1.1
463	US429	104	2.8	-1.16	0.29	0.91	-0.2	0.91	-0.2
464	US429	304	2.6	-1.45	0.26	1.11	0.5	1.11	0.5
465	US430	104	3.0	-0.96	0.33	0.72	-0.9	0.72	-0.9

466	US431	104	2.6	-1.45	0.32	0.49	-2.1	0.49	-2.1
467	US432	104	2.2	-1.99	0.38	0.81	-0.5	0.81	-0.4
468	US433	104	3.7	-0.38	0.33	1.29	1.0	1.31	1.0
469	US434	104	3.4	-1.26	0.36	0.73	-0.8	0.73	-0.8
470	US435	104	3.0	-1.44	0.32	1.38	1.2	1.38	1.2
471	US436	104	4.4	0.08	0.28	1.39	1.4	1.40	1.4
472	US436	304	4.4	-0.03	0.26	2.04	3.6	2.10	3.7
473	US437	104	2.3	-2.01	0.32	1.09	0.4	1.10	0.4
474	US438	106	3.5	0.51	0.33	0.92	-0.1	0.92	-0.1
475	US439	106	3.2	-0.40	0.29	0.86	-0.4	0.87	-0.4
476	US439	306	3.4	-0.85	0.27	1.06	0.3	1.06	0.3
477	US440	106	5.0	2.10	0.30	0.84	-0.5	0.86	-0.4
478	US440	306	5.4	2.14	0.30	1.06	0.3	1.11	0.4
479	US441	106	3.6	0.30	0.33	0.72	-0.9	0.73	-0.9
480	US442	106	4.4	0.43	0.33	1.01	0.1	0.96	0
481	US443	106	4.5	1.74	0.29	1.22	0.9	1.24	0.9
482	US443	306	4.8	2.03	0.27	1.51	1.9	1.47	1.8
483	US444	106	4.8	1.71	0.31	1.26	1.0	1.23	0.9
484	US445	106	3.4	0.73	0.33	0.81	-0.5	0.81	-0.5
485	US446	106	4.3	1.10	0.32	1.05	0.2	1.04	0.2
486	US447	106	4.5	1.43	0.31	1.50	1.6	1.47	1.5
487	US448	106	2.5	-2.02	0.29	1.40	1.4	1.39	1.4
488	US448	306	2.6	-2.63	0.26	1.34	1.4	1.36	1.4
489	US449	106	4.7	1.25	0.32	0.93	-0.1	0.96	0
490	US450	106	3.4	-0.04	0.29	0.80	-0.7	0.81	-0.7
491	US450	306	3.7	-0.21	0.27	0.87	-0.5	0.87	-0.4
492	US451	106	4.0	0.35	0.32	1.78	2.2	1.81	2.3
493	US452	106	4.5	0.94	0.31	0.95	-0.1	0.93	-0.1
494	US453	106	4.0	1.18	0.33	0.62	-1.3	0.64	-1.2
495	US454	106	4.1	-0.63	0.29	1.44	1.5	1.48	1.6
496	US454	306	4.3	-0.86	0.26	1.77	2.7	1.86	2.9
497	US455	106	2.9	-0.42	0.33	0.90	-0.2	0.90	-0.2
498	US456	106	3.4	-0.28	0.33	1.42	1.3	1.44	1.4
499	US457	106	3.7	0.17	0.35	0.87	-0.3	0.88	-0.2
500	US458	106	3.9	0.74	0.32	0.92	-0.1	0.92	-0.1
501	US459	106	2.8	-0.71	0.29	1.56	1.9	1.59	2.0
502	US459	306	3.1	-0.79	0.27	2.24	3.9	2.27	4.0
503	US460	106	3.1	0	0.29	1.67	2.2	1.66	2.2

504	US460	306	3.4	-0.28	0.27	2.26	4.0	2.29	4.1
505	US461	106	3.3	0.43	0.33	0.70	-1.0	0.71	-0.9
506	US462	106	3.5	-0.50	0.29	1.21	0.8	1.21	0.8
507	US462	306	4.2	0.12	0.26	1.04	0.2	1.06	0.3
508	US463	106	3.7	-0.02	0.29	1.33	1.2	1.32	1.2
509	US463	306	4.3	0.72	0.26	0.88	-0.4	0.87	-0.4
510	US464	106	3.9	0.90	0.32	1.46	1.4	1.44	1.4
511	US465	106	2.3	-2.82	0.34	2.38	3.5	2.34	3.2
512	US466	106	4.2	0.98	0.28	1.40	1.5	1.38	1.4
513	US466	306	4.6	0.87	0.26	1.90	3.2	1.91	3.2
514	US467	106	3.6	-0.36	0.33	0.75	-0.8	0.74	-0.9
515	US468	106	3.8	0.55	0.33	0.87	-0.3	0.88	-0.3
516	US469	106	3.5	-0.36	0.33	0.82	-0.5	0.82	-0.5
517	US470	106	4.0	1.16	0.29	1.06	0.3	1.05	0.2
518	US470	306	4.4	1.13	0.26	1.38	1.5	1.40	1.6
519	US471	106	4.2	1.07	0.32	0.92	-0.1	0.92	-0.1
520	US472	106	3.5	0.55	0.32	0.67	-1.1	0.67	-1.1
521	US473	112	4.0	1.03	0.32	0.51	-1.9	0.51	-1.9
522	US474	112	3.9	1.29	0.32	0.54	-1.8	0.53	-1.8
523	US475	112	3.8	1.35	0.33	0.62	-1.4	0.62	-1.4
524	US476	112	3.9	1.21	0.32	0.75	-0.8	0.73	-0.9
525	US477	112	4.1	1.70	0.32	0.82	-0.5	0.79	-0.6
526	US478	112	4.1	1.51	0.33	1.19	0.6	1.19	0.7
527	US479	112	4.2	1.54	0.32	0.40	-2.6	0.40	-2.6
528	US480	112	3.6	0.52	0.32	0.80	-0.6	0.79	-0.6
529	US481	112	4.3	1.42	0.32	0.72	-0.9	0.71	-1.0
530	US482	112	4.1	1.51	0.32	0.83	-0.5	0.82	-0.5
531	US483	112	4.5	1.55	0.32	1.11	0.4	1.07	0.3
532	US484	112	3.7	1.02	0.33	0.92	-0.1	0.92	-0.1
533	US485	112	3.8	1.35	0.33	0.73	-0.8	0.73	-0.8
534	US486	112	4.5	1.74	0.32	0.84	-0.4	0.81	-0.6
535	US487	112	3.8	0.80	0.36	0.75	-0.7	0.75	-0.7
536	US488	112	4.2	1.31	0.32	0.64	-1.3	0.62	-1.4
537	US489	112	4.7	2.20	0.31	0.82	-0.6	0.79	-0.7
538	US490	112	3.8	1.09	0.33	1.14	0.5	1.12	0.5
539	US491	112	4.0	1.37	0.32	1.13	0.5	1.15	0.5
540	US492	112	2.8	0.14	0.32	0.85	-0.4	0.85	-0.4
541	US493	112	4.1	1.27	0.32	1.00	0	0.98	0

542	US494	112	3.4	-0.09	0.33	0.91	-0.2	0.92	-0.1
543	US495	112	3.5	1.14	0.33	0.82	-0.5	0.82	-0.5
544	US496	112	3.3	0.12	0.33	0.76	-0.7	0.77	-0.7
545	US497	112	3.5	0.66	0.33	0.75	-0.8	0.75	-0.8
546	US498	112	3.2	0.68	0.32	0.60	-1.5	0.59	-1.5
547	US499	112	2.4	0.01	0.32	1.00	0	1.01	0.1
548	US500	112	3.3	0.24	0.32	0.59	-1.5	0.59	-1.5
549	US501	112	3.9	0.25	0.32	1.09	0.4	1.10	0.4
550	US502	112	2.6	0.13	0.32	0.31	-3.2	0.32	-3.2
551	US503	112	3.6	1.07	0.33	0.60	-1.4	0.60	-1.4
552	US504	112	2.1	-0.45	0.32	0.45	-2.5	0.45	-2.5
553	US505	112	2.3	-0.06	0.35	0.99	0	0.98	0
554	US506	112	2.8	-0.47	0.32	1.16	0.6	1.18	0.7
555	US507	112	3.3	0.48	0.33	0.51	-1.9	0.51	-1.9
556	US508	112	3.4	0.35	0.32	0.73	-0.9	0.74	-0.9
557	US509	112	3.4	0.50	0.32	0.91	-0.2	0.90	-0.2
558	US510	112	3.8	1.01	0.32	0.76	-0.8	0.76	-0.8
559	US511	112	2.9	0.16	0.35	0.70	-0.9	0.70	-0.9
560	US512	112	3.5	1.21	0.33	0.91	-0.2	0.91	-0.1
561	US513	112	3.7	0.57	0.32	1.54	1.7	1.54	1.7
562	US514	112	3.4	1.35	0.33	0.70	-1.0	0.70	-1.0
563	US515	112	2.9	0.49	0.32	0.54	-1.8	0.54	-1.8
564	US516	112	3.1	0.88	0.32	1.05	0.2	1.04	0.2
565	US517	112	5.1	2.86	0.33	0.73	-1.0	0.76	-0.9
566	US518	112	3.8	1.79	0.33	1.12	0.4	1.13	0.5
567	US519	112	3.8	1.39	0.33	0.71	-1.0	0.71	-0.9
568	US520	112	3.3	1.21	0.32	0.58	-1.6	0.57	-1.6
569	US521	112	4.4	1.68	0.32	0.90	-0.2	0.87	-0.3
570	US522	112	4.1	1.65	0.32	0.39	-2.6	0.39	-2.6
571	US523	112	5.6	3.85	0.41	0.61	-1.3	0.79	-0.5
572	US524	112	4.3	2.05	0.32	0.61	-1.4	0.61	-1.4
573	US525	112	4.5	2.29	0.31	0.55	-1.8	0.55	-1.8
574	US526	112	4.3	1.69	0.32	0.67	-1.2	0.65	-1.3
575	US527	112	4.2	1.87	0.32	0.86	-0.4	0.86	-0.4
576	US528	112	4.3	1.97	0.32	0.82	-0.5	0.83	-0.5
577	US529	112	3.8	1.22	0.33	0.57	-1.6	0.57	-1.6
578	US530	112	3.9	0.79	0.32	0.88	-0.3	0.89	-0.2
579	US531	112	4.6	2.24	0.31	1.01	0.1	1.01	0.1

580	US532	112	4.1	2.19	0.32	0.82	-0.5	0.83	-0.5
581	US533	112	5.2	3.22	0.33	0.96	0	0.96	0
582	US534	112	4.3	1.47	0.33	0.94	-0.1	0.91	-0.2
583	US535	112	3.8	1.04	0.33	0.81	-0.5	0.82	-0.5
584	US536	112	4.2	1.22	0.32	0.85	-0.4	0.84	-0.5
585	US537	112	4.0	1.55	0.32	0.81	-0.5	0.83	-0.5
586	US538	112	3.5	1.12	0.33	0.66	-1.2	0.66	-1.2
587	US539	112	4.9	2.15	0.31	1.58	1.9	1.55	1.8
588	US540	112	5.2	3.28	0.32	1.26	1.0	1.30	1.1
589	US541	112	4.0	1.77	0.32	0.91	-0.2	0.92	-0.1
590	US542	112	3.9	1.24	0.32	0.57	-1.6	0.57	-1.5
591	US543	112	4.4	1.77	0.32	0.67	-1.2	0.69	-1.1
592	US544	112	4.6	1.94	0.31	0.74	-0.9	0.75	-0.8
593	US545	112	4.1	0.84	0.32	0.73	-0.9	0.75	-0.8
594	US546	112	4.1	1.40	0.32	0.58	-1.5	0.58	-1.5
595	US547	112	3.8	0.78	0.33	0.53	-1.7	0.53	-1.7
596	US548	112	4.3	1.31	0.32	0.46	-2.2	0.45	-2.2
597	US549	112	3.3	0.29	0.33	0.33	-3.0	0.33	-3.0
598	US550	112	3.9	0.98	0.36	0.70	-0.9	0.70	-0.9
599	US551	112	3.5	0.50	0.33	1.16	0.6	1.14	0.5
600	US552	112	4.5	2.03	0.32	0.57	-1.7	0.58	-1.7
601	US553	112	4.4	2.55	0.31	0.75	-0.8	0.75	-0.8
602	US554	112	3.6	0.96	0.33	0.78	-0.7	0.78	-0.7
603	US555	112	4.9	2.35	0.34	0.43	-2.4	0.44	-2.4
604	US556	112	3.9	1.09	0.33	0.46	-2.1	0.45	-2.1
605	US557	112	3.5	0.82	0.33	0.61	-1.4	0.61	-1.4
606	US558	112	2.5	-0.31	0.32	1.00	0	1.02	0.1
607	US559	112	3.1	0.34	0.32	0.59	-1.5	0.58	-1.5
608	US560	112	3.3	0.43	0.33	0.42	-2.4	0.42	-2.4
609	US561	112	2.2	-0.37	0.32	0.35	-3.1	0.35	-3.1
610	US562	112	3.3	1.01	0.33	0.68	-1.1	0.67	-1.1
611	US563	112	2.5	0.13	0.32	0.71	-1.0	0.71	-1.0
612	US564	112	3.5	1.36	0.32	0.55	-1.7	0.55	-1.7
613	US565	112	3.9	1.16	0.33	0.76	-0.7	0.76	-0.7
614	US566	112	3.2	0.58	0.33	0.58	-1.5	0.58	-1.6
615	US567	112	3.8	0.71	0.33	1.07	0.3	1.06	0.3
616	US568	112	2.3	-0.65	0.32	1.37	1.3	1.37	1.3
617	US569	112	2.9	-0.19	0.32	0.74	-0.8	0.74	-0.8

618	US570	112	2.5	-0.48	0.32	0.84	-0.5	0.84	-0.5
619	US571	104	4.7	1.16	0.31	0.82	-0.6	0.81	-0.6
620	US572	104	3.2	-1.13	0.33	2.42	3.6	2.42	3.6
621	US573	104	4.5	0.39	0.31	1.46	1.5	1.50	1.6
622	US574	104	3.5	-0.21	0.33	0.88	-0.3	0.88	-0.3
623	US575	104	3.8	-0.39	0.33	0.54	-1.7	0.53	-1.7
624	US576	104	4.0	0.18	0.33	0.72	-0.9	0.71	-0.9
625	US577	104	3.1	-1.43	0.30	3.84	6.4	3.89	6.5
626	US577	304	3.3	-0.55	0.28	2.45	4.1	2.47	4.2
627	US578	104	5.0	1.66	0.34	1.06	0.3	1.06	0.3
628	US579	104	3.8	-1.14	0.33	7.47	9.0	6.85	9.0
629	US580	104	3.5	-0.44	0.33	1.13	0.5	1.15	0.5
630	US581	104	3.1	-1.52	0.33	2.06	2.8	2.07	2.8
631	US582	104	3.9	-0.20	0.36	0.55	-1.5	0.54	-1.6
632	US583	104	4.0	0.46	0.33	0.78	-0.6	0.77	-0.7
633	US584	104	3.8	-0.56	0.33	0.49	-2.0	0.49	-2.0
634	US585	104	3.5	-2.42	0.32	0.89	-0.3	0.91	-0.2
635	US586	104	2.7	-0.96	0.29	1.31	1.1	1.33	1.2
636	US586	304	2.7	-0.56	0.26	1.13	0.6	1.15	0.6
637	US587	104	2.7	-1.78	0.32	0.76	-0.8	0.75	-0.8
638	US588	104	3.5	0.38	0.36	0.51	-1.8	0.50	-1.8
639	US589	104	3.5	-0.96	0.33	0.59	-1.5	0.59	-1.5
640	HK001	206	1.3	-7.32	1.17	0.49	-0.4	0.43	-0.4
641	HK002	206	3.2	-2.82	0.65	0.94	0	0.94	0
642	HK003	206	3.8	-1.35	0.65	0.93	0	0.95	0.1
643	HK004	206	3.3	-2.39	0.66	1.15	0.4	1.15	0.4
644	HK005	206	1.3	-8.67	1.17	0.74	0	0.49	0
645	HK006	206	3.5	-1.96	0.66	0.57	-0.6	0.57	-0.6
646	HK007	206	5.3	2.35	0.68	0.18	-2.2	0.21	-1.9
647	HK008	206	3.2	-2.82	0.65	0.87	0	0.87	0
648	HK009	206	3.7	-1.78	0.66	2.46	1.9	2.46	1.9
649	HK010	206	4.8	1.34	0.62	0.69	-0.4	0.69	-0.4
650	HK011	206	2.3	-5.11	0.63	2.51	2.1	2.59	2.1
651	HK012	206	4.7	0.95	0.62	0.35	-1.4	0.36	-1.4
652	HK013	206	3.7	-1.78	0.66	0.80	-0.1	0.80	-0.1
653	HK014	206	4.2	-0.26	0.64	1.38	0.7	1.34	0.7
654	HK015	206	5.3	2.35	0.68	0.44	-1.1	0.60	-0.6
655	HK016	206	4.2	-0.25	0.65	0.11	-2.4	0.11	-2.4

656	HK017	206	3.0	-3.49	0.64	0.41	-1.1	0.41	-1.1
657	HK018	206	3.3	-2.39	0.66	0.21	-1.8	0.21	-1.8
658	HK019	206	3.8	-1.35	0.65	1.78	1.2	1.80	1.3
659	HK020	206	3.0	-3.24	0.65	0.70	-0.3	0.71	-0.3
660	HK021	206	3.8	-1.35	0.65	8.33	5.5	8.14	5.4
661	HK022	206	1.6	-6.68	0.78	0.57	-0.6	0.68	-0.4
662	HK023	206	4.0	-0.93	0.65	1.85	1.3	1.80	1.3
663	HK024	206	4.8	1.34	0.62	1.15	0.4	1.17	0.4
664	HK025	206	4.7	0.95	0.63	0.75	-0.3	0.75	-0.3
665	HK026	206	3.7	-1.78	0.66	0.32	-1.4	0.32	-1.4
666	HK027	206	2.8	-3.65	0.64	1.84	1.3	1.88	1.4
667	HK028	206	3.2	-3.07	0.65	0.36	-1.2	0.35	-1.3
668	HK029	206	3.0	-3.24	0.65	0.78	-0.2	0.78	-0.2
669	HK030	206	4.2	-0.51	0.64	0.95	0.1	0.98	0.1
670	HK031	206	3.2	-2.82	0.65	0.92	0	0.92	0
671	HK032	206	4.8	1.09	0.63	0.56	-0.8	0.57	-0.7
672	HK033	206	3.0	-3.24	0.65	0.31	-1.4	0.30	-1.4
673	HK034	206	4.2	-0.51	0.64	1.61	1.0	1.67	1.1
674	HK035	206	3.3	-2.39	0.66	0.21	-1.8	0.21	-1.8
675	HK036	206	3.8	-1.40	0.71	8.79	5.2	8.85	5.2
676	HK037	206	3.0	-2.76	0.80	1.19	0.4	1.19	0.4
677	HK038	206	3.7	-1.78	0.66	0.84	0	0.85	0
678	HK039	206	3.0	-3.24	0.65	1.44	0.8	1.46	0.8
679	HK040	206	4.0	-0.93	0.65	0.64	-0.5	0.64	-0.5
680	HK041	206	2.8	-3.65	0.64	0.72	-0.3	0.73	-0.3
681	HK042	206	4.2	-0.51	0.64	0.92	0	0.95	0.1
682	HK043	206	3.0	-3.24	0.65	0.34	-1.3	0.34	-1.3
683	HK044	206	3.8	-1.35	0.65	3.28	2.7	3.25	2.6
684	HK045	206	3.2	-2.82	0.65	0.61	-0.5	0.60	-0.5
685	HK046	206	2.8	-3.90	0.64	0.66	-0.4	0.66	-0.4
686	HK047	206	3.5	-1.96	0.66	0.19	-1.9	0.19	-1.9
687	HK048	206	5.2	1.90	0.65	2.09	1.7	2.38	2.0
688	HK049	206	3.7	-1.53	0.66	2.35	1.8	2.34	1.8
689	HK050	206	3.5	-2.12	0.80	0.89	0	0.90	0.1
690	HK051	206	2.8	-3.65	0.64	0.76	-0.2	0.77	-0.2
691	HK052	206	3.2	-3.07	0.65	0.36	-1.2	0.35	-1.3
692	HK053	206	2.8	-3.65	0.64	0.72	-0.3	0.73	-0.3
693	HK054	206	3.8	-1.35	0.65	2.65	2.1	2.61	2.1

694	HK055	206	3.8	-1.10	0.65	2.91	2.3	2.96	2.4
695	HK056	206	2.5	-4.71	0.63	1.68	1.1	1.70	1.2
696	HK057	206	2.0	-5.67	0.64	0.27	-1.8	0.28	-1.8
697	HK058	206	3.3	-2.64	0.66	2.36	1.8	2.37	1.8
698	HK059	206	3.3	-2.39	0.66	2.54	2.0	2.53	2.0
699	HK060	206	2.7	-4.31	0.64	0.49	-0.9	0.50	-0.8
700	HK061	206	2.8	-3.65	0.64	0.37	-1.2	0.37	-1.2
701	HK062	206	2.7	-4.31	0.64	2.85	2.4	2.95	2.4
702	HK063	206	4.2	-0.25	0.65	2.24	1.7	2.24	1.7
703	HK064	206	2.5	-4.71	0.63	1.49	0.9	1.52	0.9
704	HK065	206	4.0	-0.67	0.65	1.91	1.4	1.82	1.3
705	HK066	206	3.2	-3.07	0.65	1.23	0.5	1.24	0.5
706	HK067	206	4.0	-0.68	0.65	0.76	-0.2	0.76	-0.2
707	HK068	206	2.8	-3.90	0.64	0.77	-0.2	0.78	-0.2
708	HK069	206	2.8	-3.65	0.64	1.80	1.3	1.76	1.2
709	HK070	206	4.0	-0.93	0.65	1.06	0.2	1.08	0.3
710	HK071	206	2.3	-4.86	0.63	0.55	-0.8	0.57	-0.7
711	HK072	206	2.3	-5.16	0.90	2.77	1.8	2.74	1.7
712	HK073	206	2.8	-3.65	0.64	3.11	2.6	3.03	2.5
713	HK074	206	2.0	-5.93	0.65	2.01	1.6	2.00	1.6
714	HK075	206	4.2	-0.25	0.65	1.15	0.4	1.17	0.4
715	HK076	206	1.8	-6.36	0.67	1.23	0.5	1.14	0.4
716	HK077	206	2.2	-5.10	0.69	0.92	0	0.91	0
717	HK078	206	2.4	-5.09	0.70	3.25	2.5	3.28	2.5
718	HK079	206	3.8	-1.10	0.65	0.84	0	0.84	0
719	HK080	206	3.0	-3.49	0.64	3.41	2.8	3.35	2.7
720	HK081	206	1.0	-8.30	1.97	Minim	um		
721	HK082	206	1.0	-8.43	1.95	Minim	um		
722	HK083	206	3.5	-1.96	0.66	1.54	0.9	1.53	0.9
723	HK084	206	4.3	-0.43	0.78	4.14	2.8	4.42	2.9
724	HK085	206	2.5	-4.46	0.63	0.43	-1.1	0.43	-1.1
725	HK086	206	2.2	-5.58	0.70	7.92	5.2	7.79	5.1
726	HK087	206	3.7	-1.53	0.66	2.08	1.6	2.06	1.5
727	HK088	206	1.5	-7.36	0.76	1.05	0.2	1.04	0.2
728	HK089	206	1.8	-6.28	0.73	0.49	-0.9	0.50	-0.9
729	HK090	206	3.3	-2.64	0.66	3.66	2.9	3.62	2.8
730	HK091	206	2.2	-5.26	0.64	1.23	0.5	1.23	0.5
731	HK092	206	2.3	-5.11	0.63	1.06	0.2	1.07	0.3

732	HK093	206	2.7	-4.06	0.64	0.42	-1.1	0.41	-1.1
733	HK094	206	4.0	-0.93	0.65	2.59	2.1	2.67	2.1
734	HK095	206	5.0	1.73	0.63	2.12	1.8	2.05	1.7
735	HK096	206	3.5	-2.21	0.66	0.14	-2.2	0.14	-2.2
736	HK097	206	3.3	-2.39	0.66	1.05	0.2	1.04	0.2
737	HK098	206	1.0	-10.08	1.91	Minim	um		
738	HK099	206	2.8	-3.65	0.64	0.48	-0.9	0.47	-0.9
739	HK100	206	2.3	-5.11	0.63	0.93	0	0.91	0
740	HK101	206	3.7	-1.53	0.66	0.77	-0.2	0.77	-0.2
741	HK102	206	2.3	-5.11	0.63	0.45	-1.0	0.45	-1.0
742	HK103	206	3.8	-1.10	0.65	0.92	0	0.93	0
743	HK104	206	2.6	-4.22	0.69	1.42	0.8	1.41	0.7
744	HK105	206	2.7	-4.06	0.64	0.42	-1.1	0.41	-1.1
745	HK106	206	2.0	-5.93	0.65	1.31	0.7	1.35	0.7
746	HK107	206	2.5	-4.46	0.63	0.52	-0.8	0.51	-0.8
747	HK108	206	1.4	-7.87	0.88	0.82	0	0.70	-0.1
748	HK109	206	4.0	-0.68	0.65	0.34	-1.3	0.34	-1.3
749	HK110	206	2.8	-3.90	0.64	1.86	1.4	1.91	1.4
750	HK111	206	2.5	-4.46	0.63	0.60	-0.6	0.63	-0.5
751	HK112	206	3.3	-2.64	0.66	4.42	3.3	4.39	3.3
752	HK113	206	4.8	1.34	0.62	1.66	1.2	1.65	1.2
753	HK114	206	3.8	-1.35	0.65	5.06	3.8	5.05	3.8
754	HK115	206	3.2	-2.82	0.65	2.10	1.6	2.09	1.6
755	HK116	206	4.2	-0.51	0.64	0.25	-1.7	0.23	-1.7
756	HK117	206	3.0	-3.24	0.65	2.10	1.6	2.10	1.6
757	HK118	206	3.8	-1.35	0.65	0.94	0	0.95	0.1
758	HK119	206	3.7	-1.53	0.66	0.93	0	0.93	0
759	HK120	206	3.0	-3.49	0.64	2.77	2.3	2.76	2.3
760	HK121	206	3.0	-3.24	0.65	2.64	2.1	2.69	2.2
761	HK122	206	2.0	-5.93	0.65	0.54	-0.8	0.55	-0.8
762	HK123	206	3.3	-2.39	0.66	0.23	-1.7	0.23	-1.7
763	HK124	206	1.3	-8.01	0.86	0.29	-1.3	0.25	-1.1
764	HK125	206	3.2	-2.82	0.65	0.94	0	0.94	0.1
765	HK126	206	4.7	0.69	0.63	1.54	1.0	1.54	1.0
766	HK127	206	5.0	1.74	0.63	0.98	0.1	0.98	0.1
767	HK128	206	2.7	-4.31	0.64	2.44	2.0	2.55	2.1
768	HK129	206	5.2	2.14	0.65	0.45	-1.1	0.58	-0.7
769	HK130	206	4.2	-0.51	0.64	1.42	0.8	1.50	0.9

770	HK131	206	1.8	-6.64	0.82	0.20	-1.7	0.23	-1.6
771	HK132	206	4.0	-0.93	0.65	0.56	-0.7	0.56	-0.7
772	HK133	206	3.0	-3.24	0.65	1.92	1.4	1.92	1.4
773	HK134	206	3.0	-3.49	0.64	0.44	-1.0	0.44	-1.0
774	HK135	206	5.8	4.53	1.08	0.99	0.2	0.82	0.2
775	HK136	206	3.0	-3.49	0.64	1.03	0.2	1.04	0.2
776	HK137	206	2.3	-4.86	0.63	1.04	0.2	1.07	0.2

Separation: 4.38 Reliability: 0.95

Fixed (all same) chi-square: 12555.1 significance *p*: .00

Note: 104 = Grade 4 US student, 106 = Grade 6 US student, 108 = Grade 8 US student, 110 = Grade 10 US student, 112 = Grade 12 US student, 206 = Grade 6 HK student, 304 = Grade 4 US sub-sample student, 306 = Grade 6 US sub-sample student

Calibration of the students for all student essays using US and HK raters and IEA (129 misfitting students)

Student Number		Group	Observed score	Measure (Logit)	S.E.	Infit MnSq	Infit ZStd	Outfit MnSq	Outfit ZStd
197	US182	108	3.3	0.63	0.36	1.41	1.2	1.39	1.1
266	US243	106	4.0	0.92	0.29	1.41	1.4	1.42	1.4
78	US064	306	4.2	-0.25	0.26	1.42	1.5	1.37	1.4
300	US272	104	4.5	-0.76	0.32	1.42	1.4	1.44	1.4
309	US280	104	4.4	-0.22	0.39	1.42	1.1	1.40	1.1
498	US456	106	3.4	-0.28	0.33	1.42	1.3	1.44	1.4
743	HK104	206	2.6	-4.22	0.69	1.42	0.8	1.41	0.7
769	HK130	206	4.2	-0.51	0.64	1.42	0.8	1.50	0.9
495	US454	106	4.1	-0.63	0.29	1.44	1.5	1.48	1.6
678	HK039	206	3.0	-3.24	0.65	1.44	0.8	1.46	0.8
510	US464	106	3.9	0.90	0.32	1.46	1.4	1.44	1.4
621	US573	104	4.5	0.39	0.31	1.46	1.5	1.50	1.6
56	US046	304	2.7	-1.67	0.26	1.47	1.8	1.50	1.9
287	US260	104	3.0	-1.72	0.32	1.47	1.5	1.47	1.5
307	US278	104	4.8	1.28	0.31	1.47	1.6	1.45	1.5
5	US005	104	3.1	-2.29	0.32	1.48	1.5	1.50	1.6
227	US212	104	3.3	-1.01	0.29	1.49	1.7	1.50	1.7
373	US343	110	4.5	1.39	0.31	1.49	1.5	1.47	1.5
390	US360	110	3.4	-0.38	0.33	1.49	1.5	1.51	1.5
703	HK064	206	2.5	-4.71	0.63	1.49	0.9	1.52	0.9
486	US447	106	4.5	1.43	0.31	1.50	1.6	1.47	1.5

298	US270	104	3.9	0.04	0.36	1.51	1.4	1.51	1.4
482	US443	306	4.8	2.03	0.27	1.51	1.9	1.47	1.8
314	US285	104	1.2	-5.69	0.57	1.52	1.1	2.18	1.6
150	US135	108	5.3	3.18	0.33	1.53	1.8	1.59	2.0
293	US266	104	2.4	-1.83	0.28	1.53	1.9	1.52	1.8
561	US513	112	3.7	0.57	0.32	1.54	1.7	1.54	1.7
722	HK083	206	3.5	-1.96	0.66	1.54	0.9	1.53	0.9
765	HK126	206	4.7	0.69	0.63	1.54	1.0	1.54	1.0
176	US161	108	3.4	0.32	0.36	1.56	1.5	1.58	1.6
267	US243	306	4.4	1.35	0.26	1.56	2.1	1.55	2.0
501	US459	106	2.8	-0.71	0.29	1.56	1.9	1.59	2.0
236	US218	304	4.2	0.27	0.26	1.57	2.0	1.57	2.0
255	US235	306	2.6	-1.00	0.26	1.57	2.1	1.55	2.1
318	US288	104	4.5	0.74	0.31	1.57	1.8	1.58	1.8
322	US292	106	5.2	2.27	0.34	1.58	1.8	1.32	0.9
587	US539	112	4.9	2.15	0.31	1.58	1.9	1.55	1.8
673	HK034	206	4.2	-0.51	0.64	1.61	1.0	1.67	1.1
39	US031	106	3.5	-0.75	0.32	1.63	1.9	1.63	1.9
84	US070	106	2.3	-1.35	0.29	1.63	2.1	1.63	2.1
274	US248	306	4.9	2.06	0.26	1.65	2.5	1.59	2.3
387	US357	110	4.1	1.08	0.32	1.65	1.9	1.67	2.0
461	US427	304	3.3	-0.59	0.27	1.65	2.3	1.63	2.2
752	HK113	206	4.8	1.34	0.62	1.66	1.2	1.65	1.2
73	US060	304	2.0	-3.22	0.26	1.67	2.6	1.61	2.4
297	US269	104	5.0	1.30	0.31	1.67	2.2	1.66	2.2
503	US460	106	3.1	0	0.29	1.67	2.2	1.66	2.2
695	HK056	206	2.5	-4.71	0.63	1.68	1.1	1.70	1.2
70	US058	104	2.5	-2.00	0.31	1.72	2.2	1.73	2.2
63	US052	104	3.2	-0.93	0.33	1.73	2.1	1.74	2.2
203	US188	108	3.3	0.89	0.33	1.73	2.1	1.74	2.2
116	US101	108	2.9	-0.56	0.32	1.77	2.2	1.77	2.2
210	US195	108	3.6	-0.33	0.36	1.77	2.1	1.77	2.0
283	US256	304	1.8	-2.85	0.34	1.77	2.4	1.71	2.2
496	US454	306	4.3	-0.86	0.26	1.77	2.7	1.86	2.9
492	US451	106	4.0	0.35	0.32	1.78	2.2	1.81	2.3
658	HK019	206	3.8	-1.35	0.65	1.78	1.2	1.80	1.3
708	HK069	206	2.8	-3.65	0.64	1.80	1.3	1.76	1.2
85	US070	306	2.4	-2.00	0.26	1.82	3.0	1.89	3.2

43	US034	306	3.3	-0.11	0.27	1.83	2.8	1.83	2.8
228	US212	304	3.0	-1.54	0.26	1.83	2.8	1.84	2.8
666	HK027	206	2.8	-3.65	0.64	1.84	1.3	1.88	1.4
662	HK023	206	4.0	-0.93	0.65	1.85	1.3	1.80	1.3
749	HK110	206	2.8	-3.90	0.64	1.86	1.4	1.91	1.4
242	US224	104	2.4	-1.60	0.39	1.88	2.1	1.88	2.1
19	US017	104	1.6	-3.46	0.34	1.89	2.9	1.85	2.7
513	US466	306	4.6	0.87	0.26	1.90	3.2	1.91	3.2
704	HK065	206	4.0	-0.67	0.65	1.91	1.4	1.82	1.3
772	HK133	206	3.0	-3.24	0.65	1.92	1.4	1.92	1.4
258	US237	106	3.3	-0.94	0.29	1.93	2.8	1.97	3.0
20	US017	304	1.5	-3.63	0.30	1.94	3.4	1.84	3.0
37	US030	106	3.3	-0.33	0.29	1.95	2.9	1.96	2.9
336	US306	106	3.0	-1.55	0.39	1.95	2.2	2.01	2.3
38	US030	306	3.5	-0.69	0.27	1.97	3.2	1.96	3.2
392	US362	110	3.3	-0.64	0.33	1.98	2.7	1.96	2.6
713	HK074	206	2.0	-5.93	0.65	2.01	1.6	2.00	1.6
472	US436	304	4.4	-0.03	0.26	2.04	3.6	2.10	3.7
294	US266	304	2.1	-2.26	0.26	2.05	3.8	2.02	3.7
630	US581	104	3.1	-1.52	0.33	2.06	2.8	2.07	2.8
726	HK087	206	3.7	-1.53	0.66	2.08	1.6	2.06	1.5
687	HK048	206	5.2	1.90	0.65	2.09	1.7	2.38	2.0
754	HK115	206	3.2	-2.82	0.65	2.10	1.6	2.09	1.6
756	HK117	206	3.0	-3.24	0.65	2.10	1.6	2.10	1.6
273	US248	106	4.0	0.99	0.28	2.12	3.4	2.07	3.3
734	HK095	206	5.0	1.73	0.63	2.12	1.8	2.05	1.7
68	US057	104	3.7	-0.84	0.35	2.14	2.7	2.18	2.8
24	US021	104	3.3	-1.49	0.29	2.16	3.4	2.16	3.4
502	US459	306	3.1	-0.79	0.27	2.24	3.9	2.27	4.0
702	HK063	206	4.2	-0.25	0.65	2.24	1.7	2.24	1.7
504	US460	306	3.4	-0.28	0.27	2.26	4.0	2.29	4.1
310	US281	104	4.2	-0.63	0.32	2.27	3.2	2.33	3.3
230	US214	104	2.9	-2.16	0.32	2.29	3.5	2.32	3.5
688	HK049	206	3.7	-1.53	0.66	2.35	1.8	2.34	1.8
697	HK058	206	3.3	-2.64	0.66	2.36	1.8	2.37	1.8
511	US465	106	2.3	-2.82	0.34	2.38	3.5	2.34	3.2
329	US299	106	3.8	0.43	0.35	2.39	3.3	2.50	3.4
69	US057	304	3.5	-0.36	0.33	2.41	3.5	2.40	3.5

620	US572	104	3.2	-1.13	0.33	2.42	3.6	2.42	3.6
259	US237	306	3.4	-1.51	0.27	2.43	4.4	2.44	4.4
767	HK128	206	2.7	-4.31	0.64	2.44	2.0	2.55	2.1
626	US577	304	3.3	-0.55	0.28	2.45	4.1	2.47	4.2
648	HK009	206	3.7	-1.78	0.66	2.46	1.9	2.46	1.9
396	US366	110	3.9	1.14	0.32	2.50	3.6	2.47	3.5
650	HK011	206	2.3	-5.11	0.63	2.51	2.1	2.59	2.1
698	HK059	206	3.3	-2.39	0.66	2.54	2.0	2.53	2.0
733	HK094	206	4.0	-0.93	0.65	2.59	2.1	2.67	2.1
760	HK121	206	3.0	-3.24	0.65	2.64	2.1	2.69	2.2
693	HK054	206	3.8	-1.35	0.65	2.65	2.1	2.61	2.1
25	US021	304	3.1	-1.65	0.27	2.75	5.0	2.76	5.1
711	HK072	206	2.3	-5.16	0.90	2.77	1.8	2.74	1.7
759	HK120	206	3.0	-3.49	0.64	2.77	2.3	2.76	2.3
59	US049	104	2.7	-1.81	0.29	2.81	4.8	2.85	4.9
701	HK062	206	2.7	-4.31	0.64	2.85	2.4	2.95	2.4
694	HK055	206	3.8	-1.10	0.65	2.91	2.3	2.96	2.4
712	HK073	206	2.8	-3.65	0.64	3.11	2.6	3.03	2.5
60	US049	304	2.7	-1.85	0.26	3.13	6.0	3.16	6.0
717	HK078	206	2.4	-5.09	0.70	3.25	2.5	3.28	2.5
683	HK044	206	3.8	-1.35	0.65	3.28	2.7	3.25	2.6
719	HK080	206	3.0	-3.49	0.64	3.41	2.8	3.35	2.7
729	HK090	206	3.3	-2.64	0.66	3.66	2.9	3.62	2.8
625	US577	104	3.1	-1.43	0.30	3.84	6.4	3.89	6.5
723	HK084	206	4.3	-0.43	0.78	4.14	2.8	4.42	2.9
299	US271	104	3.2	-1.36	0.36	4.22	5.8	4.27	5.9
751	HK112	206	3.3	-2.64	0.66	4.42	3.3	4.39	3.3
753	HK114	206	3.8	-1.35	0.65	5.06	3.8	5.05	3.8
628	US579	104	3.8	-1.14	0.33	7.47	9.0	6.85	9.0
725	HK086	206	2.2	-5.58	0.70	7.92	5.2	7.79	5.1
660	HK021	206	3.8	-1.35	0.65	8.33	5.5	8.14	5.4
675	HK036	206	3.8	-1.40	0.71	8.79	5.2	8.85	5.2

Separation: 4.38 Reliability: 0.95

Fixed (all same) chi-square: 12555.1 significance *p*: .00

Note: 104 = Grade 4 US student, 106 = Grade 6 US student, 108 = Grade 8 US student, 110 = Grade 10 US student, 112 = Grade 12 US student, 206 = Grade 6 HK student, 304 = Grade 4 US sub-sample student, 306 = Grade 6 US sub-sample student

Calibration of the prompts for all student essays using US and HK raters and IEA

Prompt	Observed score	Measure (Logit)	S.E.	Infit MnSq	Infit ZStd	Outfit MnSq	Outfit ZStd
4AI	3.6	-2.39	0.07	0.76	-4.3	0.75	-4.4
4AN	3.7	-2.34	0.07	0.75	-4.5	0.79	-3.6
4AP	3.6	-2.44	0.07	1.61	8.3	1.58	7.9
4BI	3.6	-1.83	0.04	1.26	7.7	1.26	7.5
4BN	3.6	-1.75	0.04	0.97	-0.9	0.98	-0.6
4BP	3.5	-1.69	0.04	1.57	9.0	1.57	9.0
8AI	3.5	0.16	0.05	1.02	0.3	1.02	0.4
8AN	3.7	-0.13	0.05	0.63	-9.0	0.62	-9.0
8AP	2.9	1.69	0.05	0.83	-3.9	0.84	-3.8
8BI	3.8	0.33	0.05	0.87	-2.8	0.87	-2.9
8BN	3.7	0.60	0.05	0.63	-8.8	0.64	-8.7
8BP	3.6	0.91	0.05	1.09	1.9	1.09	1.8
12AI	3.8	2.12	0.06	0.76	-5.0	0.76	-5.0
12AN	3.9	0.56	0.06	0.69	-6.8	0.69	-6.7
12AP	3.4	2.16	0.06	0.80	-4.1	0.80	-4.0
12BI	3.5	2.07	0.08	0.82	-2.6	0.81	-2.7
12BN	4.0	0.94	0.08	0.76	-3.6	0.75	-3.7
12BP	4.0	1.04	0.08	0.99	0	0.99	-0.1

Separation: 26.99 Reliability: 1.00

Fixed (all same) chi-square: 15602.0 significance *p*: .00

Calibration of the US raters for all student essays using US and HK raters and IEA

US rater	Observed score	Measure (Logit)	S.E.	Infit MnSq	Infit ZStd	Outfit MnSq	Outfit ZStd
USR101	3.8	0.09	0.05	0.70	-7.2	0.73	-6.3
USR102	3.8	0.29	0.05	0.66	-8.7	0.67	-8.5
USR103	3.6	0.78	0.09	0.61	-5.7	0.61	-5.5
USR104	3.6	0.84	0.05	0.62	-9.0	0.63	-9.0
USR105	3.8	0.30	0.05	0.80	-4.7	0.79	-4.8
USR106	4.0	-0.32	0.06	0.58	-9.0	0.58	-8.9
USR108	3.3	-1.70	0.37	0.56	-1.4	0.55	-1.5
USR200	3.2	1.04	0.06	1.15	2.5	1.15	2.5
USR201	3.6	-0.20	0.05	1.08	1.8	1.08	1.6

USR202	3.6	-0.49	0.06	1.09	1.6	1.08	1.4
USR203	3.5	-0.03	0.05	1.07	1.7	1.08	1.8
USR204	3.6	-0.04	0.05	0.85	-3.6	0.85	-3.5
USR205	3.5	-0.12	0.06	0.96	-0.8	0.96	-0.7
USR206	3.4	0.23	0.05	0.80	-4.8	0.80	-4.8
USR207	3.6	-0.47	0.06	0.95	-1.0	0.94	-1.1
USR208	3.7	-0.41	0.06	0.95	-0.9	0.95	-1.0
USR209	3.4	0.56	0.06	1.00	0	0.99	0
USR210	3.7	-0.17	0.05	1.20	4.2	1.20	4.1
USR211	3.2	0.59	0.35	0.73	-0.8	0.70	-0.9
Separation: 4.89			Reliability: 0.96				
Fixed (all same) chi-square: 1173.0			significance *p*: .00				

Calibration of the HK raters for all student essays using US and HK raters and IEA

HK rater	Observed score	Measure (Logit)	S.E.	Infit		Outfit	
				MnSq	ZStd	MnSq	ZStd
HKR001	3.1	0.22	0.10	1.63	6.1	1.66	6.2
HKR002	3.5	-0.56	0.10	1.48	4.8	1.47	4.6
HKR003	3.8	-1.66	0.10	2.01	9.0	1.98	8.7
HKR004	3.0	0.63	0.10	1.72	6.9	1.70	6.7
Separation: 2.06			Reliability: 0.81				
Fixed (all same) chi-square: 411.4			significance *p*: .00				

Calibration of IEA for all student essays using US and HK raters and IEA

IEA	Observed score	Measure (Logit)	S.E.	Infit		Outfit	
				MnSq	ZStd	MnSq	ZStd
Category 1	2.2	3.18	0.05	1.10	2.2	1.12	2.5
Category 2	2.8	2.36	0.04	0.93	-1.7	0.94	-1.6
Category 3	3.1	1.56	0.03	1.08	2.4	1.08	2.5
Category 4	3.3	0.98	0.03	1.10	3.2	1.11	3.4
Category 5	3.7	-0.02	0.03	0.89	-4.5	0.88	-4.7
Category 6	4.0	-0.85	0.03	0.92	-2.8	0.92	-2.9
Category 7	4.2	-1.47	0.04	0.97	-0.9	0.96	-1.2
Category 8	4.6	-2.36	0.04	0.96	-1.1	0.95	-1.1
Category 9	5.0	-3.39	0.06	0.91	-1.7	0.92	-1.6
Separation: 49.53			Reliability: 1.00				
Fixed (all same) chi-square: 17741.1			significance *p*: .00				

Appendix 21

Calibration of the students for all student essays using US and HK raters, Lexile Analyzer and IEA

	Student		Observed	Measure	S.E.	Infit		Outfit	
	Number	Group	score	(Logit)		MnSq	ZStd	MnSq	ZStd
1	US001	104	3.7	-0.05	0.33	0.76	-0.7	0.76	-0.7
2	US002	104	3.8	0.10	0.36	0.96	0	0.96	0
3	US003	104	4.1	0.40	0.33	0.89	-0.2	0.89	-0.2
4	US004	104	4.2	0.01	0.32	0.49	-2.1	0.49	-2.0
5	US005	104	3.1	-1.79	0.33	1.32	1.1	1.32	1.1
6	US006	104	4.1	0.05	0.29	0.55	-1.9	0.53	-2.0
7	US006	304	3.9	-0.03	0.27	0.61	-1.7	0.58	-1.9
8	US007	104	3.3	-1.17	0.33	1.03	0.1	1.01	0.1
9	US008	104	3.2	-0.48	0.33	0.51	-1.9	0.51	-1.9
10	US009	104	4.1	0.54	0.33	1.21	0.7	1.21	0.7
11	US010	104	3.5	-0.56	0.33	0.59	-1.5	0.59	-1.5
12	US011	104	5.3	1.62	0.33	0.75	-0.9	0.72	-1.0
13	US012	104	1.7	-3.22	0.44	0.80	-0.5	0.96	0
14	US013	104	4.3	0.86	0.29	0.99	0	1.00	0
15	US013	304	4.3	0.80	0.26	1.09	0.4	1.11	0.5
16	US014	104	4.6	0.54	0.32	0.59	-1.6	0.57	-1.7
17	US015	104	4.1	-0.19	0.36	0.39	-2.3	0.38	-2.3
18	US016	104	3.4	-1.16	0.33	0.65	-1.2	0.65	-1.2
19	US017	104	1.6	-2.90	0.34	1.88	2.8	1.87	2.8
20	US017	304	1.5	-3.02	0.30	2.06	3.7	1.98	3.4
21	US018	104	3.8	-0.68	0.33	1.30	1.0	1.29	1.0
22	US019	104	3.7	-0.53	0.33	0.55	-1.7	0.55	-1.7
23	US020	104	4.2	-0.63	0.33	0.93	-0.1	0.96	0
24	US021	104	3.3	-1.14	0.30	1.95	2.9	1.95	2.9
25	US021	304	3.1	-1.50	0.27	2.16	3.7	2.17	3.7
26	US022	106	2.7	-0.99	0.29	0.61	-1.7	0.62	-1.6
27	US022	306	3.2	-1.17	0.27	0.67	-1.4	0.67	-1.4
28	US023	106	4.0	0.48	0.33	0.67	-1.1	0.67	-1.1
29	US024	106	3.1	-0.15	0.33	0.58	-1.6	0.60	-1.5
30	US025	106	2.8	-0.81	0.33	1.02	0.1	1.02	0.1
31	US026	106	3.5	-0.35	0.29	0.76	-0.9	0.76	-0.9
32	US026	306	3.9	-0.33	0.27	0.88	-0.4	0.88	-0.4

33	US027	106	4.3	1.45	0.29	1.35	1.3	1.34	1.2
34	US027	306	4.8	1.58	0.27	1.26	1.1	1.22	1.0
35	US028	106	3.3	-0.06	0.33	0.60	-1.4	0.60	-1.4
36	US029	106	3.3	0.50	0.33	0.61	-1.4	0.62	-1.4
37	US030	106	3.3	-0.19	0.30	2.09	3.2	2.11	3.3
38	US030	306	3.5	-0.51	0.27	2.25	3.9	2.25	3.9
39	US031	106	3.5	-0.42	0.33	1.72	2.1	1.71	2.1
40	US032	106	3.1	-1.09	0.33	0.74	-0.8	0.73	-0.9
41	US033	106	3.1	0.09	0.33	0.43	-2.4	0.44	-2.4
42	US034	106	3.1	0.69	0.29	1.54	1.8	1.57	1.9
43	US034	306	3.3	0.10	0.27	2.10	3.6	2.10	3.6
44	US035	106	4.3	0.68	0.32	1.32	1.1	1.30	1.0
45	US036	106	3.7	-0.14	0.32	0.75	-0.8	0.74	-0.9
46	US037	106	3.8	0.18	0.33	0.89	-0.2	0.9	-0.2
47	US038	106	3.3	0.47	0.33	0.58	-1.5	0.58	-1.6
48	US039	106	2.5	-0.87	0.32	0.95	0	0.97	0
49	US040	106	3.8	-0.09	0.32	0.78	-0.7	0.78	-0.7
50	US041	104	3.8	-0.51	0.33	0.73	-0.9	0.73	-0.9
51	US042	104	2.6	-1.08	0.32	0.37	-2.8	0.36	-2.9
52	US043	104	3.6	0.19	0.33	0.86	-0.4	0.87	-0.3
53	US044	104	3.8	0.07	0.33	0.99	0	1.01	0.1
54	US045	104	4.0	0.16	0.33	0.57	-1.6	0.57	-1.6
55	US046	104	2.8	-1.03	0.29	0.92	-0.2	0.92	-0.2
56	US046	304	2.7	-1.34	0.26	1.22	0.9	1.25	1.0
57	US047	104	3.6	-0.05	0.33	0.58	-1.5	0.56	-1.6
58	US048	104	3.4	-0.78	0.41	0.93	0	0.92	-0.1
59	US049	104	2.7	-1.71	0.29	2.50	4.2	2.51	4.2
60	US049	304	2.7	-1.80	0.26	2.71	5.1	2.70	5.1
61	US050	104	3.0	-0.88	0.36	1.01	0.1	1.02	0.1
62	US051	104	4.5	0.73	0.32	1.01	0.1	1.02	0.1
63	US052	104	3.2	-0.86	0.33	1.79	2.3	1.80	2.3
64	US053	104	3.8	-0.04	0.36	0.58	-1.4	0.57	-1.4
65	US054	104	3.6	-0.50	0.33	1.23	0.8	1.25	0.9
66	US055	104	3.0	-1.01	0.33	1.02	0.1	1.02	0.1
67	US056	104	2.9	-1.95	0.33	1.34	1.1	1.35	1.1
68	US057	104	3.7	-0.19	0.36	1.98	2.4	2.02	2.5
69	US057	304	3.5	0.15	0.33	2.27	3.2	2.27	3.2
70	US058	104	2.5	-1.38	0.32	2.04	2.9	2.07	3.0

71	US059	104	3.3	-0.99	0.36	0.40	-2.3	0.41	-2.3
72	US060	104	2.3	-2.01	0.29	1.23	0.9	1.22	0.9
73	US060	304	2.0	-2.77	0.27	1.27	1.1	1.21	0.9
74	US061	104	2.4	-2.19	0.39	0.72	-0.8	0.71	-0.8
75	US062	106	3.7	-0.39	0.33	1.33	1.0	1.32	1.0
76	US063	106	3.9	0.52	0.40	0.62	-1.0	0.59	-1.2
77	US064	106	3.8	-0.40	0.30	1.04	0.2	1.00	0
78	US064	306	4.2	-0.11	0.27	1.46	1.7	1.41	1.5
79	US065	106	3.7	-0.20	0.33	1.10	0.4	1.12	0.4
80	US066	106	4.5	1.33	0.32	0.76	-0.8	0.76	-0.8
81	US067	106	3.9	1.08	0.33	0.74	-0.9	0.73	-0.9
82	US068	106	3.8	0.24	0.33	1.09	0.3	1.09	0.3
83	US069	106	3.5	-0.26	0.33	0.84	-0.4	0.84	-0.4
84	US070	106	2.3	-0.97	0.30	1.67	2.2	1.68	2.1
85	US070	306	2.4	-1.73	0.26	1.75	2.8	1.81	2.9
86	US071	106	3.8	0.12	0.33	1.26	0.9	1.26	0.9
87	US072	106	3.0	-0.21	0.33	1.08	0.3	1.09	0.3
88	US073	106	2.7	-0.69	0.33	1.16	0.6	1.16	0.6
89	US074	106	3.6	0.47	0.33	1.07	0.3	1.08	0.3
90	US075	106	4.3	1.34	0.32	1.23	0.8	1.24	0.8
91	US076	106	4.0	0.82	0.33	0.45	-2.2	0.44	-2.3
92	US077	106	2.5	0.02	0.36	0.56	-1.5	0.56	-1.5
93	US078	106	4.2	-0.10	0.33	1.04	0.2	1.04	0.2
94	US079	106	4.1	-0.01	0.33	0.46	-2.2	0.45	-2.2
95	US080	106	2.5	-1.38	0.32	0.61	-1.5	0.62	-1.4
96	US081	106	3.5	0.06	0.33	0.78	-0.7	0.78	-0.7
97	US082	108	3.4	0.54	0.33	0.48	-2.1	0.48	-2.1
98	US083	108	3.1	-0.47	0.33	0.63	-1.3	0.63	-1.3
99	US084	108	1.5	-1.52	0.42	0.87	-0.3	0.84	-0.4
100	US085	108	3.0	0	0.33	0.72	-1.0	0.73	-0.9
101	US086	108	2.8	-0.72	0.32	0.59	-1.5	0.60	-1.5
102	US087	108	3.0	-0.37	0.33	1.04	0.2	1.05	0.2
103	US088	108	3.0	-0.03	0.33	0.48	-2.1	0.48	-2.1
104	US089	108	3.0	0.40	0.33	0.52	-1.9	0.52	-1.8
105	US090	108	3.3	-0.39	0.33	0.52	-1.9	0.52	-1.9
106	US091	108	2.2	-0.47	0.32	0.99	0	0.99	0
107	US092	108	3.5	-0.32	0.33	0.89	-0.3	0.90	-0.2
108	US093	108	2.8	-0.61	0.32	1.39	1.2	1.40	1.3

109	US094	108	3.7	0.84	0.33	0.50	-1.9	0.50	-1.9
110	US095	108	3.3	0.39	0.33	1.26	0.8	1.24	0.8
111	US096	108	3.1	-0.44	0.33	1.03	0.2	1.03	0.1
112	US097	108	3.6	-0.26	0.33	0.75	-0.8	0.76	-0.8
113	US098	108	2.8	-0.38	0.32	0.54	-1.8	0.53	-1.9
114	US099	108	2.8	-1.07	0.33	0.80	-0.6	0.81	-0.6
115	US100	108	3.0	-0.21	0.33	0.70	-1.1	0.70	-1.0
116	US101	108	2.9	-0.38	0.33	1.76	2.2	1.75	2.2
117	US102	108	3.0	0.70	0.32	0.69	-1.1	0.68	-1.2
118	US103	108	2.6	-0.49	0.32	0.83	-0.5	0.84	-0.5
119	US104	108	3.6	0.19	0.33	0.82	-0.5	0.82	-0.5
120	US105	108	2.9	0.24	0.33	0.48	-2.1	0.49	-2
121	US106	108	3.6	0.84	0.33	0.62	-1.4	0.62	-1.4
122	US107	108	3.3	0.16	0.33	0.82	-0.5	0.82	-0.5
123	US108	108	3.7	0.70	0.33	0.62	-1.4	0.62	-1.3
124	US109	108	3.2	-0.36	0.33	0.75	-0.8	0.76	-0.8
125	US110	108	3.2	-0.08	0.36	1.09	0.3	1.10	0.3
126	US111	108	3.4	0.27	0.33	0.46	-2.1	0.46	-2.1
127	US112	108	3.3	0.60	0.33	0.99	0	0.99	0
128	US113	108	3.4	0.14	0.33	0.76	-0.7	0.77	-0.7
129	US114	108	3.1	0.29	0.33	0.76	-0.7	0.76	-0.7
130	US115	108	3.2	0.22	0.33	0.58	-1.6	0.58	-1.6
131	US116	108	3.1	-0.02	0.33	0.76	-0.7	0.77	-0.7
132	US117	108	2.3	0.31	0.32	0.36	-3.0	0.35	-3.1
133	US118	108	2.3	-1.91	0.32	0.61	-1.5	0.61	-1.5
134	US119	108	3.0	0.18	0.33	0.42	-2.5	0.41	-2.5
135	US120	108	2.8	0.31	0.32	0.57	-1.6	0.57	-1.6
136	US121	108	2.1	-0.91	0.32	0.81	-0.6	0.81	-0.6
137	US122	108	2.6	0.06	0.32	0.40	-2.6	0.40	-2.6
138	US123	108	2.5	-0.98	0.35	0.28	-3.3	0.28	-3.2
139	US124	108	3.2	0.79	0.33	0.33	-3.0	0.33	-3.0
140	US125	108	2.3	-0.30	0.32	0.82	-0.6	0.82	-0.5
141	US126	108	2.5	-0.37	0.32	0.62	-1.4	0.61	-1.5
142	US127	108	2.7	0.69	0.33	0.72	-1.0	0.71	-1.0
143	US128	108	3.1	-0.15	0.33	0.28	-3.4	0.28	-3.4
144	US129	108	3.5	0.63	0.41	0.83	-0.4	0.83	-0.4
145	US130	108	2.4	-0.02	0.35	0.17	-4.2	0.17	-4.2
146	US131	108	4.6	1.24	0.41	0.49	-1.6	0.51	-1.6

147	US132	108	4.1	1.33	0.32	0.92	-0.1	0.94	-0.1
148	US133	108	4.8	1.43	0.32	1.15	0.6	1.14	0.5
149	US134	108	4.0	0.93	0.33	0.63	-1.3	0.63	-1.3
150	US135	108	5.3	2.64	0.34	1.81	2.5	1.96	2.9
151	US136	108	5.5	2.10	0.37	1.18	0.6	1.05	0.2
152	US137	108	4.9	1.78	0.32	1.44	1.5	1.47	1.6
153	US138	108	3.8	0.71	0.33	0.55	-1.7	0.55	-1.7
154	US139	108	4.5	1.40	0.32	1.23	0.8	1.19	0.7
155	US140	108	4.3	1.43	0.32	0.53	-1.9	0.52	-1.9
156	US141	108	4.1	1.24	0.33	0.50	-2.0	0.49	-2.0
157	US142	108	4.5	1.41	0.32	0.46	-2.4	0.47	-2.3
158	US143	108	4.6	1.02	0.33	0.75	-0.9	0.75	-0.8
159	US144	108	5.4	2.50	0.35	0.81	-0.6	0.85	-0.5
160	US145	108	4.6	1.40	0.32	0.83	-0.5	0.94	-0.1
161	US146	108	4.5	1.18	0.31	0.56	-1.8	0.57	-1.7
162	US147	108	4.4	1.20	0.36	1.27	0.9	1.27	0.8
163	US148	108	5.0	1.84	0.32	0.72	-1.0	0.74	-1.0
164	US149	108	5.5	2.80	0.42	1.20	0.6	1.13	0.4
165	US150	108	5.4	2.42	0.35	0.49	-2.3	0.52	-2.1
166	US151	108	4.6	1.28	0.32	0.99	0	0.98	0
167	US152	108	5.0	1.58	0.36	0.97	0	0.97	0
168	US153	108	2.3	0.08	0.43	0.49	-1.6	0.60	-1.1
169	US154	108	3.3	-0.12	0.36	0.98	0	0.97	0
170	US155	108	3.0	0.06	0.33	0.76	-0.8	0.76	-0.7
171	US156	108	3.3	0.76	0.33	0.44	-2.3	0.44	-2.3
172	US157	108	3.4	0.69	0.33	0.69	-1.1	0.68	-1.1
173	US158	108	3.5	0	0.36	0.40	-2.3	0.40	-2.3
174	US159	108	3.4	-0.16	0.33	0.66	-1.2	0.65	-1.2
175	US160	108	2.4	-0.14	0.32	0.31	-3.3	0.31	-3.4
176	US161	108	3.4	0.08	0.36	1.57	1.6	1.60	1.6
177	US162	108	4.0	0.78	0.33	0.96	0	0.96	0
178	US163	108	4.1	1.20	0.36	0.70	-0.8	0.71	-0.8
179	US164	108	2.1	-0.71	0.33	0.72	-1.0	0.72	-1.0
180	US165	108	2.5	-0.25	0.34	0.89	-0.2	0.86	-0.4
181	US166	108	1.8	-0.02	0.41	0.82	-0.4	0.83	-0.4
182	US167	108	3.1	0.78	0.32	0.55	-1.8	0.55	-1.7
183	US168	108	3.5	-0.56	0.33	0.67	-1.2	0.67	-1.2
184	US169	108	2.4	-0.45	0.36	0.68	-1.1	0.68	-1.1

185	US170	108	2.5	-0.67	0.32	0.62	-1.4	0.61	-1.5
186	US171	108	3.1	-0.48	0.33	0.90	-0.2	0.89	-0.2
187	US172	108	3.2	0.49	0.33	0.96	0	0.97	0
188	US173	108	3.5	0.91	0.33	0.78	-0.7	0.78	-0.7
189	US174	108	3.2	0.31	0.33	0.42	-2.5	0.42	-2.5
190	US175	108	4.0	0.96	0.36	0.55	-1.5	0.54	-1.5
191	US176	108	4.1	0.45	0.33	0.45	-2.2	0.46	-2.2
192	US177	108	3.8	1.03	0.33	0.81	-0.6	0.82	-0.5
193	US178	108	4.2	1.23	0.33	0.91	-0.2	0.91	-0.2
194	US179	108	4.0	1.85	0.36	0.79	-0.5	0.79	-0.6
195	US180	108	3.5	1.00	0.33	1.28	0.9	1.30	1.0
196	US181	108	3.3	0.25	0.33	0.89	-0.2	0.90	-0.2
197	US182	108	3.3	0.57	0.36	1.27	0.8	1.26	0.8
198	US183	108	5.0	1.76	0.33	1.30	1.1	1.29	1.0
199	US184	108	3.3	0.35	0.33	0.77	-0.7	0.77	-0.7
200	US185	108	3.6	0.65	0.36	0.78	-0.6	0.78	-0.6
201	US186	108	4.2	1.35	0.33	1.45	1.4	1.46	1.4
202	US187	108	3.4	0.07	0.33	0.59	-1.5	0.59	-1.5
203	US188	108	3.3	0.74	0.33	1.61	1.8	1.62	1.9
204	US189	108	3.9	0.71	0.33	1.36	1.2	1.35	1.2
205	US190	108	3.9	1.16	0.36	0.47	-1.9	0.46	-1.9
206	US191	108	3.6	0.56	0.33	0.62	-1.3	0.62	-1.4
207	US192	108	2.8	0.01	0.33	0.54	-1.8	0.54	-1.8
208	US193	108	3.0	0.33	0.33	0.67	-1.1	0.67	-1.1
209	US194	108	3.6	0.48	0.33	0.47	-2.1	0.47	-2.1
210	US195	108	3.6	-0.36	0.36	1.82	2.2	1.81	2.1
211	US196	108	4.2	0.81	0.33	1.03	0.1	1.04	0.2
212	US197	108	3.6	-0.43	0.33	0.86	-0.4	0.86	-0.3
213	US198	108	3.7	0.34	0.33	0.36	-2.8	0.35	-2.8
214	US199	108	3.5	0.14	0.33	0.73	-0.9	0.73	-0.9
215	US200	108	3.1	-0.33	0.33	0.37	-2.7	0.37	-2.7
216	US201	108	3.9	1.06	0.37	1.02	0.1	1.00	0.1
217	US202	108	3.4	-0.21	0.33	0.59	-1.5	0.59	-1.5
218	US203	108	3.9	0.68	0.33	0.87	-0.3	0.88	-0.3
219	US204	108	3.4	0.60	0.33	0.58	-1.6	0.58	-1.6
220	US205	108	4.0	0.39	0.33	0.65	-1.3	0.64	-1.3
221	US206	108	4.3	1.21	0.33	1.13	0.5	1.13	0.5
222	US207	108	4.1	2.10	0.33	1.04	0.2	1.03	0.2

223	US208	108	4.8	1.51	0.32	1.14	0.5	1.12	0.5
224	US209	108	4.0	0.11	0.33	0.98	0	0.98	0
225	US210	108	3.8	0.85	0.37	0.98	0	0.98	0
226	US211	104	4.0	-0.30	0.33	0.81	-0.5	0.81	-0.5
227	US212	104	3.3	-0.88	0.30	1.43	1.5	1.43	1.5
228	US212	304	3.0	-1.23	0.27	1.81	2.8	1.82	2.8
229	US213	104	4.1	0.26	0.40	0.52	-1.5	0.53	-1.4
230	US214	104	2.9	-1.75	0.32	1.66	2.0	1.66	2.0
231	US215	104	3.7	0.17	0.30	0.85	-0.5	0.84	-0.5
232	US215	304	3.5	-0.40	0.27	0.99	0	0.99	0
233	US216	104	3.7	0.41	0.33	0.58	-1.6	0.58	-1.5
234	US217	104	3.5	-0.41	0.33	1.17	0.6	1.17	0.6
235	US218	104	4.2	0.65	0.29	1.33	1.2	1.32	1.1
236	US218	304	4.2	0.72	0.27	1.57	2.1	1.58	2.1
237	US219	104	4.3	0.88	0.33	0.38	-2.6	0.38	-2.6
238	US220	104	3.6	-0.32	0.33	0.86	-0.4	0.87	-0.3
239	US221	104	2.5	-1.64	0.32	1.11	0.4	1.11	0.4
240	US222	104	4.1	-0.08	0.32	0.56	-1.7	0.54	-1.8
241	US223	104	3.8	0.07	0.33	0.69	-1.0	0.68	-1.1
242	US224	104	2.4	-0.86	0.39	2.16	2.7	2.16	2.7
243	US225	104	3.3	-0.98	0.33	0.83	-0.5	0.84	-0.4
244	US226	104	4.3	0.22	0.32	0.63	-1.4	0.65	-1.3
245	US227	104	3.1	-0.79	0.32	0.84	-0.5	0.83	-0.5
246	US227	304	2.6	-1.28	0.31	0.82	-0.6	0.82	-0.6
247	US228	104	3.0	-1.21	0.33	0.79	-0.6	0.79	-0.6
248	US229	104	3.5	-0.43	0.40	0.89	-0.1	0.89	-0.1
249	US230	106	3.0	-0.63	0.33	1.16	0.6	1.18	0.7
250	US231	106	3.1	-0.53	0.33	0.58	-1.6	0.56	-1.6
251	US232	106	3.3	0.11	0.33	0.51	-1.9	0.50	-1.9
252	US233	106	2.7	-0.68	0.32	0.58	-1.6	0.58	-1.6
253	US234	106	3.0	-0.72	0.33	0.54	-1.8	0.54	-1.8
254	US235	106	2.8	0.28	0.29	1.28	1.0	1.27	1.0
255	US235	306	2.6	-0.27	0.26	1.49	1.9	1.47	1.8
256	US236	106	2.4	-1.09	0.29	0.99	0	0.97	0
257	US236	306	2.6	-1.47	0.26	1.19	0.8	1.17	0.7
258	US237	106	3.3	-0.68	0.29	1.69	2.2	1.72	2.3
259	US237	306	3.4	-1.19	0.27	2.11	3.6	2.12	3.7
260	US238	106	3.3	-0.27	0.33	0.74	-0.8	0.73	-0.9

261	US239	106	3.5	-0.01	0.33	0.97	0	0.95	0
262	US240	106	3.1	0.17	0.33	0.51	-1.9	0.51	-1.9
263	US241	106	3.7	0.03	0.29	1.27	1.0	1.27	1.0
264	US241	306	4.3	0.67	0.27	0.73	-1.1	0.75	-1.0
265	US242	106	3.7	0.67	0.33	1.06	0.2	1.09	0.3
266	US243	106	4.0	0.79	0.29	1.33	1.2	1.33	1.2
267	US243	306	4.4	1.07	0.26	1.55	2.0	1.54	2.0
268	US244	106	3.8	0.67	0.33	0.70	-1.1	0.70	-1.0
269	US245	106	2.4	-0.87	0.33	0.62	-1.4	0.63	-1.4
270	US246	106	3.0	-1.15	0.29	0.85	-0.5	0.87	-0.4
271	US246	306	3.6	-1.15	0.27	1.22	0.9	1.23	0.9
272	US247	106	3.7	0.49	0.33	0.90	-0.2	0.91	-0.2
273	US248	106	4.0	0.72	0.29	2.20	3.6	2.15	3.5
274	US248	306	4.9	1.77	0.27	1.82	3.0	1.75	2.8
275	US249	106	3.7	0.10	0.33	0.97	0	0.96	0
276	US250	104	3.4	0.04	0.33	0.43	-2.4	0.43	-2.4
277	US251	104	3.2	-1.20	0.33	0.72	-0.9	0.73	-0.9
278	US252	104	3.4	-0.93	0.33	1.11	0.4	1.11	0.4
279	US253	104	3.1	-0.83	0.33	0.93	-0.1	0.93	-0.1
280	US254	104	3.9	0.81	0.33	0.50	-2.0	0.50	-2.0
281	US255	104	2.7	-1.16	0.32	0.47	-2.2	0.46	-2.2
282	US256	104	1.8	-2.31	0.34	1.37	1.3	1.34	1.2
283	US256	304	1.8	-2.22	0.35	1.95	2.8	1.88	2.7
284	US257	104	3.4	-1.65	0.33	0.68	-1.1	0.68	-1.1
285	US258	104	3.5	0.04	0.33	1.10	0.4	1.09	0.4
286	US259	104	4.4	0.42	0.32	0.49	-2.1	0.47	-2.2
287	US260	104	3.0	-1.36	0.33	1.49	1.5	1.48	1.5
288	US261	104	4.1	0.04	0.33	0.80	-0.6	0.80	-0.5
289	US262	104	2.2	-2.39	0.35	0.67	-1.1	0.66	-1.2
290	US263	104	2.1	-2.06	0.32	0.81	-0.6	0.81	-0.6
291	US264	104	2.9	-0.74	0.32	0.82	-0.5	0.83	-0.5
292	US265	104	3.9	-0.33	0.33	1.32	1.1	1.30	1.0
293	US266	104	2.4	-1.59	0.29	1.53	1.9	1.53	1.9
294	US266	304	2.1	-2.13	0.26	1.76	2.9	1.74	2.8
295	US267	104	2.5	-1.03	0.32	1.06	0.2	1.03	0.2
296	US268	104	2.2	-2.20	0.33	0.89	-0.3	0.88	-0.3
297	US269	104	5.0	1.13	0.31	1.70	2.3	1.70	2.3
298	US270	104	3.9	0.03	0.36	1.50	1.4	1.50	1.4

299	US271	104	3.2	-1.43	0.36	4.30	5.9	4.34	6.0
300	US272	104	4.5	-0.95	0.32	1.54	1.7	1.53	1.6
301	US273	104	4.1	0.57	0.29	0.64	-1.4	0.62	-1.5
302	US273	304	4.1	0.47	0.27	0.84	-0.6	0.82	-0.7
303	US274	104	3.6	-0.01	0.33	0.74	-0.8	0.74	-0.8
304	US275	104	4.7	1.02	0.31	0.42	-2.7	0.42	-2.7
305	US276	104	3.9	0.12	0.33	0.91	-0.2	0.89	-0.2
306	US277	104	4.4	0.85	0.32	0.61	-1.5	0.62	-1.4
307	US278	104	4.8	1.19	0.31	1.45	1.5	1.42	1.4
308	US279	104	3.5	0.69	0.33	0.41	-2.5	0.41	-2.5
309	US280	104	4.4	0.02	0.39	1.45	1.2	1.43	1.1
310	US281	104	4.2	-0.34	0.33	1.88	2.4	1.94	2.5
311	US282	104	4.1	0.54	0.33	0.37	-2.7	0.36	-2.8
312	US283	104	3.8	-0.58	0.33	1.73	2.1	1.75	2.1
313	US284	104	4.1	-0.42	0.33	1.15	0.5	1.16	0.6
314	US285	104	1.2	-5.66	0.59	1.69	1.4	3.12	2.0
315	US285	304	1.1	-4.98	0.76	1.18	0.4	1.68	0.9
316	US286	104	4.3	-0.07	0.32	0.93	-0.1	0.91	-0.2
317	US287	104	4.3	0.32	0.32	0.73	-0.9	0.73	-0.9
318	US288	104	4.5	0.66	0.32	1.39	1.3	1.39	1.3
319	US289	104	3.0	-0.39	0.33	0.64	-1.3	0.64	-1.3
320	US290	104	3.4	-1.27	0.41	0.43	-1.9	0.43	-1.9
321	US291	106	3.3	0.08	0.33	1.18	0.6	1.18	0.6
322	US292	106	5.2	1.66	0.35	1.79	2.3	1.50	1.4
323	US293	106	4.3	1.06	0.32	0.65	-1.3	0.64	-1.4
324	US294	106	4.4	0.88	0.35	1.04	0.2	1.03	0.1
325	US295	106	3.1	-0.29	0.33	0.64	-1.3	0.64	-1.3
326	US296	106	4.2	0.54	0.32	1.20	0.7	1.18	0.7
327	US297	106	3.9	0.45	0.33	1.12	0.4	1.12	0.4
328	US298	106	4.7	0.78	0.32	1.07	0.3	1.06	0.2
329	US299	106	3.8	0.21	0.35	2.04	2.6	2.16	2.9
330	US300	106	4.4	1.63	0.32	0.82	-0.5	0.82	-0.5
331	US301	106	3.5	-0.52	0.33	1.08	0.3	1.08	0.3
332	US302	106	5.2	1.64	0.37	0.78	-0.6	0.69	-0.8
333	US303	106	4.1	-0.04	0.32	0.70	-1.0	0.68	-1.1
334	US304	106	5.0	1.91	0.35	0.89	-0.2	0.91	-0.2
335	US305	106	4.5	0.74	0.36	0.78	-0.6	0.82	-0.5
336	US306	106	3.0	-0.81	0.40	2.19	2.6	2.26	2.7

337	US307	110	4.8	2.06	0.31	0.46	-2.4	0.46	-2.5
338	US308	110	3.9	0.61	0.33	0.56	-1.6	0.54	-1.7
339	US309	110	4.1	1.49	0.33	0.46	-2.2	0.45	-2.2
340	US310	110	3.8	0.95	0.33	0.84	-0.4	0.83	-0.5
341	US311	110	4.9	1.38	0.33	0.89	-0.3	0.88	-0.3
342	US312	110	5.0	2.49	0.32	1.03	0.2	1.02	0.1
343	US313	110	4.7	2.09	0.31	0.69	-1.2	0.68	-1.2
344	US314	110	4.5	1.46	0.32	0.65	-1.3	0.63	-1.4
345	US315	110	5.0	1.85	0.32	1.01	0.1	0.99	0
346	US316	110	4.0	1.22	0.33	0.85	-0.4	0.83	-0.4
347	US317	110	4.4	1.22	0.33	1.02	0.1	1.00	0
348	US318	110	4.1	0.98	0.33	1.31	1.0	1.32	1.0
349	US319	110	4.8	1.67	0.31	0.78	-0.8	0.76	-0.8
350	US320	110	5.3	1.75	0.34	0.97	0	0.99	0
351	US321	110	3.1	0.24	0.33	0.49	-2.1	0.49	-2.0
352	US322	110	3.0	0.06	0.33	0.95	0	0.94	-0.1
353	US323	110	2.8	0.07	0.33	1.06	0.2	1.07	0.3
354	US324	110	4.3	1.01	0.32	0.59	-1.5	0.60	-1.5
355	US325	110	3.3	0.10	0.33	0.45	-2.3	0.44	-2.3
356	US326	110	3.1	-0.32	0.36	1.19	0.6	1.20	0.7
357	US327	110	3.1	-0.65	0.33	1.10	0.4	1.10	0.4
358	US328	110	3.9	1.04	0.33	0.61	-1.4	0.61	-1.4
359	US329	110	3.0	-0.44	0.33	0.81	-0.6	0.81	-0.6
360	US330	110	3.2	0.60	0.36	0.06	-5.2	0.06	-5.3
361	US331	110	4.0	0.60	0.33	1.14	0.5	1.11	0.4
362	US332	110	2.6	-1.05	0.39	0.78	-0.5	0.77	-0.5
363	US333	110	3.3	0.62	0.33	0.71	-1.0	0.71	-0.9
364	US334	110	4.0	0.41	0.33	1.04	0.2	1.03	0.1
365	US335	110	4.9	1.31	0.31	1.06	0.3	1.06	0.2
366	US336	110	4.3	0.93	0.33	0.62	-1.4	0.63	-1.3
367	US337	110	3.4	0.59	0.33	0.60	-1.4	0.60	-1.4
368	US338	110	3.5	-0.03	0.33	0.63	-1.3	0.63	-1.3
369	US339	110	4.5	1.14	0.32	0.82	-0.5	0.84	-0.5
370	US340	110	3.5	-0.09	0.33	1.01	0.1	1.01	0.1
371	US341	110	3.3	-0.11	0.33	1.29	1.0	1.29	1.0
372	US342	110	4.8	2.12	0.34	0.89	-0.3	0.90	-0.2
373	US343	110	4.5	1.19	0.32	1.53	1.7	1.49	1.5
374	US344	110	4.8	1.59	0.32	1.40	1.4	1.49	1.6

375	US345	110	4.7	1.55	0.32	0.63	-1.5	0.65	-1.4
376	US346	110	4.1	0.68	0.33	0.94	-0.1	0.94	-0.1
377	US347	110	4.1	1.29	0.32	0.63	-1.3	0.64	-1.3
378	US348	110	4.1	1.17	0.33	0.62	-1.3	0.61	-1.4
379	US349	110	3.8	0.60	0.33	1.32	1.1	1.31	1.0
380	US350	110	4.0	1.09	0.33	0.74	-0.9	0.73	-0.9
381	US351	110	4.6	1.86	0.32	1.20	0.7	1.22	0.8
382	US352	110	3.4	-0.31	0.33	1.18	0.6	1.18	0.6
383	US353	110	4.8	2.05	0.31	1.30	1.1	1.31	1.1
384	US354	110	4.1	0.86	0.33	1.14	0.5	1.15	0.5
385	US355	110	3.9	0.49	0.33	1.06	0.3	1.07	0.3
386	US356	110	4.5	1.20	0.32	1.34	1.1	1.37	1.2
387	US357	110	4.1	0.82	0.32	1.72	2.1	1.73	2.1
388	US358	110	3.9	0.65	0.33	0.68	-1.0	0.68	-1.0
389	US359	110	4.1	0.44	0.33	1.13	0.5	1.14	0.5
390	US360	110	3.4	-0.83	0.33	1.47	1.4	1.48	1.5
391	US361	110	3.9	0.80	0.33	1.11	0.4	1.13	0.5
392	US362	110	3.3	-0.70	0.33	1.95	2.6	1.93	2.6
393	US363	110	3.6	0.90	0.40	0.77	-0.5	0.78	-0.5
394	US364	110	3.0	-0.36	0.33	0.75	-0.8	0.75	-0.8
395	US365	110	2.2	-0.06	0.41	0.90	-0.2	0.88	-0.2
396	US366	110	3.9	0.31	0.33	2.53	3.6	2.49	3.5
397	US367	110	3.5	0.49	0.33	0.65	-1.3	0.65	-1.3
398	US368	110	4.3	1.37	0.32	0.88	-0.3	0.89	-0.2
399	US369	110	3.1	0.16	0.33	0.60	-1.5	0.61	-1.4
400	US370	110	3.6	0.45	0.33	1.27	0.9	1.25	0.8
401	US371	110	4.2	1.49	0.33	0.59	-1.5	0.57	-1.6
402	US372	110	3.2	0.62	0.32	0.91	-0.2	0.89	-0.2
403	US373	110	3.3	-0.22	0.33	0.78	-0.7	0.79	-0.7
404	US374	110	2.8	0.12	0.33	0.51	-2.0	0.51	-2.0
405	US375	110	3.0	-0.16	0.33	0.39	-2.7	0.39	-2.6
406	US376	110	2.6	-1.13	0.38	0.80	-0.5	0.75	-0.6
407	US377	110	3.3	0.65	0.33	0.61	-1.4	0.62	-1.3
408	US378	110	3.3	-0.05	0.33	0.74	-0.8	0.74	-0.8
409	US379	110	3.3	0.35	0.33	0.71	-1.0	0.71	-0.9
410	US380	110	4.3	1.65	0.32	1.02	0.1	0.99	0
411	US381	110	3.3	-0.25	0.33	0.59	-1.5	0.59	-1.5
412	US382	110	3.1	-0.10	0.33	0.68	-1.2	0.67	-1.2

413	US383	110	3.4	0.30	0.36	1.07	0.3	1.06	0.2
414	US384	110	3.7	0.79	0.33	0.48	-2.1	0.48	-2.1
415	US385	110	4.4	1.54	0.32	0.90	-0.2	0.87	-0.3
416	US386	110	4.0	1.11	0.33	0.91	-0.2	0.93	-0.1
417	US387	110	4.6	1.57	0.32	1.44	1.5	1.44	1.5
418	US388	110	3.8	0.97	0.33	1.03	0.1	1.03	0.1
419	US389	110	4.6	1.95	0.32	0.87	-0.3	0.87	-0.3
420	US390	110	4.1	0.63	0.33	1.12	0.5	1.13	0.5
421	US391	110	4.1	0.52	0.32	0.85	-0.4	0.87	-0.4
422	US392	110	4.2	0.98	0.32	1.03	0.1	1.02	0.1
423	US393	110	4.7	2.00	0.32	1.26	0.9	1.28	1.0.
424	US394	110	3.6	0.45	0.33	1.02	0.1	1.01	0.1
425	US395	110	3.6	0.72	0.33	0.74	-0.8	0.76	-0.7
426	US396	110	5.3	2.25	0.35	0.74	-0.9	0.76	-0.7
427	US397	110	4.5	1.59	0.32	0.80	-0.6	0.80	-0.6
428	US398	110	4.0	0.12	0.33	0.66	-1.1	0.66	-1.1
429	US399	106	3.5	0.22	0.33	0.56	-1.7	0.55	-1.7
430	US400	106	3.7	0.69	0.33	0.80	-0.6	0.82	-0.5
431	US401	106	4.3	1.01	0.32	1.40	1.3	1.41	1.3
432	US402	106	3.8	0.95	0.32	0.89	-0.3	0.90	-0.2
433	US403	106	3.3	-0.15	0.32	0.45	-2.3	0.45	-2.3
434	US404	106	2.3	-0.30	0.30	0.54	-2.1	0.53	-2.1
435	US404	306	2.7	-0.54	0.26	0.77	-1.0	0.77	-0.9
436	US405	106	3.7	0.51	0.33	1.11	0.4	1.14	0.5
437	US406	106	3.4	0.18	0.33	1.01	0.1	1.02	0.1
438	US407	106	2.9	-0.29	0.33	0.59	-1.6	0.60	-1.5
439	US408	106	4.1	0.46	0.32	0.44	-2.3	0.45	-2.3
440	US409	106	4.5	1.16	0.34	0.60	-1.5	0.61	-1.4
441	US410	106	5.2	2.17	0.35	1.03	0.2	0.92	-0.1
442	US411	106	4.1	1.15	0.33	0.97	0	1.00	0.1
443	US412	106	3.4	-0.13	0.33	0.41	-2.5	0.41	-2.5
444	US413	106	2.8	-0.27	0.33	0.78	-0.7	0.79	-0.7
445	US414	106	2.9	-1.01	0.33	0.96	0	0.96	0
446	US415	106	3.5	-0.12	0.32	0.44	-2.3	0.43	-2.4
447	US416	106	4.2	0.99	0.32	0.52	-2.0	0.53	-2.0
448	US417	106	3.2	0.08	0.33	0.54	-1.8	0.53	-1.8
449	US418	104	3.1	-0.31	0.33	1.12	0.5	1.12	0.5
450	US419	104	2.9	-0.35	0.29	0.83	-0.6	0.84	-0.5

451	US419	304	2.8	-0.08	0.26	0.54	-2.3	0.54	-2.2
452	US420	104	2.2	-1.63	0.29	0.43	-3.0	0.43	-3.0
453	US420	304	2.2	-1.71	0.26	0.55	-2.3	0.55	-2.3
454	US421	104	3.6	-0.17	0.33	1.14	0.5	1.14	0.5
455	US422	104	3.8	-0.34	0.33	1.01	0.1	0.98	0
456	US423	104	3.1	0.47	0.33	1.13	0.5	1.11	0.4
457	US424	104	3.5	0.36	0.33	0.53	-1.9	0.53	-1.9
458	US425	104	3.9	0.14	0.33	0.42	-2.3	0.42	-2.4
459	US426	104	3.6	-0.06	0.33	0.85	-0.4	0.85	-0.4
460	US427	104	3.3	-0.54	0.30	1.26	0.9	1.24	0.9
461	US427	304	3.3	-0.21	0.28	1.70	2.4	1.68	2.4
462	US428	104	4.3	0.24	0.32	0.83	-0.5	0.80	-0.6
463	US429	104	2.8	-0.72	0.29	0.88	-0.4	0.86	-0.4
464	US429	304	2.6	-0.82	0.26	1.15	0.7	1.15	0.6
465	US430	104	3.0	-0.85	0.33	0.80	-0.6	0.80	-0.6
466	US431	104	2.6	-0.80	0.32	0.52	-1.9	0.53	-1.9
467	US432	104	2.2	-1.36	0.39	0.88	-0.2	0.89	-0.2
468	US433	104	3.7	-0.25	0.33	1.33	1.1	1.35	1.1
469	US434	104	3.4	-0.56	0.36	0.64	-1.1	0.64	-1.1
470	US435	104	3.0	-1.06	0.33	1.16	0.6	1.15	0.5
471	US436	104	4.4	-0.03	0.29	1.51	1.8	1.52	1.8
472	US436	304	4.4	-0.13	0.26	2.17	3.9	2.24	4.0
473	US437	104	2.3	-1.63	0.32	1.03	0.2	1.04	0.2
474	US438	106	3.5	0.44	0.33	0.91	-0.2	0.91	-0.2
475	US439	106	3.2	-0.18	0.29	0.86	-0.4	0.87	-0.4
476	US439	306	3.4	-0.60	0.27	1.08	0.3	1.09	0.4
477	US440	106	5.0	1.61	0.31	0.86	-0.4	0.91	-0.2
478	US440	306	5.4	1.47	0.30	1.06	0.3	1.06	0.3
479	US441	106	3.6	0.59	0.33	0.59	-1.5	0.58	-1.5
480	US442	106	4.4	0.27	0.33	1.09	0.4	1.05	0.2
481	US443	106	4.5	1.76	0.29	1.23	0.9	1.25	1.0
482	US443	306	4.8	2.03	0.28	1.55	2.0	1.51	1.9
483	US444	106	4.8	1.46	0.32	1.26	0.9	1.23	0.8
484	US445	106	3.4	0.71	0.33	0.78	-0.6	0.78	-0.6
485	US446	106	4.3	0.94	0.32	0.93	-0.1	0.92	-0.2
486	US447	106	4.5	0.98	0.32	1.55	1.8	1.53	1.7
487	US448	106	2.5	-1.30	0.29	1.43	1.5	1.42	1.5
488	US448	306	2.6	-2.03	0.26	1.25	1.1	1.27	1.1

489	US449	106	4.7	1.11	0.32	0.97	0	0.99	0
490	US450	106	3.4	0.27	0.30	0.84	-0.5	0.85	-0.5
491	US450	306	3.7	0.05	0.27	0.96	0	0.97	0
492	US451	106	4.0	0.24	0.32	1.93	2.6	1.95	2.6
493	US452	106	4.5	0.81	0.31	0.98	0	0.96	0
494	US453	106	4.0	1.33	0.33	0.67	-1.1	0.68	-1.0
495	US454	106	4.1	-0.69	0.29	1.46	1.6	1.48	1.6
496	US454	306	4.3	-0.85	0.26	1.81	2.8	1.88	3.0
497	US455	106	2.9	-0.22	0.33	0.83	-0.5	0.83	-0.5
498	US456	106	3.4	-0.30	0.33	1.45	1.4	1.48	1.5
499	US457	106	3.7	0.18	0.36	0.92	-0.1	0.93	-0.1
500	US458	106	3.9	0.53	0.33	0.96	0	0.96	0
501	US459	106	2.8	-0.12	0.29	1.56	1.9	1.60	2.0
502	US459	306	3.1	-0.23	0.27	2.23	3.9	2.25	3.9
503	US460	106	3.1	0.14	0.30	1.69	2.3	1.67	2.2
504	US460	306	3.4	0.03	0.27	2.26	4.0	2.28	4.1
505	US461	106	3.3	0.82	0.33	0.89	-0.2	0.91	-0.2
506	US462	106	3.5	-0.40	0.30	1.19	0.7	1.19	0.7
507	US462	306	4.2	0.21	0.27	1.07	0.3	1.09	0.4
508	US463	106	3.7	0.31	0.30	1.29	1.1	1.29	1.0
509	US463	306	4.3	1.01	0.27	0.73	-1.1	0.72	-1.2
510	US464	106	3.9	0.83	0.33	1.40	1.3	1.39	1.2
511	US465	106	2.3	-2.37	0.34	2.96	4.5	2.97	4.1
512	US466	106	4.2	0.77	0.29	1.49	1.8	1.49	1.8
513	US466	306	4.6	0.54	0.26	2.00	3.5	2.05	3.7
514	US467	106	3.6	-0.32	0.33	0.73	-0.9	0.73	-0.9
515	US468	106	3.8	0.45	0.33	0.8	-0.6	0.81	-0.5
516	US469	106	3.5	-0.45	0.33	0.84	-0.4	0.84	-0.4
517	US470	106	4.0	1.34	0.29	1.04	0.2	1.04	0.2
518	US470	306	4.4	1.31	0.26	1.37	1.5	1.40	1.5
519	US471	106	4.2	0.94	0.32	0.89	-0.3	0.88	-0.3
520	US472	106	3.5	0.55	0.33	0.69	-1.0	0.70	-1.0
521	US473	112	4.0	0.66	0.32	0.49	-2.0	0.49	-2.0
522	US474	112	3.9	0.77	0.32	0.54	-1.8	0.54	-1.8
523	US475	112	3.8	1.09	0.33	0.64	-1.3	0.64	-1.2
524	US476	112	3.9	0.99	0.33	0.75	-0.8	0.73	-0.9
525	US477	112	4.1	1.49	0.33	0.84	-0.4	0.82	-0.5
526	US478	112	4.1	1.18	0.33	1.20	0.7	1.21	0.7

527	US479	112	4.2	1.32	0.33	0.41	-2.5	0.41	-2.5
528	US480	112	3.6	0.27	0.33	0.79	-0.6	0.77	-0.7
529	US481	112	4.3	0.86	0.32	0.68	-1.1	0.67	-1.2
530	US482	112	4.1	1.25	0.32	0.82	-0.5	0.81	-0.6
531	US483	112	4.5	1.04	0.32	1.04	0.2	1.00	0.1
532	US484	112	3.7	0.81	0.33	0.92	-0.1	0.92	-0.1
533	US485	112	3.8	1.04	0.33	0.78	-0.6	0.78	-0.7
534	US486	112	4.5	1.39	0.32	0.70	-1.0	0.68	-1.1
535	US487	112	3.8	0.53	0.37	0.76	-0.6	0.76	-0.6
536	US488	112	4.2	1.04	0.33	0.70	-1.0	0.67	-1.1
537	US489	112	4.7	2.09	0.31	0.84	-0.5	0.82	-0.6
538	US490	112	3.8	0.73	0.33	1.16	0.6	1.15	0.5
539	US491	112	4.0	0.95	0.33	1.19	0.7	1.21	0.7
540	US492	112	2.8	0.35	0.33	0.85	-0.4	0.86	-0.4
541	US493	112	4.1	0.98	0.33	1.00	0	0.99	0
542	US494	112	3.4	-0.30	0.33	0.92	-0.1	0.93	-0.1
543	US495	112	3.5	0.88	0.33	0.76	-0.8	0.75	-0.8
544	US496	112	3.3	0.36	0.33	0.84	-0.4	0.85	-0.4
545	US497	112	3.5	0.19	0.33	0.68	-1.1	0.68	-1.1
546	US498	112	3.2	0.55	0.33	0.62	-1.4	0.61	-1.4
547	US499	112	2.4	-0.18	0.32	0.91	-0.2	0.92	-0.2
548	US500	112	3.3	0.08	0.33	0.61	-1.4	0.62	-1.4
549	US501	112	3.9	-0.02	0.33	1.27	0.9	1.28	0.9
550	US502	112	2.6	0.13	0.32	0.37	-2.9	0.37	-2.9
551	US503	112	3.6	1.01	0.33	0.61	-1.4	0.61	-1.4
552	US504	112	2.1	-0.70	0.33	0.47	-2.3	0.47	-2.3
553	US505	112	2.3	-0.13	0.35	0.86	-0.4	0.85	-0.4
554	US506	112	2.8	-0.37	0.32	1.19	0.7	1.21	0.8
555	US507	112	3.3	0.29	0.33	0.48	-2.0	0.48	-2.0
556	US508	112	3.4	0.37	0.33	0.71	-1.0	0.72	-1.0
557	US509	112	3.4	0.36	0.33	0.84	-0.5	0.83	-0.5
558	US510	112	3.8	0.77	0.33	0.69	-1.1	0.69	-1.1
559	US511	112	2.9	0.30	0.36	0.90	-0.2	0.90	-0.2
560	US512	112	3.5	0.90	0.33	0.85	-0.4	0.86	-0.4
561	US513	112	3.7	0.02	0.32	1.51	1.6	1.50	1.6
562	US514	112	3.4	1.16	0.33	0.68	-1.0	0.69	-1.0
563	US515	112	2.9	0.36	0.33	0.58	-1.6	0.58	-1.6
564	US516	112	3.1	0.55	0.32	1.01	0.1	1.00	0

565	US517	112	5.1	2.45	0.33	0.73	-1.0	0.79	-0.7
566	US518	112	3.8	1.39	0.33	1.14	0.5	1.14	0.5
567	US519	112	3.8	1.19	0.33	0.73	-0.9	0.73	-0.9
568	US520	112	3.3	0.83	0.33	0.53	-1.8	0.52	-1.9
569	US521	112	4.4	1.33	0.32	0.94	-0.1	0.92	-0.1
570	US522	112	4.1	1.15	0.33	0.45	-2.2	0.46	-2.2
571	US523	112	5.6	3.61	0.41	0.58	-1.4	0.73	-0.7
572	US524	112	4.3	1.63	0.33	0.63	-1.3	0.64	-1.3
573	US525	112	4.5	2.02	0.32	0.63	-1.4	0.64	-1.3
574	US526	112	4.3	1.11	0.32	0.65	-1.3	0.64	-1.4
575	US527	112	4.2	1.32	0.33	0.85	-0.4	0.85	-0.4
576	US528	112	4.3	1.61	0.32	0.75	-0.8	0.76	-0.8
577	US529	112	3.8	0.89	0.33	0.51	-1.9	0.50	-2.0
578	US530	112	3.9	0.25	0.33	0.88	-0.3	0.90	-0.2
579	US531	112	4.6	1.79	0.32	0.95	-0.1	0.95	0
580	US532	112	4.1	1.66	0.33	0.89	-0.2	0.90	-0.2
581	US533	112	5.2	2.87	0.34	1.04	0.2	1.09	0.3
582	US534	112	4.3	0.99	0.33	1.07	0.3	1.05	0.2
583	US535	112	3.8	0.73	0.33	0.84	-0.4	0.84	-0.4
584	US536	112	4.2	0.83	0.32	0.92	-0.1	0.91	-0.2
585	US537	112	4.0	1.19	0.33	0.83	-0.5	0.86	-0.4
586	US538	112	3.5	0.73	0.33	0.69	-1.1	0.69	-1.1
587	US539	112	4.9	1.54	0.32	1.36	1.3	1.35	1.2
588	US540	112	5.2	2.63	0.32	1.18	0.7	1.22	0.8
589	US541	112	4.0	1.36	0.33	0.84	-0.5	0.86	-0.4
590	US542	112	3.9	0.98	0.33	0.62	-1.3	0.63	-1.3
591	US543	112	4.4	1.64	0.32	0.61	-1.5	0.62	-1.4
592	US544	112	4.6	1.71	0.32	0.73	-0.9	0.75	-0.8
593	US545	112	4.1	0.49	0.32	0.74	-0.9	0.77	-0.8
594	US546	112	4.1	1.13	0.33	0.55	-1.6	0.55	-1.6
595	US547	112	3.8	0.54	0.33	0.51	-1.9	0.51	-1.9
596	US548	112	4.3	0.58	0.33	0.42	-2.4	0.42	-2.4
597	US549	112	3.3	0.06	0.33	0.34	-2.9	0.34	-2.9
598	US550	112	3.9	0.77	0.37	0.77	-0.6	0.77	-0.6
599	US551	112	3.5	0.03	0.33	1.12	0.4	1.11	0.4
600	US552	112	4.5	1.71	0.32	0.63	-1.4	0.65	-1.3
601	US553	112	4.4	1.94	0.32	0.72	-0.9	0.72	-0.9
602	US554	112	3.6	0.61	0.33	0.80	-0.6	0.81	-0.6

603	US555	112	4.9	1.95	0.34	0.51	-2.0	0.51	-2.0
604	US556	112	3.9	0.74	0.33	0.46	-2.1	0.46	-2.1
605	US557	112	3.5	0.80	0.33	0.60	-1.5	0.60	-1.5
606	US558	112	2.5	-0.69	0.33	0.90	-0.2	0.93	-0.1
607	US559	112	3.1	0.65	0.33	0.58	-1.5	0.57	-1.5
608	US560	112	3.3	-0.09	0.33	0.39	-2.6	0.39	-2.6
609	US561	112	2.2	-0.50	0.32	0.39	-2.8	0.39	-2.8
610	US562	112	3.3	0.74	0.33	0.67	-1.1	0.67	-1.1
611	US563	112	2.5	-0.33	0.32	0.74	-0.9	0.74	-0.8
612	US564	112	3.5	1.11	0.33	0.57	-1.6	0.57	-1.6
613	US565	112	3.9	0.86	0.33	0.67	-1.1	0.68	-1.1
614	US566	112	3.2	0.12	0.33	0.61	-1.4	0.60	-1.4
615	US567	112	3.8	0.45	0.33	1.12	0.4	1.11	0.4
616	US568	112	2.3	-1.03	0.32	1.07	0.3	1.09	0.4
617	US569	112	2.9	-0.35	0.33	0.75	-0.8	0.75	-0.8
618	US570	112	2.5	-0.89	0.32	0.95	0	0.95	0
619	US571	104	4.7	1.15	0.31	0.79	-0.7	0.79	-0.7
620	US572	104	3.2	-0.89	0.33	2.90	4.5	2.90	4.5
621	US573	104	4.5	0.56	0.32	1.67	2.0	1.73	2.2
622	US574	104	3.5	-0.30	0.33	0.81	-0.6	0.81	-0.6
623	US575	104	3.8	-0.05	0.34	0.65	-1.2	0.65	-1.2
624	US576	104	4.0	0.48	0.33	0.65	-1.2	0.63	-1.3
625	US577	104	3.1	-1.24	0.31	4.11	6.6	4.19	6.7
626	US577	304	3.3	-0.48	0.29	2.00	3.1	2.00	3.1
627	US578	104	5.0	1.71	0.36	0.85	-0.4	0.84	-0.4
628	US579	104	3.8	-1.12	0.33	7.70	9.0	7.08	9.0
629	US580	104	3.5	-0.30	0.33	1.22	0.7	1.23	0.8
630	US581	104	3.1	-1.42	0.33	2.20	3.1	2.21	3.1
631	US582	104	3.9	0.20	0.36	0.65	-1.1	0.64	-1.2
632	US583	104	4.0	0.70	0.33	0.80	-0.6	0.78	-0.7
633	US584	104	3.8	-0.46	0.33	0.46	-2.2	0.46	-2.2
634	US585	104	3.5	-1.99	0.33	0.98	0	1.01	0.1
635	US586	104	2.7	-0.57	0.29	1.24	0.9	1.26	1.0
636	US586	304	2.7	-0.19	0.26	1.02	0.1	1.04	0.2
637	US587	104	2.7	-1.26	0.32	0.82	-0.5	0.80	-0.6
638	US588	104	3.5	1.03	0.37	0.42	-2.2	0.42	-2.2
639	US589	104	3.5	-0.86	0.33	0.71	-1.0	0.71	-1.0
640	HK001	206	1.3	-6.13	1.17	0.50	-0.4	0.45	-0.5

641	HK002	206	3.2	-2.26	0.66	0.99	0.1	0.97	0.1
642	HK003	206	3.8	-1.54	0.66	0.99	0.1	1.02	0.2
643	HK004	206	3.3	-1.71	0.67	0.85	0	0.84	0
644	HK005	206	1.3	-7.42	1.17	0.82	0	0.54	0
645	HK006	206	3.5	-1.97	0.67	0.64	-0.5	0.64	-0.5
646	HK007	206	5.3	2.10	0.69	0.24	-1.8	0.28	-1.6
647	HK008	206	3.2	-1.96	0.66	0.79	-0.1	0.79	-0.1
648	HK009	206	3.7	-1.58	0.67	2.14	1.6	2.15	1.6
649	HK010	206	4.8	1.15	0.63	0.60	-0.7	0.61	-0.7
650	HK011	206	2.3	-3.86	0.64	2.71	2.3	2.81	2.3
651	HK012	206	4.7	0.93	0.63	0.40	-1.2	0.41	-1.2
652	HK013	206	3.7	-1.23	0.66	0.86	0	0.86	0
653	HK014	206	4.2	-0.02	0.65	1.21	0.5	1.17	0.4
654	HK015	206	5.3	1.77	0.69	0.42	-1.2	0.52	-0.9
655	HK016	206	4.2	-0.39	0.65	0.13	-2.3	0.13	-2.3
656	HK017	206	3.0	-3.33	0.65	0.42	-1.1	0.42	-1.1
657	HK018	206	3.3	-2.21	0.67	0.16	-2.1	0.16	-2.1
658	HK019	206	3.8	-1.23	0.66	2.29	1.8	2.34	1.8
659	HK020	206	3.0	-2.69	0.66	0.49	-0.9	0.49	-0.9
660	HK021	206	3.8	-1.43	0.66	9.00	5.8	8.79	5.7
661	HK022	206	1.6	-5.48	0.78	0.52	-0.7	0.61	-0.5
662	HK023	206	4.0	-1.25	0.65	1.69	1.1	1.63	1.1
663	HK024	206	4.8	1.53	0.63	1.03	0.2	1.07	0.2
664	HK025	206	4.7	0.83	0.64	0.73	-0.3	0.73	-0.3
665	HK026	206	3.7	-1.03	0.67	0.36	-1.2	0.36	-1.2
666	HK027	206	2.8	-3.02	0.65	2.28	1.8	2.33	1.8
667	HK028	206	3.2	-2.60	0.66	0.33	-1.3	0.32	-1.4
668	HK029	206	3.0	-2.59	0.66	0.88	0	0.88	0
669	HK030	206	4.2	-0.74	0.65	0.88	0	0.91	0
670	HK031	206	3.2	-2.81	0.66	0.77	-0.2	0.78	-0.2
671	HK032	206	4.8	0.56	0.64	0.62	-0.6	0.63	-0.6
672	HK033	206	3.0	-2.24	0.66	0.26	-1.6	0.25	-1.6
673	HK034	206	4.2	-0.27	0.65	1.57	1.0	1.65	1.1
674	HK035	206	3.3	-2.31	0.67	0.19	-1.9	0.19	-1.9
675	HK036	206	3.8	-1.72	0.72	8.05	4.9	8.10	5.0
676	HK037	206	3.0	-3.08	0.81	1.08	0.3	1.08	0.3
677	HK038	206	3.7	-1.78	0.66	0.75	-0.2	0.75	-0.2
678	HK039	206	3.0	-2.24	0.66	1.32	0.6	1.33	0.7

679	HK040	206	4.0	-0.70	0.66	0.65	-0.4	0.65	-0.5
680	HK041	206	2.8	-3.42	0.65	0.71	-0.3	0.72	-0.3
681	HK042	206	4.2	-0.84	0.65	1.00	0.2	1.04	0.2
682	HK043	206	3.0	-2.97	0.66	0.29	-1.5	0.29	-1.5
683	HK044	206	3.8	-0.64	0.66	3.17	2.5	3.15	2.5
684	HK045	206	3.2	-2.45	0.66	0.59	-0.6	0.59	-0.6
685	HK046	206	2.8	-3.22	0.65	0.33	-1.4	0.32	-1.4
686	HK047	206	3.5	-1.57	0.66	0.28	-1.5	0.28	-1.5
687	HK048	206	5.2	2.43	0.67	2.69	2.3	3.14	2.7
688	HK049	206	3.7	-1.13	0.67	2.51	2.0	2.51	2.0
689	HK050	206	3.5	-2.06	0.82	0.86	0	0.87	0
690	HK051	206	2.8	-2.82	0.65	0.79	-0.2	0.79	-0.1
691	HK052	206	3.2	-2.80	0.66	0.36	-1.2	0.35	-1.3
692	HK053	206	2.8	-3.00	0.65	0.86	0	0.88	0
693	HK054	206	3.8	-0.64	0.66	2.44	1.9	2.38	1.9
694	HK055	206	3.8	-1.18	0.66	2.79	2.2	2.84	2.3
695	HK056	206	2.5	-3.60	0.64	1.63	1.1	1.66	1.1
696	HK057	206	2.0	-4.74	0.65	0.33	-1.5	0.33	-1.5
697	HK058	206	3.3	-1.97	0.66	2.80	2.2	2.82	2.3
698	HK059	206	3.3	-1.71	0.66	2.53	2.0	2.52	2.0
699	HK060	206	2.7	-3.33	0.65	0.56	-0.7	0.58	-0.6
700	HK061	206	2.8	-3.01	0.65	0.47	-0.9	0.47	-0.9
701	HK062	206	2.7	-3.98	0.65	2.86	2.3	2.98	2.4
702	HK063	206	4.2	-0.57	0.66	2.68	2.1	2.69	2.1
703	HK064	206	2.5	-4.29	0.64	1.34	0.7	1.37	0.7
704	HK065	206	4.0	-0.55	0.66	2.11	1.6	1.99	1.5
705	HK066	206	3.2	-1.91	0.66	1.16	0.4	1.18	0.4
706	HK067	206	4.0	-0.74	0.66	0.84	0	0.85	0
707	HK068	206	2.8	-3.46	0.65	0.82	-0.1	0.82	-0.1
708	HK069	206	2.8	-3.41	0.65	2.10	1.6	2.02	1.5
709	HK070	206	4.0	-0.62	0.65	1.40	0.8	1.41	0.8
710	HK071	206	2.3	-4.85	0.64	0.59	-0.6	0.61	-0.6
711	HK072	206	2.3	-4.00	0.91	2.72	1.7	2.69	1.7
712	HK073	206	2.8	-2.85	0.65	2.72	2.2	2.65	2.2
713	HK074	206	2.0	-5.00	0.66	2.00	1.6	1.94	1.5
714	HK075	206	4.2	-0.40	0.66	1.13	0.4	1.14	0.4
715	HK076	206	1.8	-5.44	0.67	1.33	0.7	1.24	0.6
716	HK077	206	2.2	-4.78	0.72	1.46	0.8	1.43	0.8

717	HK078	206	2.4	-5.07	0.70	3.03	2.3	3.00	2.3
718	HK079	206	3.8	-0.90	0.66	0.81	-0.1	0.82	-0.1
719	HK080	206	3.0	-2.84	0.65	2.97	2.4	2.91	2.4
720	HK081	206	1.0	-7.23	1.97	Minim	um		
721	HK082	206	1.0	-7.35	1.95	Minim	um		
722	HK083	206	3.5	-1.86	0.66	1.23	0.5	1.25	0.5
723	HK084	206	4.3	-0.40	0.79	4.05	2.7	4.33	2.9
724	HK085	206	2.5	-3.49	0.64	0.30	-1.6	0.29	-1.6
725	HK086	206	2.2	-4.64	0.71	7.48	5.0	7.45	5
726	HK087	206	3.7	-1.02	0.67	1.86	1.3	1.83	1.3
727	HK088	206	1.5	-6.3	0.77	1.02	0.2	1.04	0.2
728	HK089	206	1.8	-5.37	0.75	0.39	-1.1	0.41	-1.1
729	HK090	206	3.3	-2.07	0.66	3.34	2.6	3.34	2.6
730	HK091	206	2.2	-4.97	0.65	1.42	0.8	1.42	0.8
731	HK092	206	2.3	-4.01	0.64	1.12	0.3	1.13	0.4
732	HK093	206	2.7	-3.82	0.64	0.39	-1.2	0.38	-1.2
733	HK094	206	4.0	-1.19	0.66	2.48	2.0	2.55	2.0
734	HK095	206	5.0	1.72	0.63	2.02	1.7	1.95	1.6
735	HK096	206	3.5	-2.12	0.67	0.14	-2.2	0.14	-2.2
736	HK097	206	3.3	-1.71	0.67	1.07	0.3	1.06	0.2
737	HK098	206	1.0	-8.97	1.92	Minim	um		
738	HK099	206	2.8	-3.26	0.65	0.58	-0.6	0.59	-0.6
739	HK100	206	2.3	-4.01	0.64	0.88	0	0.86	0
740	HK101	206	3.7	-1.42	0.67	0.81	-0.1	0.81	-0.1
741	HK102	206	2.3	-4.62	0.64	0.33	-1.5	0.33	-1.5
742	HK103	206	3.8	-1.09	0.66	1.07	0.3	1.09	0.3
743	HK104	206	2.6	-3.53	0.71	1.24	0.5	1.23	0.5
744	HK105	206	2.7	-3.63	0.64	0.59	-0.6	0.57	-0.7
745	HK106	206	2.0	-5.39	0.66	0.94	0	0.95	0
746	HK107	206	2.5	-3.69	0.64	0.86	0	0.86	0
747	HK108	206	1.4	-6.82	0.90	0.96	0.1	0.81	0
748	HK109	206	4.0	0.06	0.66	0.46	-0.9	0.46	-0.9
749	HK110	206	2.8	-3.65	0.65	1.75	1.2	1.80	1.3
750	HK111	206	2.5	-4.04	0.64	0.59	-0.6	0.61	-0.6
751	HK112	206	3.3	-1.92	0.66	3.78	2.9	3.75	2.9
752	HK113	206	4.8	1.32	0.63	1.60	1.1	1.57	1.0
753	HK114	206	3.8	-1.03	0.66	5.10	3.9	5.11	3.9
754	HK115	206	3.2	-2.54	0.66	2.06	1.5	2.06	1.5

755	HK116	206	4.2	0.03	0.65	0.43	-1.0	0.40	-1.1
756	HK117	206	3.0	-2.78	0.65	2.41	1.9	2.39	1.9
757	HK118	206	3.8	-1.13	0.66	0.92	0	0.93	0
758	HK119	206	3.7	-1.23	0.66	0.86	0	0.86	0
759	HK120	206	3.0	-2.98	0.65	1.83	1.3	1.79	1.3
760	HK121	206	3.0	-2.39	0.66	2.72	2.2	2.77	2.2
761	HK122	206	2.0	-5.20	0.66	0.67	-0.5	0.69	-0.4
762	HK123	206	3.3	-1.90	0.66	0.27	-1.5	0.27	-1.5
763	HK124	206	1.3	-6.80	0.87	0.33	-1.2	0.28	-1.0
764	HK125	206	3.2	-2.35	0.66	1.03	0.2	1.04	0.2
765	HK126	206	4.7	0.89	0.65	1.40	0.8	1.44	0.8
766	HK127	206	5.0	1.97	0.66	1.73	1.2	1.74	1.2
767	HK128	206	2.7	-3.18	0.65	2.53	2.0	2.64	2.1
768	HK129	206	5.2	1.98	0.65	0.46	-1.1	0.62	-0.6
769	HK130	206	4.2	-0.58	0.65	1.43	0.8	1.51	0.9
770	HK131	206	1.8	-5.50	0.83	0.21	-1.6	0.23	-1.5
771	HK132	206	4.0	-0.70	0.66	0.65	-0.5	0.64	-0.5
772	HK133	206	3.0	-2.44	0.66	1.67	1.1	1.65	1.1
773	HK134	206	3.0	-2.84	0.65	0.40	-1.1	0.40	-1.1
774	HK135	206	5.8	3.92	1.10	1.12	0.4	0.99	0.4
775	HK136	206	3.0	-3.14	0.65	0.82	-0.1	0.82	-0.1
776	HK137	206	2.3	-4.06	0.64	1.00	0.1	1.03	0.2

Separation: 4.38 Reliability: 0.95

Fixed (all same) chi-square: 12555.1 significance *p*: .00

Note: 104 = Grade 4 US student, 106 = Grade 6 US student, 108 = Grade 8 US student, 110 = Grade 10 US student, 112 = Grade 12 US student, 206 = Grade 6 HK student, 304 = Grade 4 US sub-sample student, 306 = Grade 6 US sub-sample student

Calibration of the students for all student essays using US and HK raters, Lexile Analyzer and IEA (127 misfitting students)

Student			Observed	Measure	S.E.	Infit		Outfit	
Number		Group	score	(Logit)		MnSq	ZStd	MnSq	ZStd
730	HK091	206	2.2	-4.97	0.65	1.42	0.8	1.42	0.8
487	US448	106	2.5	-1.30	0.29	1.43	1.5	1.42	1.5
227	US212	104	3.3	-0.88	0.30	1.43	1.5	1.43	1.5
769	HK130	206	4.2	-0.58	0.65	1.43	0.8	1.51	0.9
417	US387	110	4.6	1.57	0.32	1.44	1.5	1.44	1.5
152	US137	108	4.9	1.78	0.32	1.44	1.5	1.47	1.6

498	US456	106	3.4	-0.30	0.33	1.45	1.4	1.48	1.5
309	US280	104	4.4	0.02	0.39	1.45	1.2	1.43	1.1
307	US278	104	4.8	1.19	0.31	1.45	1.5	1.42	1.4
201	US186	108	4.2	1.35	0.33	1.45	1.4	1.46	1.4
716	HK077	206	2.2	-4.78	0.72	1.46	0.8	1.43	0.8
495	US454	106	4.1	-0.69	0.29	1.46	1.6	1.48	1.6
78	US064	306	4.2	-0.11	0.27	1.46	1.7	1.41	1.5
390	US360	110	3.4	-0.83	0.33	1.47	1.4	1.48	1.5
287	US260	104	3.0	-1.36	0.33	1.49	1.5	1.48	1.5
255	US235	306	2.6	-0.27	0.26	1.49	1.9	1.47	1.8
512	US466	106	4.2	0.77	0.29	1.49	1.8	1.49	1.8
298	US270	104	3.9	0.03	0.36	1.50	1.4	1.50	1.4
471	US436	104	4.4	-0.03	0.29	1.51	1.8	1.52	1.8
561	US513	112	3.7	0.02	0.32	1.51	1.6	1.50	1.6
293	US266	104	2.4	-1.59	0.29	1.53	1.9	1.53	1.9
373	US343	110	4.5	1.19	0.32	1.53	1.7	1.49	1.5
300	US272	104	4.5	-0.95	0.32	1.54	1.7	1.53	1.6
42	US034	106	3.1	0.69	0.29	1.54	1.8	1.57	1.9
486	US447	106	4.5	0.98	0.32	1.55	1.8	1.53	1.7
267	US243	306	4.4	1.07	0.26	1.55	2.0	1.54	2.0
482	US443	306	4.8	2.03	0.28	1.55	2.0	1.51	1.9
501	US459	106	2.8	-0.12	0.29	1.56	1.9	1.60	2.0
673	HK034	206	4.2	-0.27	0.65	1.57	1.0	1.65	1.1
176	US161	108	3.4	0.08	0.36	1.57	1.6	1.60	1.6
236	US218	304	4.2	0.72	0.27	1.57	2.1	1.58	2.1
752	HK113	206	4.8	1.32	0.63	1.60	1.1	1.57	1.0
203	US188	108	3.3	0.74	0.33	1.61	1.8	1.62	1.9
695	HK056	206	2.5	-3.60	0.64	1.63	1.1	1.66	1.1
230	US214	104	2.9	-1.75	0.32	1.66	2.0	1.66	2.0
772	HK133	206	3.0	-2.44	0.66	1.67	1.1	1.65	1.1
84	US070	106	2.3	-0.97	0.30	1.67	2.2	1.68	2.1
621	US573	104	4.5	0.56	0.32	1.67	2.0	1.73	2.2
314	US285	104	1.2	-5.66	0.59	1.69	1.4	3.12	2.0
662	HK023	206	4.0	-1.25	0.65	1.69	1.1	1.63	1.1
258	US237	106	3.3	-0.68	0.29	1.69	2.2	1.72	2.3
503	US460	106	3.1	0.14	0.30	1.69	2.3	1.67	2.2
461	US427	304	3.3	-0.21	0.28	1.70	2.4	1.68	2.4
297	US269	104	5.0	1.13	0.31	1.70	2.3	1.70	2.3

39	US031	106	3.5	-0.42	0.33	1.72	2.1	1.71	2.1
387	US357	110	4.1	0.82	0.32	1.72	2.1	1.73	2.1
312	US283	104	3.8	-0.58	0.33	1.73	2.1	1.75	2.1
766	HK127	206	5.0	1.97	0.66	1.73	1.2	1.74	1.2
749	HK110	206	2.8	-3.65	0.65	1.75	1.2	1.80	1.3
85	US070	306	2.4	-1.73	0.26	1.75	2.8	1.81	2.9
294	US266	304	2.1	-2.13	0.26	1.76	2.9	1.74	2.8
116	US101	108	2.9	-0.38	0.33	1.76	2.2	1.75	2.2
63	US052	104	3.2	-0.86	0.33	1.79	2.3	1.80	2.3
322	US292	106	5.2	1.66	0.35	1.79	2.3	1.50	1.4
228	US212	304	3.0	-1.23	0.27	1.81	2.8	1.82	2.8
496	US454	306	4.3	-0.85	0.26	1.81	2.8	1.88	3.0
150	US135	108	5.3	2.64	0.34	1.81	2.5	1.96	2.9
210	US195	108	3.6	-0.36	0.36	1.82	2.2	1.81	2.1
274	US248	306	4.9	1.77	0.27	1.82	3.0	1.75	2.8
759	HK120	206	3.0	-2.98	0.65	1.83	1.3	1.79	1.3
726	HK087	206	3.7	-1.02	0.67	1.86	1.3	1.83	1.3
19	US017	104	1.6	-2.90	0.34	1.88	2.8	1.87	2.8
310	US281	104	4.2	-0.34	0.33	1.88	2.4	1.94	2.5
492	US451	106	4.0	0.24	0.32	1.93	2.6	1.95	2.6
283	US256	304	1.8	-2.22	0.35	1.95	2.8	1.88	2.7
24	US021	104	3.3	-1.14	0.30	1.95	2.9	1.95	2.9
392	US362	110	3.3	-0.70	0.33	1.95	2.6	1.93	2.6
68	US057	104	3.7	-0.19	0.36	1.98	2.4	2.02	2.5
713	HK074	206	2.0	-5.00	0.66	2.00	1.6	1.94	1.5
626	US577	304	3.3	-0.48	0.29	2.00	3.1	2.00	3.1
513	US466	306	4.6	0.54	0.26	2.00	3.5	2.05	3.7
734	HK095	206	5.0	1.72	0.63	2.02	1.7	1.95	1.6
70	US058	104	2.5	-1.38	0.32	2.04	2.9	2.07	3.0
329	US299	106	3.8	0.21	0.35	2.04	2.6	2.16	2.9
20	US017	304	1.5	-3.02	0.30	2.06	3.7	1.98	3.4
754	HK115	206	3.2	-2.54	0.66	2.06	1.5	2.06	1.5
37	US030	106	3.3	-0.19	0.30	2.09	3.2	2.11	3.3
708	HK069	206	2.8	-3.41	0.65	2.10	1.6	2.02	1.5
43	US034	306	3.3	0.10	0.27	2.10	3.6	2.10	3.6
259	US237	306	3.4	-1.19	0.27	2.11	3.6	2.12	3.7
704	HK065	206	4.0	-0.55	0.66	2.11	1.6	1.99	1.5
648	HK009	206	3.7	-1.58	0.67	2.14	1.6	2.15	1.6

25	US021	304	3.1	-1.50	0.27	2.16	3.7	2.17	3.7
242	US224	104	2.4	-0.86	0.39	2.16	2.7	2.16	2.7
472	US436	304	4.4	-0.13	0.26	2.17	3.9	2.24	4.0
336	US306	106	3.0	-0.81	0.40	2.19	2.6	2.26	2.7
630	US581	104	3.1	-1.42	0.33	2.2	3.1	2.21	3.1
273	US248	106	4.0	0.72	0.29	2.2	3.6	2.15	3.5
502	US459	306	3.1	-0.23	0.27	2.23	3.9	2.25	3.9
38	US030	306	3.5	-0.51	0.27	2.25	3.9	2.25	3.9
504	US460	306	3.4	0.03	0.27	2.26	4.0	2.28	4.1
69	US057	304	3.5	0.15	0.33	2.27	3.2	2.27	3.2
666	HK027	206	2.8	-3.02	0.65	2.28	1.8	2.33	1.8
658	HK019	206	3.8	-1.23	0.66	2.29	1.8	2.34	1.8
756	HK117	206	3.0	-2.78	0.65	2.41	1.9	2.39	1.9
693	HK054	206	3.8	-0.64	0.66	2.44	1.9	2.38	1.9
733	HK094	206	4.0	-1.19	0.66	2.48	2.0	2.55	2.0
59	US049	104	2.7	-1.71	0.29	2.50	4.2	2.51	4.2
688	HK049	206	3.7	-1.13	0.67	2.51	2.0	2.51	2.0
767	HK128	206	2.7	-3.18	0.65	2.53	2.0	2.64	2.1
698	HK059	206	3.3	-1.71	0.66	2.53	2.0	2.52	2.0
396	US366	110	3.9	0.31	0.33	2.53	3.6	2.49	3.5
702	HK063	206	4.2	-0.57	0.66	2.68	2.1	2.69	2.1
687	HK048	206	5.2	2.43	0.67	2.69	2.3	3.14	2.7
650	HK011	206	2.3	-3.86	0.64	2.71	2.3	2.81	2.3
60	US049	304	2.7	-1.80	0.26	2.71	5.1	2.70	5.1
711	HK072	206	2.3	-4.00	0.91	2.72	1.7	2.69	1.7
712	HK073	206	2.8	-2.85	0.65	2.72	2.2	2.65	2.2
760	HK121	206	3.0	-2.39	0.66	2.72	2.2	2.77	2.2
694	HK055	206	3.8	-1.18	0.66	2.79	2.2	2.84	2.3
697	HK058	206	3.3	-1.97	0.66	2.80	2.2	2.82	2.3
701	HK062	206	2.7	-3.98	0.65	2.86	2.3	2.98	2.4
620	US572	104	3.2	-0.89	0.33	2.90	4.5	2.90	4.5
511	US465	106	2.3	-2.37	0.34	2.96	4.5	2.97	4.1
719	HK080	206	3.0	-2.84	0.65	2.97	2.4	2.91	2.4
717	HK078	206	2.4	-5.07	0.70	3.03	2.3	3.00	2.3
683	HK044	206	3.8	-0.64	0.66	3.17	2.5	3.15	2.5
729	HK090	206	3.3	-2.07	0.66	3.34	2.6	3.34	2.6
751	HK112	206	3.3	-1.92	0.66	3.78	2.9	3.75	2.9
723	HK084	206	4.3	-0.40	0.79	4.05	2.7	4.33	2.9

625	US577	104	3.1	-1.24	0.31	4.11	6.6	4.19	6.7
299	US271	104	3.2	-1.43	0.36	4.30	5.9	4.34	6.0
753	HK114	206	3.8	-1.03	0.66	5.10	3.9	5.11	3.9
725	HK086	206	2.2	-4.64	0.71	7.48	5.0	7.45	5.0
628	US579	104	3.8	-1.12	0.33	7.70	9.0	7.08	9.0
675	HK036	206	3.8	-1.72	0.72	8.05	4.9	8.10	5.0
660	HK021	206	3.8	-1.43	0.66	9.00	5.8	8.79	5.7

Separation: 3.68 Reliability: 0.93

Fixed (all same) chi-square: 8841.3 significance *p*: .00

Note: 104 = Grade 4 US student, 106 = Grade 6 US student, 108 = Grade 8 US student, 110 = Grade 10 US student, 112 = Grade 12 US student, 206 = Grade 6 HK student, 304 = Grade 4 US sub-sample student, 306 = Grade 6 US sub-sample student

Appendix 22

Bias calibration report for the US rater-student interactions

US rater	Student Number	Group	Observed score	Expected score	Obs-Exp average	Bias (logit)	S.E.	z-score	Fit MnSq
USR101	19	104	3	1.5	1.54	3.91	1.63	2.41	0
USR101	20	304	3	1.4	1.58	4.04	1.63	2.49	0
USR101	45	106	11	8.8	1.11	3.02	1.33	2.27	0.1
USR101	314	104	4	2.2	0.92	3.70	1.11	3.33	0.1
USR101	315	304	2	1.1	0.92	3.53	1.56	2.26	0
USR101	383	110	8	9.9	-0.96	-2.35	1.16	-2.03	0.1
USR101	417	110	8	9.8	-0.90	-2.27	1.12	-2.03	0.6
USR101	587	112	8	9.9	-0.94	-2.29	1.15	-1.99	1.7
USR101	625	104	8	6.1	0.95	2.52	1.17	2.16	0
USR101	628	104	10	7.7	1.16	2.86	1.07	2.68	0
USR102	2	104	2	3.4	-1.37	-3.44	1.56	-2.20	0
USR102	19	104	5	3.1	0.93	2.37	1.11	2.14	1.2
USR102	20	304	3	1.4	1.64	4.25	1.63	2.62	0
USR102	90	106	7	8.8	-0.92	-2.41	1.15	-2.09	0.2
USR102	116	108	8	6.3	0.85	2.21	1.11	1.99	0.9
USR102	152	108	8	10.3	-1.14	-2.78	1.15	-2.42	0.1
USR102	224	108	10	8.0	0.99	2.48	1.15	2.15	0.7
USR102	242	104	4	2.6	1.36	3.57	1.65	2.16	0
USR102	261	106	3	4.3	-1.29	-3.42	1.63	-2.10	0
USR102	314	104	3	2.1	0.44	2.55	1.27	2.01	0.4
USR102	318	104	8	9.9	-0.95	-2.31	1.17	-1.98	0
USR102	390	110	10	7.9	1.07	2.60	1.07	2.44	1.9
USR102	538	112	3	4.3	-1.28	-3.40	1.63	-2.09	0
USR102	561	112	5	6.8	-0.91	-2.33	1.15	-2.02	1.2
USR103	516	106	5	3.7	1.26	3.14	1.51	2.08	0
USR103	628	104	9	7.1	0.94	2.46	1.09	2.26	0.4
USR104	81	106	3	4.3	-1.29	-3.43	1.63	-2.11	0
USR104	163	108	8	9.9	-0.93	-2.27	1.15	-1.97	0.1
USR104	223	108	8	10	-0.98	-2.39	1.16	-2.05	0
USR104	242	104	4	2.4	1.59	4.14	1.65	2.51	0
USR104	249	106	8	6.2	0.91	2.37	1.11	2.14	0.5
USR104	314	104	2	1.1	0.95	3.98	1.56	2.55	0
USR104	554	112	7	5.2	0.90	2.30	1.16	1.99	0

USR105	5	104	8	5.9	1.06	2.77	1.12	2.47	0.5
USR105	89	106	6	7.8	-0.91	-2.40	1.14	-2.11	0.2
USR105	195	108	6	8.0	-1.01	-2.68	1.14	-2.34	0.1
USR105	204	108	5	8.3	-1.63	-4.18	1.13	-3.70	0
USR105	210	108	5	7.8	-1.39	-3.62	1.12	-3.24	1.9
USR105	240	104	3	4.4	-1.36	-3.60	1.63	-2.21	0
USR105	242	104	4	2.6	1.36	3.59	1.65	2.17	0
USR105	248	104	5	3.6	1.44	3.62	1.51	2.40	0
USR105	292	104	3	4.4	-1.45	-3.82	1.63	-2.35	0
USR105	298	104	3	4.3	-1.28	-3.39	1.63	-2.09	0
USR105	300	104	3	4.2	-1.20	-3.20	1.63	-1.97	0
USR105	431	106	7	9.0	-0.99	-2.54	1.16	-2.19	0
USR105	454	104	6	8.0	-0.99	-2.57	1.12	-2.29	0.5
USR105	538	112	6	7.9	-0.97	-2.58	1.13	-2.28	0.2
USR105	578	112	6	7.9	-0.97	-2.59	1.15	-2.25	0
USR105	611	112	7	5.1	0.97	2.51	1.15	2.18	0.3
USR105	615	112	6	7.7	-0.87	-2.30	1.13	-2.03	1.2
USR106	396	110	1	2.9	-1.93	-3.72	1.79	-2.08	0.7
USR200	49	106	2	3.3	-1.29	-3.23	1.56	-2.07	0
USR200	79	106	2	3.5	-1.47	-3.71	1.56	-2.38	0
USR200	107	108	5	3.2	1.84	4.66	1.51	3.09	0
USR200	110	108	4	2.4	1.56	4.07	1.65	2.46	0
USR200	149	108	5	6.9	-0.93	-2.39	1.11	-2.15	0.2
USR200	169	108	4	2.7	1.28	3.39	1.65	2.05	0
USR200	197	108	4	2.7	1.32	3.48	1.65	2.11	0
USR200	203	108	8	6.2	0.89	2.38	1.17	2.04	2.4
USR200	263	106	3	5.1	-1.06	-2.67	1.25	-2.14	0.6
USR200	266	106	4	5.9	-0.96	-2.42	1.15	-2.12	0.5
USR200	273	106	3	5.5	-1.23	-3.45	1.38	-2.51	0.1
USR200	383	110	5	3.7	1.35	3.36	1.51	2.23	0
USR200	385	110	5	3.5	1.53	3.85	1.51	2.56	0
USR200	403	110	2	3.4	-1.43	-3.61	1.56	-2.31	0
USR200	460	104	4	2.5	1.51	3.95	1.65	2.39	0
USR200	461	304	4	2.6	1.38	3.62	1.65	2.19	0
USR200	471	104	2	3.5	-1.54	-3.89	1.56	-2.49	0
USR200	472	304	2	3.5	-1.50	-3.79	1.56	-2.43	0
USR200	492	106	10	7.8	1.09	2.65	1.07	2.49	0
USR200	511	106	4	6.9	-0.95	-3.71	1.33	-2.80	2.2

USR200	618	112	4	2.5	1.47	3.85	1.65	2.33	0
USR201	68	104	5	3.7	1.27	3.15	1.51	2.09	0
USR201	70	104	4	2.7	1.28	3.38	1.65	2.05	0
USR201	219	108	7	9.2	-0.72	-1.83	0.91	-2.01	0.1
USR201	239	104	4	2.7	1.26	3.33	1.65	2.01	0
USR201	279	104	7	5.1	0.93	2.42	1.15	2.11	0.4
USR201	321	106	5	3.8	1.21	2.99	1.51	1.98	0
USR201	340	110	5	7.1	-1.06	-2.70	1.15	-2.35	0.2
USR201	353	110	6	8.5	-0.82	-2.17	0.98	-2.22	0.9
USR201	371	110	2	3.6	-1.60	-4.04	1.56	-2.59	0
USR201	373	110	3	4.8	-1.78	-4.58	1.63	-2.82	0
USR201	374	110	4	5.5	-1.52	-3.71	1.65	-2.25	0
USR201	376	110	3	4.4	-1.39	-3.66	1.63	-2.25	0
USR201	379	110	2	3.5	-1.50	-3.79	1.56	-2.43	0
USR201	380	110	3	4.2	-1.23	-3.28	1.63	-2.02	0
USR201	382	110	2	3.7	-1.72	-4.35	1.56	-2.79	0
USR201	406	110	2	3.4	-1.41	-3.53	1.56	-2.27	0
USR201	508	106	2	3.4	-1.43	-3.61	1.56	-2.31	0
USR201	515	106	5	3.8	1.20	2.98	1.51	1.97	0
USR201	535	112	5	3.7	1.26	3.12	1.51	2.07	0
USR201	538	112	5	3.5	1.46	3.67	1.51	2.43	0
USR201	549	112	5	3.4	1.63	4.11	1.51	2.72	0
USR201	558	112	5	3.6	1.41	3.54	1.51	2.35	0
USR201	625	104	2	3.8	-1.78	-4.54	1.56	-2.91	0
USR201	628	104	1	4.7	-3.66	-8.48	1.89	-4.48	0.5
USR201	634	104	2	3.7	-1.72	-4.38	1.56	-2.81	0
USR202	21	104	5	3.6	1.35	3.38	1.51	2.24	0
USR202	24	104	2	3.3	-1.25	-3.13	1.56	-2.00	0
USR202	93	106	3	4.4	-1.42	-3.75	1.63	-2.31	0
USR202	176	108	2	3.3	-1.30	-3.25	1.56	-2.09	0
USR202	298	104	6	4.1	1.86	3.61	1.84	1.96	0.6
USR202	299	104	1	3.7	-2.71	-5.74	1.78	-3.23	0.7
USR202	310	104	6	3.9	2.11	4.41	1.91	2.31	0.5
USR202	312	104	6	3.8	2.21	4.58	1.86	2.46	0.5
USR202	313	104	6	4.1	1.93	3.68	1.80	2.05	0.6
USR202	345	110	3	4.6	-1.64	-4.27	1.63	-2.62	0
USR202	346	110	3	4.3	-1.32	-3.51	1.63	-2.16	0
USR202	468	104	10	7.6	1.18	2.92	1.07	2.74	2.3

USR202	556	112	4	5.8	-0.91	-2.28	1.13	-2.02	0.3
USR202	586	112	3	4.2	-1.22	-3.24	1.63	-1.99	0
USR202	620	104	5	7.1	-1.05	-2.71	1.10	-2.46	1.0
USR202	625	104	3	7.1	-2.05	-5.41	1.32	-4.11	2.1
USR202	626	304	2	3.5	-1.50	-3.79	1.56	-2.43	0
USR203	58	104	2	3.3	-1.29	-3.23	1.56	-2.07	0
USR203	59	104	6	8.5	-0.82	-2.02	0.91	-2.23	1.1
USR203	68	104	5	7.7	-1.33	-3.48	1.10	-3.15	1.1
USR203	83	106	5	3.6	1.40	3.50	1.51	2.32	0
USR203	89	106	6	4.0	1.99	3.74	1.76	2.12	0.7
USR203	110	108	8	10.3	-0.77	-1.98	0.91	-2.18	0.6
USR203	241	104	12	14.4	-0.61	-1.61	0.81	-2.00	0.6
USR203	258	106	5	7.4	-1.20	-3.11	1.10	-2.82	0.4
USR203	259	306	5	7.0	-1.01	-2.61	1.10	-2.36	0.4
USR203	265	106	6	4.0	1.95	3.70	1.78	2.07	0.6
USR203	285	104	5	7.0	-0.99	-2.55	1.10	-2.31	0.4
USR203	305	104	3	4.4	-1.41	-3.72	1.63	-2.29	0
USR203	330	106	16	13.6	0.81	2.03	0.97	2.09	0.2
USR203	356	110	8	6.1	0.95	2.49	1.13	2.21	0.9
USR203	383	110	11	9.0	1.01	2.70	1.31	2.05	0.1
USR203	388	110	9	7.2	0.89	2.22	1.12	1.98	0
USR203	389	110	10	7.8	1.08	2.70	1.15	2.34	0.6
USR203	479	106	8	10.2	-0.72	-1.84	0.93	-1.99	0
USR203	504	306	2	3.5	-1.53	-3.85	1.56	-2.47	0
USR203	620	104	1	3.5	-2.52	-5.51	1.89	-2.92	0.6
USR203	628	104	5	3.8	1.24	3.08	1.51	2.04	0
USR204	144	108	5	6.7	-0.85	-2.18	1.10	-1.98	0.4
USR204	177	108	3	4.3	-1.33	-3.53	1.63	-2.17	0
USR204	236	304	5	3.8	1.21	2.99	1.51	1.98	0
USR204	258	106	2	3.6	-1.59	-4.03	1.56	-2.58	0
USR204	259	306	2	3.4	-1.40	-3.52	1.56	-2.26	0
USR204	279	104	4	2.6	1.37	3.61	1.65	2.18	0
USR204	285	104	2	3.4	-1.38	-3.46	1.56	-2.22	0
USR204	287	104	1	2.8	-1.85	-3.63	1.84	-1.97	0.6
USR204	300	104	3	4.4	-1.40	-3.71	1.63	-2.28	0
USR204	336	106	2	4.0	-2.04	-5.23	1.56	-3.35	0
USR204	480	106	10	12.8	-0.92	-2.43	0.94	-2.59	0.2
USR204	495	106	10	12.4	-0.79	-2.11	0.94	-2.25	0.6

USR204	559	112	4	2.8	1.24	3.27	1.65	1.98	0
USR205	63	104	8	6.2	0.92	2.42	1.12	2.15	0.5
USR205	150	108	8	10.8	-1.39	-3.39	1.16	-2.92	3.4
USR205	199	108	2	3.2	-1.23	-3.07	1.56	-1.97	0
USR205	204	108	5	3.7	1.33	3.31	1.51	2.20	0
USR205	300	104	11	8.9	1.05	2.50	1.22	2.05	0.7
USR205	312	104	2	3.8	-1.77	-4.49	1.56	-2.88	0
USR205	371	110	2	3.6	-1.57	-3.95	1.56	-2.53	0
USR205	373	110	3	4.7	-1.74	-4.49	1.63	-2.76	0
USR205	374	110	4	5.5	-1.48	-3.62	1.65	-2.19	0
USR205	387	110	3	4.2	-1.21	-3.21	1.63	-1.98	0
USR205	392	110	5	6.8	-0.88	-2.26	1.12	-2.01	2.8
USR205	620	104	7	5.2	0.90	2.34	1.15	2.04	0.7
USR205	625	104	8	6.2	0.89	2.36	1.15	2.04	0.1
USR206	23	104	6	8.2	-1.08	-2.87	1.14	-2.51	1.9
USR206	25	304	1	2.8	-1.83	-3.62	1.85	-1.96	0.6
USR206	73	304	6	4.1	0.95	2.35	1.15	2.05	0
USR206	122	108	2	3.5	-1.46	-3.68	1.56	-2.36	0
USR206	273	106	4	6.1	-1.04	-2.65	1.16	-2.28	0.7
USR206	296	104	4	2.6	1.42	3.72	1.65	2.25	0
USR206	322	106	3	4.6	-1.64	-4.26	1.63	-2.62	0
USR206	621	104	10	8.0	1.00	2.43	1.08	2.26	0.1
USR206	625	104	3	6.6	-1.78	-4.67	1.32	-3.55	2.1
USR206	626	304	2	3.2	-1.22	-3.05	1.56	-1.96	0
USR206	628	104	6	8.1	-1.06	-2.77	1.13	-2.46	9.0
USR206	637	104	7	4.6	1.18	3.01	1.15	2.62	0.6
USR207	24	104	2	3.2	-1.25	-3.11	1.56	-2.00	0
USR207	88	106	3	5.0	-0.98	-2.60	1.30	-2.00	0.2
USR207	260	106	5	3.4	1.56	3.93	1.51	2.61	0
USR207	486	106	11	8.3	1.33	3.44	1.30	2.65	0.1
USR207	495	106	9	7.1	0.95	2.42	1.11	2.19	0
USR207	510	106	2	3.3	-1.27	-3.17	1.56	-2.03	0
USR207	621	104	7	9.6	-1.28	-3.25	1.16	-2.80	1.7
USR207	626	304	2	3.5	-1.50	-3.78	1.56	-2.42	0
USR207	627	104	9	10.8	-0.89	-2.17	1.10	-1.98	0.1
USR207	630	104	5	6.7	-0.86	-2.21	1.10	-2.00	0.3
USR208	24	104	9	6.9	1.07	2.80	1.09	2.58	0.5
USR208	37	106	9	7.0	0.99	2.55	1.10	2.32	1.6

USR208	53	104	6	7.8	-0.91	-2.43	1.15	-2.12	0
USR208	59	104	2	5.7	-1.85	-4.65	1.74	-2.67	0.2
USR208	67	104	4	6.1	-1.03	-2.57	1.11	-2.31	0.1
USR208	162	108	7	8.8	-0.91	-2.42	1.15	-2.10	0.5
USR208	166	108	13	15.3	-0.77	-1.86	0.92	-2.03	0
USR208	176	108	10	7.7	1.13	2.78	1.07	2.59	0.1
USR208	249	106	2	3.7	-1.71	-4.33	1.56	-2.78	0
USR208	455	104	3	4.5	-1.49	-3.90	1.63	-2.40	0
USR208	487	106	3	1.7	1.31	3.25	1.63	2.00	0
USR208	532	112	2	3.3	-1.26	-3.16	1.56	-2.03	0
USR208	543	112	2	3.5	-1.49	-3.76	1.56	-2.41	0
USR208	564	112	4	2.7	1.26	3.33	1.65	2.02	0
USR208	584	112	10	12.7	-0.89	-2.26	0.93	-2.44	0.1
USR208	588	112	13	15.8	-0.92	-2.17	0.91	-2.38	0.8
USR208	620	104	8	5.4	1.28	3.39	1.17	2.90	0
USR209	50	104	5	6.7	-0.87	-2.24	1.10	-2.03	0.4
USR209	63	104	4	6.5	-1.25	-3.13	1.10	-2.83	2.1
USR209	154	108	5	3.8	1.24	3.08	1.51	2.04	0
USR209	297	104	12	9.2	1.38	3.56	1.82	1.96	0.3
USR209	298	104	5	3.5	1.49	3.72	1.51	2.47	0
USR209	331	106	5	3.5	1.47	3.68	1.51	2.44	0
USR209	373	110	16	13.3	0.90	2.10	0.93	2.26	0.8
USR209	384	110	5	3.8	1.20	2.95	1.51	1.96	0
USR209	475	106	2	3.4	-1.44	-3.63	1.56	-2.32	0
USR209	476	306	2	3.3	-1.28	-3.20	1.56	-2.05	0
USR209	492	106	8	10.4	-0.81	-2.13	0.95	-2.23	2.0
USR209	506	106	5	3.8	1.19	2.95	1.51	1.96	0
USR209	511	106	2	3.4	-1.45	-3.65	1.56	-2.34	0
USR209	542	112	11	8.9	0.71	1.86	0.94	1.97	0.1
USR209	547	112	11	7.7	1.10	2.80	0.92	3.05	0.2
USR210	39	106	4	6.5	-1.27	-3.85	1.46	-2.64	3.5
USR210	108	108	3	5.4	-1.22	-3.08	1.25	-2.47	1.0
USR210	128	108	8	6.0	0.98	2.62	1.16	2.25	0
USR210	176	108	5	3.7	1.35	3.36	1.51	2.23	0
USR210	177	108	5	3.8	1.19	2.95	1.51	1.96	0
USR210	188	108	5	3.8	1.23	3.05	1.51	2.02	0
USR210	275	106	13	10	0.99	2.54	0.89	2.84	0.4
USR210	278	104	5	3.7	1.33	3.31	1.51	2.20	0

USR210	309	104	7	9.6	-1.28	-3.27	1.15	-2.85	0.7
USR210	356	110	4	6.2	-1.11	-2.81	1.14	-2.46	0.9
USR210	361	110	6	8.0	-1.02	-2.66	1.12	-2.37	0.5
USR210	392	110	4	6.3	-1.14	-2.85	1.11	-2.57	1.7
USR210	394	110	2	3.4	-1.42	-3.56	1.56	-2.28	0
USR210	413	110	4	5.8	-0.89	-2.17	1.10	-1.97	2.0
USR210	417	110	21	18	0.75	1.94	0.87	2.23	0.9
USR210	445	106	3	5.1	-1.04	-2.65	1.26	-2.10	1.3
USR210	459	104	5	3.6	1.39	3.47	1.51	2.30	0
USR210	470	104	1	2.8	-1.84	-3.63	1.85	-1.97	0.6
USR210	531	112	11	9.0	1.00	2.54	1.27	2.00	0.2
USR210	554	112	2	3.3	-1.31	-3.29	1.56	-2.11	0
USR210	564	112	7	9.2	-0.75	-1.87	0.92	-2.03	1.2
USR210	565	112	12	14.4	-0.78	-1.93	0.95	-2.03	0
USR210	566	112	8	11.1	-1.04	-2.70	0.91	-2.98	0.4
USR210	567	112	9	11.1	-0.70	-1.85	0.93	-1.99	1.1
USR210	593	112	5	7.3	-1.15	-2.99	1.10	-2.72	0.6
USR210	621	104	6	4.0	2.00	3.75	1.75	2.14	0.7
USR210	625	104	1	3.8	-2.78	-6.33	1.94	-3.26	0.5
USR210	628	104	1	4.7	-3.66	-8.47	1.90	-4.47	0.5
USR210	629	104	2	3.8	-1.75	-4.45	1.56	-2.85	0

Note: 104 = Grade 4 US student, 106 = Grade 6 US student, 108 = Grade 8 US student,
110 = Grade 10 US student, 112 = Grade 12 US student, 304 = Grade 4 US sub-sample student,
306 = Grade 6 US sub-sample student

Bias calibration report for HK rater-student interactions

HK rater	Student Number	Student Group	Observed score	Expected score	Obs-Exp average	Bias (logit)	Error	z-score	Fit MnSq
HKR001	7	304	6	7.8	-0.92	-2.44	1.14	-2.13	0.1
HKR001	38	306	5	3.1	1.93	4.90	1.51	3.25	0
HKR001	255	306	4	2.5	1.51	3.95	1.65	2.39	0
HKR001	315	304	2	1.1	0.93	3.74	1.56	2.40	0
HKR001	502	306	3	6.0	-1.52	-3.87	1.26	-3.07	1.3
HKR001	660	206	9	7.0	0.99	2.59	1.09	2.38	0.4
HKR001	683	206	10	7.2	1.41	3.53	1.07	3.30	0
HKR001	698	206	5	3.3	1.75	4.41	1.51	2.93	0
HKR001	712	206	4	2.7	1.32	3.50	1.65	2.12	0
HKR001	753	206	4	7.1	-1.56	-3.94	1.10	-3.57	0

HKR001	769	206	6	7.8	-0.88	-2.34	1.15	-2.04	0
HKR002	20	304	3	1.6	1.39	3.45	1.63	2.12	0
HKR002	56	304	8	6.1	0.93	2.47	1.15	2.15	1.5
HKR002	60	304	1	2.9	-1.88	-3.67	1.82	-2.01	0.6
HKR002	259	306	10	7.4	1.28	3.17	1.07	2.96	2.8
HKR002	271	306	5	3.5	1.49	3.74	1.51	2.48	0
HKR002	302	304	3	4.2	-1.22	-3.26	1.63	-2.01	0
HKR002	435	306	4	2.7	1.29	3.40	1.65	2.06	0
HKR002	502	306	2	3.5	-1.48	-3.74	1.56	-2.40	0
HKR002	626	304	5	3.4	1.62	4.09	1.51	2.71	0
HKR002	636	304	4	5.8	-0.90	-2.19	1.10	-1.99	0
HKR002	694	206	6	7.8	-0.91	-2.43	1.15	-2.11	2.9
HKR002	697	206	2	3.5	-1.54	-3.89	1.56	-2.50	0
HKR002	729	206	2	3.3	-1.29	-3.22	1.56	-2.07	0
HKR003	15	304	4	5.4	-1.40	-3.37	1.65	-2.04	0
HKR003	32	306	3	4.4	-1.36	-3.60	1.63	-2.21	0
HKR003	38	306	2	4.2	-2.17	-5.58	1.56	-3.58	0
HKR003	56	304	5	3.7	1.26	3.12	1.51	2.07	0
HKR003	73	304	3	4.9	-0.96	-2.55	1.29	-1.97	0.3
HKR003	78	306	3	4.8	-1.83	-4.69	1.63	-2.89	0
HKR003	85	306	2	6.0	-2.02	-4.91	1.63	-3.02	0.3
HKR003	228	304	2	3.9	-1.93	-4.94	1.56	-3.17	0
HKR003	513	306	8	10.9	-1.45	-3.55	1.17	-3.04	2.4
HKR003	636	304	2	3.3	-1.28	-3.21	1.56	-2.06	0
HKR003	666	206	5	3.3	1.73	4.37	1.51	2.90	0
HKR003	678	206	5	3.6	1.43	3.59	1.51	2.38	0
HKR003	683	206	6	8.5	-1.24	-3.32	1.15	-2.89	0
HKR003	697	206	10	7.4	1.29	3.21	1.07	3.01	0
HKR003	701	206	4	6.2	-1.10	-2.73	1.11	-2.47	3.0
HKR003	703	206	8	6.1	0.94	2.52	1.17	2.16	0
HKR003	704	206	3	4.7	-1.65	-4.29	1.63	-2.64	0
HKR003	726	206	3	4.2	-1.24	-3.29	1.63	-2.03	0
HKR003	753	206	11	8.6	1.22	2.94	1.22	2.41	0.9
HKR003	760	206	2	3.6	-1.60	-4.04	1.56	-2.59	0
HKR003	772	206	2	3.4	-1.40	-3.53	1.56	-2.26	0
HKR004	43	306	5	3.4	1.63	4.10	1.51	2.72	0
HKR004	232	304	2	3.2	-1.25	-3.12	1.56	-2.00	0
HKR004	236	304	5	7.5	-1.26	-3.29	1.11	-2.97	1.4

HKR004	274	306	8	10	-1.00	-2.43	1.15	-2.11	3.9
HKR004	461	304	5	3.3	1.75	4.42	1.51	2.93	0
HKR004	509	306	10	8.2	0.92	2.21	1.07	2.06	0
HKR004	518	306	11	8.3	1.34	3.31	1.24	2.67	1.3
HKR004	648	206	2	3.4	-1.35	-3.40	1.56	-2.18	0
HKR004	658	206	2	3.6	-1.59	-4.02	1.56	-2.58	0
HKR004	660	206	2	3.5	-1.52	-3.84	1.56	-2.46	0
HKR004	694	206	9	7.0	0.99	2.59	1.09	2.38	0.4
HKR004	698	206	4	6.1	-1.03	-2.55	1.10	-2.31	0
HKR004	702	206	6	7.8	-0.88	-2.34	1.15	-2.03	2.8
HKR004	717	206	4	2.0	2.04	5.22	1.65	3.16	0
HKR004	719	206	4	2.8	1.25	3.30	1.65	2.00	0
HKR004	723	206	6	3.8	2.21	4.58	1.86	2.46	0.5
HKR004	729	206	5	3.2	1.84	4.65	1.51	3.08	0
HKR004	733	206	2	3.5	-1.52	-3.84	1.56	-2.46	0
HKR004	753	206	5	3.4	1.56	3.92	1.51	2.60	0
HKR004	769	206	5	3.8	1.25	3.09	1.51	2.05	0

Note: 206 = Grade 6 HK student, 304 = Grade 4 US sub-sample student, 306 = Grade 6 US sub-sample student

Bias calibration report for the US rater-prompt interactions

US rater	Prompt	Observed score	Expected score	Obs-Exp average	Bias (logit)	S.E.	z-score	Fit MnSq
USR101	4AN	326	303.3	0.28	0.72	0.18	4.03	0.7
USR101	12AN	733	756.3	-0.12	-0.32	0.12	-2.73	0.6
USR101	12BN	291	313.5	-0.29	-0.74	0.18	-4.05	0.6
USR102	4AN	313	327.7	-0.17	-0.44	0.17	-2.53	0.8
USR102	12AN	591	608.4	-0.11	-0.29	0.13	-2.26	0.6
USR104	8AN	731	760.1	-0.13	-0.35	0.11	-3.17	0.5
USR104	12AN	726	683.0	0.23	0.60	0.12	5.09	0.5
USR105	4AN	326	313.3	0.15	0.39	0.17	2.22	1.0
USR106	4BN	579	598.4	-0.13	-0.33	0.13	-2.52	0.6
USR106	8AN	723	706.5	0.09	0.24	0.12	2.00	0.5
USR200	4AI	181	171.1	0.19	0.48	0.22	2.19	0.8
USR200	4AP	58	64.5	-0.32	-0.88	0.37	-2.38	1.1
USR200	4BI	280	254.4	0.37	0.95	0.19	4.97	0.9
USR200	4BP	241	274.5	-0.39	-0.99	0.17	-5.73	1.4
USR200	8AI	192	208.1	-0.23	-0.61	0.20	-3.13	0.8

USR200	8BI	175	191.4	-0.29	-0.76	0.22	-3.50	0.8
USR200	8BP	223	207.1	0.21	0.55	0.19	2.95	1.1
USR200	12AP	269	250.9	0.24	0.62	0.18	3.34	0.8
USR201	4AP	266	255.7	0.15	0.40	0.20	2.02	1.8
USR201	4BI	549	526.0	0.15	0.40	0.13	3.02	0.7
USR201	8AI	445	409.3	0.29	0.76	0.15	5.20	0.8
USR201	8BI	407	420.7	-0.13	-0.33	0.15	-2.12	0.7
USR201	8BP	391	462.9	-0.55	-1.42	0.14	-10.09	1.0
USR201	12BP	192	165.6	0.66	1.68	0.26	6.50	0.8
USR202	4AP	115	122.7	-0.25	-0.63	0.29	-2.22	3.4
USR202	4BI	229	200.3	0.56	1.51	0.23	6.52	1.4
USR202	4BP	315	299.9	0.18	0.46	0.18	2.65	1.5
USR202	8AP	245	216.5	0.37	0.97	0.18	5.23	0.6
USR202	12AI	214	229.4	-0.27	-0.67	0.21	-3.20	0.5
USR202	12BI	114	127.6	-0.37	-0.96	0.26	-3.63	0.8
USR202	12BP	70	81.2	-0.56	-1.43	0.35	-4.03	0.8
USR203	4AI	283	298.7	-0.19	-0.50	0.18	-2.81	0.7
USR203	4BI	560	580.4	-0.13	-0.34	0.13	-2.61	1.0
USR203	8AI	313	298.6	0.17	0.45	0.18	2.54	1.2
USR203	8AP	245	271.1	-0.25	-0.65	0.16	-4.08	0.7
USR203	8BP	352	316.6	0.43	1.10	0.18	6.20	0.6
USR203	12AI	387	363.9	0.24	0.61	0.16	3.75	0.7
USR204	4AI	264	253.3	0.16	0.41	0.20	2.10	0.6
USR204	4BP	357	387.1	-0.28	-0.74	0.16	-4.70	0.9
USR204	8AI	339	361.9	-0.22	-0.57	0.16	-3.60	0.9
USR204	8BP	331	311.3	0.23	0.59	0.17	3.41	0.6
USR204	12AP	312	297.3	0.18	0.46	0.18	2.61	0.6
USR204	12BI	150	138.5	0.29	0.78	0.26	2.98	0.5
USR205	4BI	337	324.1	0.15	0.38	0.17	2.20	0.7
USR205	4BP	394	366.3	0.27	0.70	0.16	4.42	1.2
USR205	8AI	285	316.1	-0.35	-0.91	0.17	-5.34	0.8
USR205	8AP	194	204.8	-0.15	-0.38	0.19	-2.02	0.8
USR205	8BP	278	290.2	-0.16	-0.42	0.19	-2.26	0.9
USR205	12BP	156	148.6	0.21	0.52	0.26	1.97	0.6
USR206	4AI	190	179.4	0.20	0.50	0.22	2.30	0.5
USR206	4AP	194	183.5	0.20	0.52	0.22	2.33	1.9
USR206	8AP	288	314.8	-0.25	-0.65	0.16	-4.14	0.7
USR206	8BP	272	242.9	0.43	1.12	0.20	5.70	0.5

USR206	12AP	311	331.7	-0.21	-0.55	0.16	-3.37	0.5
USR207	8AI	279	260.0	0.27	0.69	0.19	3.61	1.3
USR207	8AP	250	262.9	-0.16	-0.40	0.18	-2.29	0.9
USR207	8BI	317	301.4	0.21	0.54	0.19	2.91	0.9
USR207	8BP	286	270.0	0.22	0.58	0.19	3.04	0.9
USR207	12BI	119	131.2	-0.32	-0.83	0.26	-3.20	0.5
USR207	12BP	190	205.6	-0.31	-0.78	0.22	-3.48	0.5
USR208	4BP	257	282.6	-0.34	-0.87	0.18	-4.74	1.1
USR208	8AP	174	153.6	0.39	1.01	0.22	4.55	0.5
USR208	8BP	286	274.3	0.16	0.41	0.19	2.18	1.0
USR208	12BP	140	153.5	-0.36	-0.91	0.26	-3.50	0.6
USR209	12AP	192	178.8	0.23	0.60	0.21	2.82	1.4
USR209	12BI	126	117.5	0.23	0.60	0.26	2.26	0.8
USR210	4BP	443	414.5	0.27	0.67	0.15	4.38	0.9
USR210	8AP	350	325.8	0.22	0.57	0.15	3.71	0.9
USR210	8BP	302	338.4	-0.41	-1.06	0.17	-6.20	1.0

Bias calibration report for HK rater-prompt interactions

HK rater	Prompt	Observed score	Expected score	Obs-Exp average	Bias (logit)	Error	z-score	Fit MnSq
HKR001	4BI	240	273.1	-0.37	-1.00	0.17	-5.74	1.2
HKR001	4BP	311	271.9	0.44	1.14	0.17	6.67	1.4
HKR002	4BN	332	318.5	0.15	0.39	0.17	2.29	1.4
HKR002	4BP	280	300.6	-0.23	-0.61	0.17	-3.54	1.3
HKR003	4BI	320	340.3	-0.23	-0.59	0.17	-3.46	1.9
HKR003	4BN	359	344.0	0.17	0.44	0.17	2.58	1.7
HKR004	4BN	265	281.7	-0.18	-0.49	0.17	-2.84	1.4
HKR004	4BP	287	274.2	0.14	0.37	0.17	2.18	1.8

Appendix 23

Student 660 (HK021)

4BI My Lunch Time (Human ratings: HKR003-1 and HKR004-2)

Describe what lunch time is like for you on a school day. Be sure to tell about your lunch time so that someone who has never had lunch with you on a school day can understand where you have lunch and what lunch time is like.

Our class always have our yummy lunch at noon with our class teacher.

Today, when all my classmates went to corridor to get their lunch box, I couldn't find my lunch box. "Miss Chan, I can't find my lunch box!" I shouted because I was very hungry now. "Ok. Don't be so angry. Just clam down and call your parents to give you the lunch box." Miss Chan said.

I went to teacher's room and called my mum. "Oh! Mum, you haven't given me the lunch box!" "O....I forget this one. Ha ha ha," said the silly Mum.

I went back to my classroom and sat down sadly. All the classmates were eating tastily. I was really hungry!

"Hi, Henry! Why don't you eat your lunch?" Peter, the most naughty boy in my class, pointed at me and said loudly.

All my classmates looked at me and laughed at me. I wanted to cry at this moment!

"Henry, don't be worried!" said Miss Chan kindly. "Why! Why! I was hungry and there hasn't my lunch box everywhere!" I shouted

"Look! Henry." said Miss Chan. I looked at the corridor and saw my lunch box!

"How! Happy!" I was so happy that I can eat the lunch. Tasty!

I found that I ate my lunch when I was very hungry. The lunch was the tastiest in the world!

4BN The Very Unusual Day (Human ratings: HKR001-5 and HKR002-5)

*One day you wake up and go down to breakfast. You eat what you normally eat. Your breakfast is the last normal thing that happens to you all day. Write a story called '*The Very Unusual Day*' about what happens that day, from right after breakfast until you go to bed again.*

One day, I woke up early and went down to have breakfast. My breakfast was normal and I thought it was tasty.

After eating the normal breafast, I started to went to school. When I wanted to go out, I discovered that the door was locked. I shouted loudly "Mum, why is the door locked?" But no one replied me. I was so frightened. I started to find my mum but I couldn't find her anywhere!

I rushed and hit the door. The door opened! But nothing was behide the door. It only a wall! "Oh my god! A ghost was at my home!" I thought.

I stayed at home frightened until afternoon. The wall became a street, my street! I felt happy and went out but I quickly saw that the people were horrible. They killed others and became bigger and bigger. One people rushed to me. It's my mum! She was so big and ugly. She hit me and kicked me. She said "Silly boy! You opened the door whole day! Our things were stolen by the thifves!" I hit my head and saw all the things became normal. My home is emty!

At night, I went to bed. I didn't know what happened this day. Suddenly, the emty home was on fire! I went out home and my house fell. My mum was died in this fire.

I didn't understand what happened and went to my grandpa's house......

4BP My Favourite Book (Human ratings: HKR001-4 and HKR003-6)

Your favourite book is missing from your school library. It might be a book that you like to read over and over again. Or it might be a book that your teacher or parent has read to you. Some of your friends also like to read this book. The school librarian is not sure she wants to buy the book again. Write a letter to convince your school librarian to buy the book again. In your letter, give lots of reasons why the book should be in your school library.

School librarian:

I have my favourite book 'A Super Student', I want to introduse this book to you. It might be the most famous book of all books in your library. At this moment, I want to show you this.

The book 'A Super Student' is about a student called Tom. He was the cleverest and the most polite student in his school. This book's writer wanted to show how to be good in his reader's school from his book 'A Super Student'.

I have read this book many times. And I had remommended this book to my friends. They all said it's good. I think this book is for all the students. It is funny. You can learn a lot of things too.

I want you to buy this book in your library. Would you mind buying it? I think most of our school's students will like it!

6A Henry Ng

Student 628 (US579)

4AI Describe a favorite object (Human ratings: USR204-4, USR205-4, USR207-3 and USR208-4)

We all have favorite objects that we care about and would not want to give up. Think of one object that is important to you. For example, it could be a book, a piece of clothing, a game, or any object you care about. Write about your favorite object. Be sure to describe the object and explain why it is valuable or important to you.

I like to fire guns. My family has passed it on for generation. I love first everyone of them. If you fire the double triger it will blow you away. Won time I wanted to do it, but if I did I would be hurt pretty bad. Did you know thes used a gold triger in the war? They did your won in the war and I love a god triger. It my most favorite gun yet. It kicks a little, but it will not hurt you any.

I keep my guns in my room. My hole family is youseing them. I keep them on the shelf in my bedroom. We hunt bill, and practive with these guns. With my little gun I have killed a lot of deer. My gun is a sniper bun. We all yous them for good reasons.

4AN Write a story about a castle (Human ratings: USR101-5, USR102-4, USR103-4 and USR105-4)

One morning a child looks out the window and discovers that a huge castle has appeared overnight. The child rushes outside to the castle and hears strange sounds coming from it. Someone is living in the castle! The castle door creaks open. The child goes in. Write a story about who the child meets and what happens inside the castle.

When I walked in the castle I saw santa. He was sitting on his couch relaxing. Of course he was drinking milk, and he was eating some cookies. I walked in his work room and elfs were working crazey. They had some good imagination. I walked back to santa's room and said, "Where are your deer." He said, "They were in the stable." I went looking for deer but I couldn't see them. It's like they dissappeared. I couldn't see them anywhere. I told santa about it. He said, "It would be okay."

I was out on the snow mobil. It was a merical that I saw santa, and his deer loading the sled ful of stuff. I will tell you what I saw. I saw light shoes, little bells, little trucks, skis, tubes, legos, ice sled, and little fake ice sculphures. Santa was taking off in his sled. My friend's and I went with him. Santa droped us off and we all went to sleep. The next morning we got up, and we all opened gifts.

4AP Convince a friend to become visible (Human ratings: USR201-1, USR202-1, USR206-1 and USR210-1)

Pretend you have a friend who is invisible and you would like other people to meet him or her. Write a letter to this invisible person. Convince your friend to become visible so that others may meet him or her. In your letter, use details and examples.

1717 North F. Dr. F
Corinth, MS   38834
11-15-05

Dear Stewart,

At the starting of this weekend my grandma picked me up. We played at me cousins house a lot. Then my dad got my sister and I. We went to the house. I had a couple bad grades. I sit down and I watch tv. We put pizza in the oven and ate it. We went to bed at night. That morning we got up and watched mor tv.

My dad and I got in the truck to get parts. We got the trailer, and loaded two four wheelers up. We got the parts, and we went over to our grandma's house. We would work on the four wheelers, then we would try them out. We worked on them for a couple of days. Then mine finely started working. That Saturday we went on two mazes rides. The first one got me lost. The second one got my sister lost. It was at night. Then we got in the bed, and we went to sleep.

We woke up and it was exactly 5:00 in the morning. My mom got up. We went to the donut shop to get donuts and cholate milk. We went back home and we ate donuts. Everybody got up. We all took a shower, then we got dressed. We went to church. Then we got out. Then my dad was in his coliner, and mom was in her coliner. We went to bed woke up, and we went to school, and we went to work. It should had been a good day for us.

Sincerely,
Robert Briggs

4BI Describe lunchtime (Human ratings: USR204-4, USR205-4, USR208-4 and USR209-4)

Describe what lunchtime is like for you on a school day. Be sure to tell about your lunchtime so that someone who has never had lunch with you on a school day can understand where you have lunch and what lunchtime is like.

The lunch time we have is great. When we get in the cafeteria we are starving. We set down on the fresh table. Where a specail class. Everyday we get a lot of food.

When you look in the back there materals look like they are very old. When we get done the janitors clean our trays. When we get out of the cafeteria were slam full

I like the cafeteria because the lunch ladies give us a lot of food to eat. Mostly we get alot because, we are the last class to eat everyday. The best think I like is the lasauna. They make everything so good and tasty. I know all the lunch ladies very well. They start the cooking right after right after breakfast. That's why you ort to come eat with us. We are the most specail class do come eat with us.

4BN Describe very unusual day (Human ratings: USR101-5, USR103-5, USR104-4 and USR105-5)

One day you wake up and go down to breakfast. You eat what you normally eat. Your breakfast is the last normal thing that happens to you all day. Write a story called '*The Very Unusual Day*' *about what happens that day, from right after breakfast until you go to bed again.*

I was in breakfest that day with food in front of me. Then something went wrong with me. Everything was dissapearing into nothing. I turned around to see what was happening. Nothing was happening then. I turned back around to eat, and it wasn't there. So I got up, and payed that man some money. I said, "Something is wrong with your place." He said, "Nothing is wrong with my place." I didn't want to argue with him. So I started walking out of the store and the clock started to ddissapeard. I left the building emeadeyatly.

I got in my truck cranked it up to go somewhere. I was going on the road. My tires dissapeared out of nowhere. My seat was gone. Then the truck quit on me. I got out of the truck. I was looking around. I was saying in my head, "Where an I." I turned back to my truck. It wasn't there any more. I called somebody to come get me. They picked me up and they carried me home. I made an appointment with Mr. Perry. I got there right on time. He stepped in the room. All the stools dissapeard, exrays, gloves, glasses, boxes, computer, cushens were just all gone. He started to beleive my story. I went home to get in bed. I went to sleep in the bed. The next day I woke up there was nothing wrong.

4BP Convince the school librarian (Human ratings: USR202-6, USR203-5, USR206-5 and USR207-5)

Imagine this situation: Your favorite book is missing from your school library. It might be a book that you like to read over and over again. Or it might be a book that your teacher or parent has read to you. Some of your friends also like to read this book. The school librarian is not sure she wants to buy the book again. Write a letter to *convince your school librarian to buy the book again. In your letter, give lots of reasons why the book should be in your school library.*

1715 North Fulton Dr
Corinth, MS 38834
December 8, 2005

My book is Where the Red Fern Grows. It is a good story. It sometimes make people cry a lot. Even thoe the boy dog got killed by a lion in the foest. But it wasn't necessary fore Ann not to eat anything. The reason she wasn't going to eat because Ann didn't have Dan right beside her. At the end it is very sad when they bothe died.

The reason you should buy more of the books is because they so grematic. They tell how he gets there. It's hunting mostly in the story. You publish more because it tales us how the adventure was. It was the beautiful thing you couled have ever seen. Over big Dan and little Ann graves there were to red ferns. That why I think there should be more books like that.

Your friend,
Robert Briggs

Student 69 (US057)

4BI Describe lunchtime (Human ratings: USR201-5, USR203-3, USR205-4, USR206-4, HKR002-3 and HKR004-6)

Describe what lunchtime is like for you on a school day. Be sure to tell about your lunchtime so that someone who has never had lunch with you on a school day can understand where you have lunch and what lunchtime is like.

I would tell them what where having for lunch. I would let thme pick what they want. I would tell them what time it would be. I would say what lunchtim was. The lunchtime is at 12:30. When it gets to the time we go to the lunchroom. We would get what we ordered from the lunchladys. We would go to the lunchtable. We would get our drinks. We would open the drinks taht we got. I would get my dessert that I got. We would go to the lunchtable. We would fix our food. We would pray before we eat. I would taste one bite of everything. I woudn't eat my paper or my carten of drinks. They would taste there milk or choclate milk. If the food was good I would eat it all. After we get done eating our teacher would let us wisper real soft. Thats why I call it "The Lunchtime and Lunch."

4BN Describe very unusual day (Human ratings: USR102-3, USR103-3, USR104-3, USR105-3, HKR001-3 and HKR004-2)

One day you wake up and go down to breakfast. You eat what you normally eat. Your breakfast is the last normal thing that happens to you all day. Write a story called '*The Very Unusual Day*' *about what happens that day, from right after breakfast until you go to bed again.*

One morning I wook up and I seen an elaphant on the kitchen table eating breakfast. I walked by my grandmother's house and seen a man with an elephant head. I went to school for beakfast. There were animals there, everwhere! I went back to my house and nobody was there. After I went to my bedroom. I slipped on my pajanes, I layed down on my bed and watched t.v. I went after all of that . It really was my unusual day.

4BP Convince the school librarian (Human ratings: USR200-1, USR203-1, USR208-1, USR209-1, HKR002-1 and HKR003-1)

Imagine this situation: Your favorite book is missing from your school library. It might be a book that you like to read over and over again. Or it might be a book that your teacher or parent has read to you. Some of your friends also like to read this book.

The school librarian is not sure she wants to buy the book again. Write a letter to convince your school librarian to buy the book again. In your letter, give lots of reasons why the book should be in your school library.

41 Forest School RD.
Mississippi, 38834
Corinth, Mississippi

Dear parent's.

I am writing to you. I want to tell you I am doing good in school. My grandparents are doing well to. They never fussed ontime. they don't argue anymore. We look like a family I gese.

ps, I love you, I wish I could see ou. I really mis you. I have to go to bed now. goodnight.

Sincerely Ashley Nuntey

Student 396 (US366)

8BI Which book would you save? (Human ratings: USR201-5, USR204-6, USR207-6 and USR209-5)

A novel written in the 1950's describes a world where people are not allowed to read books. A small group of people who want to save books memorize them so that the books won't be forgotten. For example, an old man who has memorized the novel The Call of the Wild helps a young boy memorize it by reciting the story to him. In this way, the book is saved for the future. If you were told that you could save just one book for future generations, which book would you choose? Write an essay in which you discuss which book you would choose to save for future generations and what it is about the book that makes it important to save. Be sure to discuss in detail why the book is important to you and why it would be important to future generations.

I would have to chose Memory, Sorrow, and Thorn by Ted Williams. It's underrated because it is so long, about 700 pages a book, and four books. So at first glance, it would seem that the book would be one long, drawn out story. The thing though, is that he is a great writer and keep your attention for days on and easily since he tends to touch on imaginative subjects in realistic ways. Also the book potrays a medeval Earth with basically everything changed that was involved in religion. Memory, Sorrow, and Thorn has a close grasp on human nature and the spread of imagination, and if imagination or ignorance is absent,you can expect less of people. It was one of the first books that I actually delved in, so it means Quite a bit. Since then, I've reread it about three times and I find something new each time. Now that's something to marvel at, I swear that as you change it will change with you. Memory, Sorrow and Thorn is actually also filled with knowledge. From old ways, military tactics, the way people act, culture, etc. You see most of this through a growing teenager, which I suppose shows a sense of growth for the book itself. A book is a thing of vast knowledge, a compendium of thoughts, codex of dreams, and guide to life. It's just what is is and who for what reson is varied among people. They derive their own essence of meanings, because we're equally unequal in a paradox of weaves.

8BN Your first day as president (Human ratings: USR101-5, USR102-5, USR104-4 and USR105-5)

Imagine that you wake up one morning to discover that you have become the President of the United States. Write a story about your first day as President.

First things first.. holy crap. I'm the President!? I mean that's like you wake up, rip your shirt off, and say "I'm ....Antonio Bandaras!" It just doesn't happen, a reality turned surreal sort of. That's not the case though, and somehow magically I am the Prez, so let's explore my day.

I wake up on just another busy day (because I'm the President of the United States) even though this seemed odd since ther was no deja` vu. I dressed myself accordingly while I ran through the tasks for today and grimaced at such daunting tasks that lay ahead of me.

"Nuclear threat!?" I screamed, "Just my luck on my first day ... wait a second, that's why I didn't have deja`vu`. Anyways, I need to talk about this."

So I hurriedly moved my way towards the person I should be in contact with, then I thought to myself, "I can just call them here," so I did.

"Russell, get yourself down here and explain to me just what the heck is going on," I said with a twinge of annoyance.

"What do you mean Mr. President?" he asked.

"About this 'nuclear threat' business, I should have known about this already," I yelled.

"sir, look at the date of the inquiry," Russell said calmly.

"Oh, that's really not anywhere close to today's date. Thank you Russell," then I jung up mainly to embaressment.

Everything just went smoothly after that. Every day order of business I had to deal with, just another day they say. After all the work had been dealt with and I could relax in the privacy of my room, I fell asleep. When I awoke the next morning, I was no longer the President of the United States.

8BP Support a school schedule (Human ratings: USR203-5, USR206-5, USR208-4 and USR210-3)

Suppose a research study showed that teenagers have low energy levels in the morning and that adults have low energy levels at night. The study recommends that teenagers should stay up later at night and sleep later in the morning. The study also recommends that adults go to bed earlier and get up earlier. Write a letter to your principal arguing for or against the proposition that classes at your school should begin much later in the day. Be sure to give detailed reasons to support your argument and make it convincing.

OK first things first, later in the day just means we go later. It's the same time frame, same amount of sleep, so why do it? If anything, you'll just get negative feedback concerning the tradition of our school time. Seriously though, quite a idea even with the statistics.

I myself sleep earlier than most and can wake at nearly the crack of dawn. Most teenagers are built with an internal alarm clock. Some of them will sleep through a World War, yes. That's not really the point, the point is should it be later in the day or not and I say no.

To work back in a sense, adults don't necessarily sleep early too. The statistics would be off bottom line, each person is unique to their routines. This is getting thin, so I will probably have to just say, no. Don't even think these things, it makes no logical sense.

My future, is coming on. As the Gorillas say, which makes a good amount of sense when you think about it. Such as the changes of what is viewed as normal. This is my closing, countering a retort of ignorance.

12AI Give writing advice to a younger student (Human ratings: USR203-4, USR205-4, USR207-4 and USR209-4)

Your school has a program in which a twelfth grader acts as a mentor for a tenth grader at the beginning of each school year. The mentor's job is to help the tenth grader have a successful experience at your school. The tenth grader you are working with is worried about being able to write well enough for high school classes. Write a letter to your tenth grader explaining what kind of writing is expected in high school classes and what the student can do to be a successful writer in high school. As you plan your responses, think about your own writing experiences. How would you describe 'goo' writing? What advice about writing has been helpful to you? What writing techniques do you use?

If you want to write well in high school classes, remember at least this one thing: You write on a specific subject. Deviation or being deviant in any way as well, will pretyy much assure the bad outcome. Along with that, remember to follow within the parameters usually provided or said, otherwise use common sense.

When you write something like a story, keep in mind it's for others to read, not personal reading. Saying that, try to write in a way that can excise or entice people in a thralling compulsion of magnificence. Don't ever leave it boring and dull so people are disinterested with it. Now if the writing is just telling about a certain thing, it's more open to change. Information is key to when writing these, keep it simple and understandable. Combined, or just general, it's best to do pre-writing. This helps ascertain the concept on which you write. When you don't, more than likely it'll take longer, and be quite sprawled out. Just keep writing, train of thought is very important. If you lose it, it takes ahile to retrieve the thoughts. To be quet accomplished in writing during high school just remember some of this. All of this in the back of your mind will do well. High school standards are sometimes different than others, so like I said, follow parameters. Finally, just keep it at. Like I say "If at first you don't succed, fail then start fresh."

12AN Write a story about a special object (Human ratings: USR101-2, USR104-1, USR105-1 and USR106-1)

The following excerpt is from a poem by Walt Whitman.
There was a child who went forth every day,
And the first object he look'd upon, that object he became,
And that object became part of him for the day or a certain part of the day,

Or for many years or stretching cycles of years. Whitman's poem suggests that certain objects become important to us and remain important to us even if we no longer have them. Write a story in which you tell about an object that remains important to the main character over a period of years. The main character could be you or someone you know. In your story, describe the main character's first encounter with the object, why the object is so important to the character, and how, over the years, it remains a part of the character's life.

Once upon a time I felt very sad during the particular day on which we do these Lexile papers. Then I thought to myself, "Wow, I've done pretty well so far, but today's paper is just not gonna happen. I only need enough writing here to fill up this and part of the second to seem like I'm accomplished for the day." So in conclusion, if you've read this far, I'm just gonna do it right now.

12AP Debate the importance of voting (Human ratings: USR202-4, USR205-4, USR207-4 and USR209-2)

Your school is sponsoring a voter registration drive for 18-year-old high school students. You and three of your friends are talking about the project. Your friends say the following,

Friend 1: ‘I’m working on the young voters’registration drive. Are you going to come to it and register? You’re all 18, so you can do it. We’re trying to help increase the number of young people who vote and it shouldn’t be too hard - I read that the percentage of 18- to 30-year-olds who vote increased in recent years. We want that percentage to keep going up.’

Friend 2: 'I'll be there. People should vote as soon as they turn 18. It's one of the responsibilities of living in a democracy.'

Friend 3: 'I don't know if people should even bother to register. One vote in an election isn't going to change anything.'

Do you agree with friend 2 or 3? Write a response to your friends in which you explain whether you will register to vote. Be sure to explain why and support your position with examples from your reading or experience. Try to convince the friend with whom you disagree that your position is the right one.

I agree with the second friend because partly that is right. Now, a responsibility of living in democracy is a bit off since you're not forced to vote; therefore you're not liable to do so. Voting though promotes a sense of attachment. A feeling of accomplishment might hit you too, it just depends on the situation. Oh, one vote doesn't count huh? Well let's see. Was it Winston Churchill who had one more? Yeah, every little person counts. We live where all men are equal, yet not so to the prequel. I've seen and heard thing's other than Churchill, that makes your vote go where it should. If you don't like a president for one thing, and right now Bush is the most hated, then vote against him next time. Basically, don't just complain about problems you help decide, do something about it. Swaying people towards one thing is a cherishable skill lies within leaders. Very now and then though people get to become their own leader in sorts, self taught majesty. That's when you can feel like it makes a difference, that's when any vote will count. So friend number three over there can just suck up and be the one who decides in his/her head. Knowing this gets you along enough for whatever's needed.

So, the government lies to us, we're all crooked in some way, and so on and henseforth, voting is a given right to us, we can do something about is, as I keep saying. Remember all the events of the past and maybe that summons the populate. It's just eh pulchritudinous of it all. A democratic county with quelling voices that reach through our TV's..

Student 725 (HK086)

4BI My Lunch Time (Human ratings: HKR003-1 and HKR004-1)

Describe what lunch time is like for you on a school day. Be sure to tell about your lunch time so that someone who has never had lunch with you on a school day can understand where you have lunch and what lunch time is like.

Dear SAKURA:

Hello, How are you? My name is Michael, is your schoolmarse, I like eating Lunch for 12:30 O'clock, I like lunch at school cantan.

Love,

Michael Lee

4BN The Very Unusual Day (Human ratings: HKR001-1 and HKR002-1)

One day you wake up and go down to breakfast. You eat what you normally eat. Your breakfast is the last normal thing that happens to you all day. Write a story called 'The Very Unusual Day' about what happens that day, from right after breakfast until you go to bed again.

Write a story called "The Very Unusual Day" about what happens that day, from right after breakfast until you go to bed again. One day, I wake up and go down to breakfast. I eat what you normally eat. Your breakfast is the last normal thing that happens to you all day

4BP My Favourite Book (Human ratings: HKR001-2 and HKR003-6)

Your favourite book is missing from your school library. It might be a book that you like to read over and over again. Or it might be a book that your teacher or parent has read to you. Some of your friends also like to read this book. The school librarian is not sure she wants to buy the book again. Write a letter to convince your school librarian to buy the book again. In your letter, give lots of reasons why the book should be in your school library.

Dear school librarian:

I was missing a favourite book from your school library, this book name is 'FULLMETAL ALCHEMIST' this book have many people to reading, this book my teachers like reading too, so you wants to buy the book again.

Appendix 24

Calibration of the HK raters for HK student essays and a sub-sample of US student essays using HK raters

HK rater	Observed score	Measure (Logit)	S.E.	Infit		Outfit	
				MnSq	ZStd	MnSq	ZStd
HKR001	3.1	0.35	0.08	0.88	-1.4	0.88	-1.4
HKR002	3.5	-0.13	0.08	0.92	-0.9	0.93	-0.8
HKR003	3.8	-0.83	0.08	1.11	1.2	1.12	1.3
HKR004	3.0	0.61	0.08	1.09	1.0	1.06	0.7
Separation: 7.03			Reliability: 0.98				
Fixed (all same) chi-square: 201.8			significance p: .00				

Appendix 25

```
+--------------------------------------------------------------------------------------------------------------+
|Measr|-Prompt                  |-Rater                                          |-Lexile|-IEA|+Student     |R6SCA|
|-----+-------------------------+------------------------------------------------+-------+----+-------------+-----|
|   5 +                         +                                                +       +    +             + (6) |
|     |                         |                                                |       |    |             |     |
|     |                         |                                                |       |    |             |     |
|   4 +                         +                                                +       +    + .           +  5  |
|     |                         |                                                |       |    | .           |     |
|     |                         |                                                |       |    |             |     |
|   3 +                         +                                                +       +    + .           +     |
|     |                         |                                                |       | 1  | .           | --- |
|     | 1-12AI  3-12AP          |                                                |       |    | .           |     |
|   2 + 1-12BI                  +                                                +       + 2  + **.         +     |
|     | 3-8AP                   |                                                |       |    | ***.        |     |
|     |                         |                                                | 1     | 3  | *****.      |  4  |
|   1 + 2-12BN  3-12BP  3-8BP   + US104  US200                                   + 2     + 4  + ******.     +     |
|     | 2-12AN  2-8BN           | HK004  US103  US209  US211                     |       |    | *********.  |     |
|     | 1-8AI   1-8BI           | HK001  US102  US105  US206                     | 3 4   |    | *******.    |     |
*   0 * 2-8AN                   * US101  US203  US204  US205                     * 5     * 5  * *********.  * --- *
|     |                         | HK002  US106  US201  US207  US208  US210       | 6     |    | *********.  |     |
|     |                         | US202                                          | 7 8   | 6  | ***.        |     |
|  -1 +                         +                                                +       +    + *****.      +     |
|     |                         | HK003                                          | 9     | 7  | **.         |  3  |
|     | 2-4BN                   | US108                                          |       |    | **.         |     |
|  -2 + 1-4BI   3-4BP           +                                                +       + 8  + *.          +     |
|     | 2-4AN                   |                                                |       |    | *.          |     |
|     | 1-4AI   3-4AP           |                                                |       |    | *           | --- |
|  -3 +                         +                                                +       + 9  + *.          +     |
|     |                         |                                                |       |    | .           |     |
|     |                         |                                                |       |    | .           |     |
|  -4 +                         +                                                +       +    + .           +  2  |
|     |                         |                                                |       |    |             |     |
|     |                         |                                                |       |    | .           |     |
|  -5 +                         +                                                +       +    + .           +     |
|     |                         |                                                |       |    | .           | --- |
|     |                         |                                                |       |    | .           |     |
|  -6 +                         +                                                +       +    + .           +     |
|     |                         |                                                |       |    | .           |     |
|     |                         |                                                |       |    | .           |     |
|  -7 +                         +                                                +       +    +             +     |
|     |                         |                                                |       |    | .           |     |
|     |                         |                                                |       |    |             |     |
|  -8 +                         +                                                +       +    + .           + (1) |
|-----+-------------------------+------------------------------------------------+-------+----+-------------+-----|
|Measr|-Prompt                  |-Rater                                          |-Lexile|-IEA| * = 10      |R6SCA|
+--------------------------------------------------------------------------------------------------------------+
```

Vertical map for all student essays with scoring of US, HK raters and Lexile Analyzer and the new measure of IEA

```
+-------------------------------------------------------------------------------------------------------------+
|Measr|-Prompt                  |-Rater                                    |-Lexile|-IEA|+Student     |R6SCA|
|-----+-------------------------+------------------------------------------+-------+----+-------------+-----|
|   5 +                         +                                          +       +    +             + (6) |
|     |                         |                                          |       |    | .           |     |
|     |                         |                                          |       |    | .           |     |
|   4 +                         +                                          +       +    +             +  5  |
|     |                         |                                          |       |    |             |     |
|     |                         |                                          |       |    |             |     |
|   3 +                         +                                          +       +    + .           +     |
|     |                         |                                          |       | 1  | .           | --- |
|     | 1-12AI  1-12BI  3-12AP  |                                          |       |    | *           |     |
|   2 +                         +                                          +       + 2  + **          +     |
|     | 3-8AP                   |                                          |       |    | ***.        |     |
|     |                         |                                          | 1     | 3  | *****.      |  4  |
|   1 + 2-12BN  3-12BP  3-8BP   + HK004  US103  US104  US200  US211        + 2     + 4  + ******.     +     |
|     | 2-12AN  2-8BN           | US209                                    | 3     |    | ********.   |     |
|     | 1-8AI   1-8BI           | HK001  US102  US105  US206               | 4     |    | *******.    |     |
*   0 * 2-8AN                   * US101  US203  US204  US205               * 5     * 5  * *********.  * --- *
|     |                         | HK002  US106  US201  US208  US210        | 6     |    | ********.   |     |
|     |                         | US202  US207                             | 7 8   | 6  | ***.        |     |
|  -1 +                         +                                          +       +    + ****.       +     |
|     |                         |                                          | 9     | 7  | **.         |  3  |
|     |                         | HK003  US108                             |       |    | *.          |     |
|  -2 + 1-4BI   2-4BN   3-4BP   +                                          +       + 8  + *.          +     |
|     | 2-4AN                   |                                          |       |    | *.          |     |
|     | 1-4AI   3-4AP           |                                          |       |    | *.          | --- |
|  -3 +                         +                                          +       + 9  + .           +     |
|     |                         |                                          |       |    | *           |     |
|     |                         |                                          |       |    | .           |     |
|  -4 +                         +                                          +       +    + .           +  2  |
|     |                         |                                          |       |    | .           |     |
|     |                         |                                          |       |    | .           |     |
|  -5 +                         +                                          +       +    + .           +     |
|     |                         |                                          |       |    | .           | --- |
|     |                         |                                          |       |    | .           |     |
|  -6 +                         +                                          +       +    + .           +     |
|     |                         |                                          |       |    | .           |     |
|     |                         |                                          |       |    | .           |     |
|  -7 +                         +                                          +       +    +             +     |
|     |                         |                                          |       |    |             |     |
|     |                         |                                          |       |    | .           |     |
|  -8 +                         +                                          +       +    +             +     |
|     |                         |                                          |       |    |             |     |
|     |                         |                                          |       |    | .           |     |
|  -9 +                         +                                          +       +    + .           + (1) |
|-----+-------------------------+------------------------------------------+-------+----+-------------+-----|
|Measr|-Prompt                  |-Rater                                    |-Lexile|-IEA| * = 10      |R6SCA|
+-------------------------------------------------------------------------------------------------------------+
```

Vertical map for all student essays with scoring of US, HK raters and IEA and the new measure of Lexile Analyzer