Social Sciences and Humanities Research Council of Canada

Conseil de recherches en sciences humaines du Canada

Application for a Grant

Internal use
486408

Identification

This page will be made available to selection committee members and external assessors.

Funding opportunity
Insight Development Grant

Grant type (Strategic Grants only)
Individual

Application title
On the Classification Consistency and Accuracy of Cognitive Diagnostic Assessment

Applicant family name	Applicant given name	Initials
Cui	Ying	

Org. code	Full name of applicant's organization and department
1480111	University of Alberta Educational Psychology

Org. code	Full name of administrative organization and department
1480111	University of Alberta Educational Psychology

Scholar type: Regular ○ New ◉

If New, specify category: 1 ◉ 2 ○ 3 ○ 4 ○

Research Field: 43005

Does your proposal require a multidisciplinary adjudication? Yes ○ No ◉

Does your proposal involve human beings as research subjects? If "Yes", consult the *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans* and submit your proposal to your organization's Research Ethics Board. Yes ○ No ◉

Does your proposal involve activity that requires a permit, licence, or approval under any federal statute; or physical interaction with the environment? If 'Yes', complete Appendices A and B. Yes ○ No ◉

	Year 1	Year 2	Total
Total funds requested from SSHRC (from page 6)	16,723	44,503	61,226

Personal information will be stored in the Personal Information Bank for the appropriate program.

PROTECTED WHEN COMPLETED

Canada

Family name, Given name
Cui, Ying

Participants

List names of your team members (co-applicants and collaborators) who will take part in the intellectual direction of the research. Do not include assistants, students or consultants.

Role
Co-applicant (●) Collaborator ( )

Family name	Given name	Initials
Gierl	Mark	J.

Org. code	Full organization name
1480111	University of Alberta

Department/Division name
Educational Psychology

Role
Co-applicant ( ) Collaborator ( )

Family name	Given name	Initials

Org. code	Full organization name

Department/Division name

Role
Co-applicant ( ) Collaborator ( )

Family name	Given name	Initials

Org. code	Full organization name

Department/Division name

Role
Co-applicant ( ) Collaborator ( )

Family name	Given name	Initials

Org. code	Full organization name

Department/Division name

Role
Co-applicant ( ) Collaborator ( )

Family name	Given name	Initials

Org. code	Full organization name

Department/Division name

Family name, Given name
Cui, Ying

Research Activity

The information provided in this section refers to your research proposal.

Keywords

List keywords that best describe your proposed research or research activity. Separate keywords with a semicolon.

Large-scale assessment; Cognitive diagsnotic assessment; Classification consistency; Classification accuracy; Real data application; Software development

Priority Areas - Priority area most relevant to your proposal.

Digital Media

Disciplines

Indicate and rank up to 5 disciplines that best correspond to your proposal.

Rank	Code	Discipline	If "Other", specify
1	61232	Measurement and Evaluation	
2	63020	Psychometrics	
3			
4			
5			

Areas of Research

Indicate and rank up to 3 areas of research related to your proposal.

Rank	Code	Area
1	140	Education
2	382	Youth
3	111	Children

Temporal Periods

If applicable, indicate up to 2 historical periods covered by your proposal.

From	To
Year _____ BC ◯ AD ◯	Year _____ BC ◯ AD ◯
_____ BC ◯ AD ◯	_____ BC ◯ AD ◯

Personal information will be stored in the Personal Information Bank for the appropriate program.

Application WEB

PROTECTED WHEN COMPLETED

Canada

Family name, Given name
Cui, Ying

Research Activity (cont'd)

Geographical Regions

If applicable, indicate and rank up to 3 geographical regions covered by or related to your proposal. Duplicate entries are not permitted.

Rank	Code	Region
1		
2		
3		

Countries

If applicable, indicate and rank up to 5 countries covered by or related to your proposal. Duplicate entries are not permitted.

Rank	Code	Country	Prov./ State
1			
2			
3			
4			
5			

Family name, Given name
Cui, Ying

Summary of Proposed Research

The summary of your research proposal should indicate clearly the problem or issue to be addressed, the potential contribution of the research both in terms of the advancement of knowledge and of the wider social benefit, etc.

The 21st century is characterized as the era of the knowledge-based economy due to rapid changes occurring in technologies and advanced computer-based systems. According to the Adult Literacy and Life Skills survey (Statistics Canada and OECD, 2005), however, 40% of Canadian adults do not have literacy skills at "the level considered by experts as a suitable minimum for coping with the increasing demands of the emerging knowledge society and information economy". As the knowledge requirements of Canadians' jobs are growing rapidly, fostering a high quality education system that helps all Canadian students develop the knowledge and skills they need to succeed in the labour market is central to the future success of Canada in the global economy.

Large-scale assessment has become an important method for monitoring student achievement and the quality of educational systems in Canada. Conventional large-scale assessments assign general scores to students on a continuous scale representing the overall amount of knowledge students have acquired within the test domain without specific information about students' strengths and weaknesses that can help teachers design effective instructional interventions. Test scores are typically useful in informing educational decisions such as grade promotion/retention, graduation, eligibility for scholarship, or certifications. However, this type of test results is not helpful in guiding learning at the classroom level. To make costly large-scale assessments more practically useful, cognitive diagnostic assessment serves as an important effort to redesign large-scale assessments so as to improve their diagnostic value in producing more instructionally relevant results that support classroom teaching and learning (cf. Leighton & Gierl, 2007). The cognitive diagnostic assessment approach is aimed at providing students with detailed information regarding whether or not they have mastered each of a set of specific skills measured on the test. These results have the potential to lead to greater utility of large-scale assessment results for informing instructional practices.

The objectives of the proposed research are to (1) evaluate the strengths and limitations of our two new statistical indexes (Cui, Gierl, & Chang, 2011) in applied settings for examining the consistency and accuracy of the results produced by cognitive diagnostic assessments, and (2) implement these new indexes into software that will be disseminated broadly to researchers and practitioners to promote more effective use of cognitive diagnostic assessment. The proposed research will address one of the most fundamental questions in the area of cognitive diagnostic assessments: how consistent and accurate are the classification results produced by a cognitive diagnostic assessment? The importance of this question cannot be overstated because it is directly linked to the future success of cognitive diagnostic assessments. Inconsistent or inaccurate skill diagnosis can cause the misinterpretation of students' skill profiles, which can lead to faulty remediation decisions. This may not only result in a waste of students' and teachers' time and effort, but may also adversely affect students' educational and future employment opportunities. The proposed research will provide researchers and test developers with a useful tool to evaluate the consistency and accuracy of diagnostic results and, therefore, promote the future success of cognitive diagnostic assessments. Through the better integration of testing, teaching and learning, we can provide higher quality education to our students to equip them with the knowledge and skills they need to succeed in the labour market, which ultimately leads to a higher quality of life for Canadians.

Detailed Description of Proposed Research

Objectives

The 21st century is characterized as the era of the knowledge-based economy due to rapid changes occurring in technologies and advanced computer-based systems. According to the Adult Literacy and Life Skills survey (Statistics Canada and OECD, 2005), however, 40% of Canadian adults do not have literacy skills at "the level considered by experts as a suitable minimum for coping with the increasing demands of the emerging knowledge society and information economy". As the knowledge requirements of Canadians' jobs are growing rapidly, fostering a high quality education system that helps all Canadian students develop the knowledge and skills they need to succeed in the labour market is central to the future success of Canada in the global economy. Large-scale assessment has become an important method for monitoring student achievement and the quality of educational systems in Canada. Currently, all 10 Canadian provinces routinely administer large-scale assessments to students at different grade levels. Every year, for example, Alberta Education measures Grade 3 student achievements in language arts and mathematics and Grades 6 and 9 student achievements in language arts, mathematics, science, and social studies. Despite their prevalence, conventional large-scale assessments have been criticized for assigning only general scores to students on a continuous scale representing the overall amount of knowledge students have acquired within the test domain without specific information about students' strengths and weaknesses that can help teachers design effective instructional interventions (e.g., National Research Council [NRC], 2001). New cognitive diagnostic assessments serve as an important effort to redesign large-scale assessments so as to improve their diagnostic value in producing more instructionally relevant results that support classroom teaching and learning (cf. Leighton & Gierl, 2007). Cognitive diagnostic assessments aim to provide students with detailed information regarding whether or not they have mastered each of a set of specific knowledge and skills. These results have the potential to lead to greater utility of large-scale assessment results for informing instructional practices.

The objectives of the proposed research are to (1) evaluate the strengths and limitations of our two new statistical indexes (Cui, Gierl, & Chang, 2011) in applied settings for examining the consistency and accuracy of the results produced by cognitive diagnostic assessments, and (2) implement these new indexes into software that will be disseminated broadly to researchers and practitioners such as test developers to promote more effective use of cognitive diagnostic assessment. This research is aimed to address one of the most fundamental questions in the area of cognitive diagnostic assessments: *how consistent and accurate are the diagnostic results?* The importance of this question cannot be overstated because it is directly linked to the future success of cognitive diagnostic assessments. Inconsistent and inaccurate skill diagnosis can cause the misinterpretation of students' skill profiles, which can lead to faulty remediation decisions. This may not only result in a waste of students' and teachers' time and effort, but may also adversely affect students' educational and future employment opportunities. The proposed research addresses this question by empirical experimentation with new statistical methods for validating the results produced by cognitive diagnostic assessments.

Context

Cognitive Diagnostic Assessment

Cognitive diagnostic assessment is designed to serve as assessment *for* learning rather than assessment *of* learning (Jang, 2008). Current large-scale assessment practices tend to focus on assessment *of* learning and are generally summative in nature. That is, assessments are designed to measure the amount of knowledge a student has acquired within the test domain and assessment results are primarily used to judge the overall sufficiency of student learning. Test scores are typically useful in informing educational decisions such as grade promotion/retention, graduation, eligibility for scholarship, or certifications. However, this type of test results is not helpful in guiding learning at the classroom level. To make costly large-scale assessments more practically useful, the cognitive diagnostic assessment approach is aimed at promoting assessment *for* learning by redesigning large-scale assessments so as to identify each student's specific strengths and weaknesses. Teachers may use these

results to adjust their teaching and remediate students' weaknesses. Cognitive diagnostic assessments are designed to explicitly measure a specific set of skills. For example, a test of fraction subtraction may measure skills such as converting a whole number to a fraction, separating a whole number from a fraction, and simplifying before subtracting (de la Torre & Douglas, 2004), and a reading test may require skills such as remembering details, separating fact from opinion, and speculating from contextual clues (McGlohen & Chang, 2008). To gauge a student's mastery of these skills, many statistical models, referred to as cognitive diagnostic models (CDMs), have been developed by measurement specialists. Recent examples of CDMs include the *DINA (deterministic input noisy and) model* (de la Torre & Douglas, 2004), the *attribute hierarchy method* (Leighton, Gierl, & Hunka, 2004; Gierl, Cui, & Hunka, 2008), the *NIDA (noisy input deterministic and) model* (Junker & Sijtsma, 2001), the *DINO (deterministic input noisy or) model* (Templin & Henson, 2006), the *Fusion model* (e.g., Roussos et al., 2007), and the *hierarchical general diagnostic model* (von Davier, 2007). Rupp and Templin (2008) provide a review of many currently existing models. By using a CDM to identify whether or not each student has mastered the set of specific skills, cognitive diagnostic assessments provide valuable feedback that helps teachers and students plan individualized remediation activities.

The development of cognitive diagnostic assessments to make inferences about student strengths and weaknesses is an exciting new area of research that holds great promise for further improving the educational benefits of large-scale assessments (e.g., Rupp, Templin, & Henson, 2010). Although tremendous progress has been made in the design and analysis of cognitive diagnostic assessments, their current applications are limited and still at an experimental stage. To fulfill their full potential, basic but important questions such as how to examine the reliability and validity of the diagnostic results yielded by the assessment must still be answered. These questions are directly linked to the future success of cognitive diagnostic assessments. During the past five years, my research has been focused on addressing some of these questions by investigating different theoretical and technical aspects of cognitive diagnostic assessment. As a new scholar, I completed my doctoral degree in winter 2007. In my dissertation research, I developed a person-fit statistic called the hierarchy consistency index to help statistically detect ill-fitting item responses for cognitive diagnostic assessment (Cui, 2007; Cui & Leighton, 2009; Cui & Roberts, 2010; Leighton, Cui, & Cor, 2009). In addition to person-fit statistics, I am also interested in other psychometrical aspects of cognitive diagnostic assessments. For example, I have proposed to use the artificial neural network approach to classify students in cognitive diagnostic assessment (Cui & Gierl, 2010; Gierl, Cui, & Hunka, 2008). I also collaborated in developing new procedures for examining the differential item functioning and skill-level reliability for cognitive diagnostic assessment (Gierl, Cui, & Zhou, 2009; Gierl, Zheng, & Cui, 2008). Most recently, I developed two new statistical indexes for evaluating the consistency and accuracy of classification results produced by the cognitive diagnostic assessment (Cui et al, 2011). The research on these two indexes is still in its initial stages as we have not evaluated them with real data yet. Hence, the proposed research addresses the SSHRC Insight Development Grants mandate by empirical experimentation with these new statistical indexes for validating the results produced by cognitive diagnostic assessments.

Theoretical Framework for Classification Consistency and Accuracy

The *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 1999), which were developed to "promote the sound and ethical use of tests and to provide a basis for evaluating the quality of testing practices" (p.1), call for estimates of the percentage of examinees that would be classified consistently using the same or alternate forms for tests with categorical decisions (p.35). Classification consistency is often referred to as the reliability of classifications. Although the definition of classification consistency requires data from repeated testing, such data are rarely available in practice. Thus, classification consistency is normally estimated based on scores from a single test administration. This is typically done by imposing a psychometric model on test data so that the observed score distribution can be estimated and consequently the probability of an examinee being classified into the same category on test occasions can be computed. It should be noted that a high level

of classification consistency does not necessarily imply the classification results are valid and indeed agree with the true nature of the examinees. Classification accuracy is more closely related to the concept of classification validity, examining the degree to which classifications based on observed scores match those based on true scores. To estimate classification accuracy, the true score distribution as well as the observed score distribution need to be estimated so that the probability of an examinee being classified into the true category can be computed.

Many classification consistency and accuracy indexes have been developed. Wilcox (1977), Huynh (1976), Berk (1984), Subkoviak (1984), Hanson and Brennan (1990), Livingston and Lewis (1995), and Lee, Brennan, and Wan (2009) employed the two- or four-parameter beta binomial model for estimating single administration estimates of classification consistency and/or accuracy where the conditional distribution of the total score given the true proportion-correct score is assumed to be binomially distributed. Huynh (1990), Schulz, Kolen, and Nicewander (1999), Wang, Kolen, and Harris (2000), Emons, Sijtsma, and Meijer (2007), and Lee (2010) studied classification consistency and accuracy based on item response theory (IRT) in which the probability of producing a correct response to an item is modeled as a logistic function of item parameters including item discriminating power, difficulty, and guessing. The classification consistency and accuracy indexes developed under the framework of the beta binomial models and IRT are designed for one-dimensional tests in which all items measure the same construct. These tests first produce scores on a generally defined continuous scale (e.g., mathematic ability) and then use cut scores for classifications. Because cognitive diagnostic assessment is designed to measure multiple skills and uses CDMs to classify students, existing indexes based on the beta binomial model and IRT cannot be directly applied.

To fill the gap in the literature on how to exam the classification consistency and accuracy in cognitive diagnostic assessments, we developed two new indexes (Cui et al., 2011): (a) the cognitive diagnostic classification consistency index, and (b) the cognitive diagnostic classification accuracy index. The ***cognitive diagnostic classification consistency index*** gives the probability of classifying a randomly selected student consistently on two administrations or forms of a cognitive diagnostic test, which serves as an index of classification consistency for a cognitive diagnostic assessment. The ***cognitive diagnostic classification accuracy index*** calculates the probability of accurately classifying a randomly selected student into the true category based on his/her responses to test items, which can serve as an index of classification accuracy for a cognitive diagnostic assessment. For assessments with known or previously calibrated item parameters, it has been shown that the sampling distributions of the new indexes are asymptotic normal when sample size approaches infinity. This is important because the sampling distribution of these two new indexes offers a way to estimate their standard errors and confidence intervals, which have important implications for evaluating the accuracy and precision of the sample estimates and for comparing different tests, or classification methods (i.e., CDMs).The two indexes can be used as important indicators of the reliability and validity of classification results produced by cognitive diagnostic assessments, which provide researchers and test developers with a path for ensuring the accuracy of the diagnostic feedback received by teachers and students.

We have conducted a preliminary simulation study to evaluate the performances for our new classification indexes. Four factors were manipulated, including the quality of test items, total number of skills measured by the test, dependency among the skills, and sample size. We hypothesized that higher quality test items, smaller number of skills measured within one test, and more dependencies among the skills would lead to more reliable and valid diagnostic results and consequently higher values of the two new indexes. The simulation results were interpretable and confirmed our hypothesis, indicating that the new classification consistency and accuracy indexes performed well on the simulated diagnostic data. We also found that sample size had little or no impact on the values of the two new indexes, which shows that they are robust to small sample sizes. Furthermore, the findings from the simulation study also suggest that the asymptotic normal theory of the new classification indexes may be safely applied even when sample size is small or moderate (e.g., 100 to 500).

Although the preliminary findings indicate our new classification indexes are statistically and logically sound, to date, they have not been applied to real data. Therefore, the first objective of the proposed research is to conduct real data studies using the new classification indexes. In this way, we can illustrate to researchers and practitioners (e.g., test developers) how to use and interpret the new classification indexes with observed student responses from real cognitive diagnostic assessments. In addition, by using the new indexes in applied settings, we can evaluate their strengths and weaknesses and identify new issues that must be addressed with these two new indexes. Results from the real data study will reveal the potential challenges with the use and interpretation of the new indexes in real diagnostic testing situations, which will lead to new research questions, directions, and projects that can further enhance the validation practices of cognitive diagnostic testing and ultimately improve the validity and reliability of the feedback yielded by cognitive diagnostic assessments.

The second objective of this research is to implement the new indexes into software to promote their use by researchers and test developers. This new computing tool will lead to fast and increased application of the new indexes for evaluating cognitive diagnostic assessments. By putting these indexes into the hands of researchers and practitioners, they can estimate the consistency and accuracy of the classification results with their own diagnostic data. It can also inspire new research that assesses the performances of the new classification consistency and accuracy indexes by conducting simulation or cross validation studies and new research that examines possible ways to improve these indexes.

Methodology

Year 1: Applying New Classification Consistency and Accuracy Indexes to Real Data

Although we conducted a simulation study to evaluate our new indexes and results suggested that they performed well with simulated data, the empirical investigation of these indexes is still needed because real data normally contain many sources of errors besides those examined in the simulation. In other words, real data are messy: observations are not always independent from one another, samples are not always representative, missing data and outliers are typically present, and so on. By conducting empirical studies, we can identify these real data problems and study their impact on the new indexes.

Our real data application of the new classification indexes will use national and international data collected from cognitive diagnostic assessments in two content domains, Mathematics and English. The national data were obtained from a pilot study designed to examine the technical and practical feasibility of obtaining more instructional relevant results from provincial achievement tests. This pilot study was carried out in Alberta. A total of 3 diagnostic mathematics tests aimed to provide teachers with diagnostic information regarding student mathematical skills were developed by seven content specialists from Alberta Education (Gierl, Alves, & Majeau, 2010). These tests measure a total number of 48 mathematical skills in the four content areas: (a) Number, (b) Patterns and Relations, (c) Shape and Space, and (d) Statistics and Probability at grade 3 and 6 levels. For each test, the number of items ranges from 16 to 24. The diagnostic tests were administered using an online computer-based testing system. Participation was voluntary. The sample size ranges from 236 to 753.

The international data that will be used in the proposed study were collected as a part of a larger study intended to create a web-based cognitive diagnostic computerized adaptive testing (CAT) system for Level 2 English Achievement Test in China (Liu et. al, 2010). A total of 352 multiple-choice items, including Listening Dialog, Short Talks, Grammar and Vocabulary, and Reading Comprehension, were developed by a group of English teachers and content specialists in China. These items measure a total of eight diagnostic skills such as understanding words, reorganizing words, and understanding grammar. A 40-item anchor test was designed and administered to 5,611 students from 78 schools in 12 cities. Data from the anchor test will be used in the proposed real data study.

The reliability and validity of each of the above diagnostic tests will be investigated using the new classification consistency and accuracy indexes. Guidelines for interpreting these indexes will be recommended. Suggestions on how to deal with real data problems (e.g., missing data, outliers) will be provided.

Year 2: Implementing New Classification Consistency and Accuracy Indexes into Software

The degree to which new statistical methods will be used in research and practical situations depends largely on their accessibility to researchers and practitioners. To make the new classification indexes more readily available, we will develop software with a user friendly interface and interactive analysis capability. The software will be developed in Visual Basic, which can run on any Windows operating system. Knowledge of programming techniques and computational and mathematical statistics will not be required in order to make use of the new software. The software will be tested using both simulated and real data so that defects, if any, can be detected and corrected. The internal design of the software will be well documented for the purpose of the ease of future maintenance and enhancement. The software, its user guide, and examples of input and output files will be available for free download. In this way, the two new indexes can be used by researchers and test developers to examine the consistency and accuracy of the classification results of different cognitive diagnostic assessments. In addition, the new software will allow researchers to conduct more comprehensive simulation studies to further study the statistical properties of the new indexes.

Summary and Implications

To summarize, the proposed research is aimed to address one of the most fundamental questions in the area of cognitive diagnostic assessments: *how consistent and accurate are the classification results produced by a cognitive diagnostic assessment?* The goal of cognitive diagnostic assessment is to improve the informational value of large-scale assessments by providing students with specific information about their strengths and weaknesses. In order for the diagnostic feedback to be truly helpful in the classroom, however, every effort must be made to ensure the accuracy of the classification results produced by the diagnostic test. It directly links to the effectiveness of teachers' and students' remediation efforts. The proposed research will provide researchers and practitioners such as test developers with a useful tool to evaluate the consistency and accuracy of diagnostic results and, therefore, promote the future success of cognitive diagnostic assessments. Through the better integration of testing, teaching and learning, we can provide higher quality education to our students to equip them with the knowledge and skills they need to succeed in the labour market, which ultimately leads to a higher quality of life for Canadians.

The proposed research will help initiate a sustained program of research on how to systematically evaluate the reliability and validity of cognitive diagnostic assessments. The results will form the theoretical foundation for a future SSHRC Insight Grant application in which we will propose to conduct external validation studies to further assess the usefulness of the new indexes in the validation of cognitive diagnostic results. A test-retest design will be used to examine the hypothesis that diagnostic feedback provided by a test with higher indexes (on consistency and accuracy) is more efficient and can lead to more significant improvement of student performance after a remediation session.

Outcomes of this research will be disseminated in both academic and applied settings to educational researchers, test developers and policymakers. Interim results will be presented at major conferences such as Canadian Society for the Study of Education (CSSE) and National Council on Measurement in Education (NCME). The final results will be valuable additions to the cognitive diagnostic assessment literature and will be submitted to major refereed journals. For example, one paper will focus on the results from the real data application of the new classification indexes and will be submitted to *Applied Measurement in Education.* Another paper will focus on the discussion of the impact of different real data problems on the new indexes and how to deal with these problems. This paper will be submitted to *Educational Measurement: Issues and Practice.* The abstract of our new software will be submitted to *Applied Psychological Measurement.* Anticipated avenues for dissemination to practitioners include the new software and its user guide as an online resource and a workshop on how to use the new software to validate the cognitive diagnostic assessment results. Intended audiences are policymakers, test developers and psychometricians from the Council of Ministers of Education, Canada, and different provincial departments of education across Canada.

References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Berk, R. A. (1984). Selecting the index of reliability. In R. A. Berk (Ed.), *A guide to criterion referenced test construction*. Baltimore: Johns Hopkins University Press.

Cui, Y. (2007). *The hierarchical consistency index: A person-fit statistic for the attribute hierarchical method.* Unpublished Doctoral Dissertation. University of Alberta, Department of Educational Psychology.

Cui, Y. & Gierl, M. J. (2010). Statistical Classification for Cognitive Diagnostic Assessment: An Artificial Neural Network Approach. Manuscript submitted for publication.

Cui, Y., Gierl, M. J., & Chang, H. (2011). Estimating classification consistency and accuracy for cognitive diagnostic assessment. Manuscript submitted for publication.

Cui, Y. & Leighton, J. P. (2009). The Hierarchy consistency index: Evaluating person fit for cognitive diagnostic assessment. *Journal of Educational Measurement, 46(4),* 429–449.

Cui, Y. & Roberts, M. R. (2010). A Person-Fit Study for Cognitive Diagnostic Assessment. Revise and resubmit to *Journal of Applied Measurement in Education.*

de la Torre, J., & Douglas, J. A. (2004). Higher-order latent trait models for cognitive diagnosis. *Psychometrika, 69*(3), 333-353.

Emons, W. H. M., Sijtsma, K., & Meijer, R. R. (2007). On the consistency of individual classification using short scales. *Psychological Methods, 12,* 105-120.

Gierl, M. J., Alves, C., & Majeau, R. T. (2010). Using the attribute hierarchy method to make diagnostic inferences about examinees' knowledge and skills in Mathematics: An operational implementation of cognitive diagnostic assessment. *International Journal of Testing, 10*, 318-341.

Gierl, M. J., Cui, Y, & Hunka, S. (2008). Using connectionist models to evaluate examinees' response patterns on tests. *Journal of Modern Applied Statistical Methods, 7,* 234-245.

Gierl, M. J., Cui, Y., & Zhou, J. (2009). Attribute reliability in cognitive diagnostic assessment. *Journal of Educational Measurement, 46(3),* 293-313.

Gierl, M. J., Zheng, Y., & Cui, Y. (2008). Using the attribute hierarchy method to identify and interpret differential group performance on tests. *Journal of Educational Measurement, 45,* 65-89.

Hanson, B. A., & Brennan, R. L. (1990). An investigation of classification consistency indexes estimated under alternative strong true score models. *Journal of Educational Measurement, 27*, 345-359.

Huynh, H. (1976). Statistical consideration of mastery scores. *Psychometrika, 41*, 65-78.

Huynh, H. (1990). Computation and statistical inference for decision consistency indexes based on the Rasch model. *Journal of Educational Statistics, 15,* 353-368.

Jang, E. E. (2008). A framework for cognitive diagnostic assessment. In C. A. Chapelle, Y. R. Chung, & J. Xu (Eds.), *Towards adaptive CALL, Natural language processing for diagnostic language assessment* (pp. 117-131). Ames, IA: Iowa State University.

Junker, B. W., & Sijtsma, K. (2001). Cognitive assessment models with few assumptions, and connections with nonparametric item response theory. *Applied Psychological Measurement, 25*, 258-272.

Lee, W. (2010). Classification consistency and accuracy for complex assessments using item response theory. *Journal of Educational Measurement, 47,* 1-17.

Lee, W., Brennan, R. L., & Wan, L. (2009). Classification consistency and accuracy for complex assessments under the compound multinomial model. *Applied Psychological Measurement, 33,* 374-390.

Leighton, J. P., Cui. Y., & Cor, M. C. (2009). Testing expert-based and student-based cognitive models: An application of the attribute hierarchy method and hierarchy consistency index. *Journal of Applied Measurement in Education, 22(3),* 229-254.

Leighton, J. P., & Gierl, M. J. (2007). Defining and evaluating models of cognition used in educational measurement to make inferences about examinees' thinking processes. *Educational Measurement: Issues and Practice, 27,* 3-16.

Leighton, J. P., Gierl, M. J., & Hunka, S. (2004). The attribute hierarchy model: An approach for integrating cognitive theory with assessment practice. *Journal of Educational Measurement, 41,* 205-236.

Liu, H., You, X., Wang, W., Ding, S., & Chang, H. (2010). The development of computerized adaptive testing with cognitive diagnosis for English Achievement Test in China. Manuscript submitted for publication.

Livingston, S. A., & Lewis, C. (1995). Estimating the consistency and accuracy of classifications based on test scores. *Journal of Educational Measurement, 32,* 179-197.

McGlohen, M., & Chang, H. (2008). Combining computer adaptive testing technology with cognitively diagnostic assessment. *Behavior Research Methods, 40 (3),* 808-21.

National Research Council (NRC). (2001). *Knowing what students know: The Science and design of educational assessment.* Washington, DC: National Academy Press.

Roussos, L., diBello, L. V., Stout, W., Hartz, S., Henson, R. A., & Templin, J. H. (2007). The fusion model skills diagnosis system. In J. P. Leighton, & Gierl, M. J. (Eds.), *Cognitive diagnostic assessment for education: Theory and applications* (pp. 275-318), Cambridge, UK: Cambridge University Press.

Rupp, A. A., & Templin, J. (2008). Unique characteristics of diagnostic classification models: A comprehensive review of the current state-of-the-art. *Measurement: Interdisciplinary Research and Perspectives, 6(4),* 219-262.

Rupp, A. A., Templin, J., & Henson, R. (2010). *Diagnostic measurement: Theory, methods, and applications.* New York: Guilford Press.

Schulz, E. M., Kolen,M. J., & Nicewander,W. A. (1999). A rationale for defining achievement levels using IRT-estimated domain scores. *Applied Psychological Measurement, 23,* 347–362.

Statistics Canada and Organisation for Economic Co-operation and Development (2005). *Learning a Living: First Results of the Adult Literacy and Life Skills Survey*. Statistics Canada catalogue no. 89-603-XIE. Accessed Nov. 19, 2010.

Subkoviak, M. J. (1976). Estimating reliability from a single administration of a criterion-referenced test. *Journal of Educational Measurement, 13,* 265-276.

Templin, J. L., & Henson, R. A. (2006). Measurement of psychological disorders using cognitive diagnosis models. *Psychological Methods, 11,* 287-305.

von Davier, M. (2007). *Hierarchical general diagnostic models* (Research Report No. RR-07-19). Princeton, NJ: Educational Testing Service.

Wang, T., Kolen, M. J., & Harris, D. J. (2000). Psychometric properties of scale scores and performance levels for performance assessments using polytomous IRT. *Journal of Educational Measurement, 37,* 141–162.

Wilcox, R. R. (1977). Estimating the likelihood of false-positive and false negative decisions in mastery testing: An empirical Bayes approach. *Journal of Educational Statistics, 2,* 289-307.

A. Description of Research Team

The research team includes the principal investigator Dr. Ying Cui and the co-investigator Dr. Mark Gierl. The formation of this team is based on team members' common interests in the area of cognitive diagnostic assessments. This research builds on the previous collaborations between Drs. Cui and Gierl on investigating different psychometric issues associated with cognitive diagnostic assessments. For example, they collaborated on using the artificial neural network approach to classify students (Cui & Gierl, 2010; Gierl, Cui, & Hunka, 2008), developing new procedures for examining the differential item functioning (Gierl, Zheng, & Cui, 2008) and skill-level reliability (Gierl, Cui, & Zhou, 2009), and, most recently, developing the new classification consistency and accuracy indexes for validating the diagnostic results (Cui, Gierl, & Chang, 2011). Drs. Cui and Gierl have developed a long-term team research agenda to examine different theoretical and technical aspects of cognitive diagnostic assessment.

Principal Investigator: Dr. Ying Cui. Dr. Cui will take the central role in organizing and completing the proposed research. Dr. Cui will work closely with the research assistants and take the leadership in overseeing the execution of the research project, especially in the mathematical and technical aspects of the research. She will ensure timelines are met so each component of the study is completed in a timely fashion. She will be responsible for setting up and coordinating the team meetings required during the project to discuss current research activities, recent research articles, and plans for future studies. Dr. Cui will share responsibility with the co-investigator and graduate research assistants for communicating research results through scholarly publications and conference presentations. If funded, this project will be the primary focus of Dr. Cui's research time (approximately 80%).

Dr. Cui completed her PhD degree in Measurement, Evaluation and Cognition from the Department of Educational Psychology at University of Alberta in 2007. Currently working as an Assistant Professor, Dr. Cui has a strong knowledge and understanding of the quantitative research methodology and psychometric issues relevant to measurement and evaluation. Her current research is focused on investigating different quantitative methods and issues related to cognitive diagnostic assessments. Since joining University of Alberta, she has been a principal or co-investigator in 5 research projects. Funding agencies include University of Alberta, CIHR, and the College Board.

Co-investigator: Dr. Mark Gierl. Dr. Gierl is committed to contributing to the intellectual direction of this research endeavor. His measurement expertise and his extensive involvement in the Alberta Education were identified as valuable for maintaining relevancy of the research for policy and practice. Dr. Gierl will provide advice and assistance regarding the practical aspect of the research, especially on how to generate and translate research findings that are interpretable and meaningful for practitioners and policy makers. Dr. Gierl is Professor of Educational Psychology. His specialization is educational and psychological testing, with an emphasis on the application of cognitive principles to assessment practices. Since joining University of Alberta, he has been a principal or co-investigator in 8 major research projects focusing on different aspects of educational measurement. Funding agencies include SSHRC, the College Board, and the Canadian Language and Literacy Research Network. His current research is focused on cognitive diagnostic assessment, differential item and bundle functioning, and assessment engineering. Dr. Gierl holds the Canada Research Chair in Educational Measurement.

B. Description of proposed student training strategies

The proposed project requires highly qualified research assistants who have strong programming skills and solid background in quantitative methodology. Two Ph.D students in the Centre for Research in Applied Measurement in Evaluation (CRAME) at University of Alberta will be asked to work on the project. One student is in the second year of his doctoral study and has a master degree in quantitative psychology. The other student is in the third year of his Ph.D study and has very strong programming skills. These two students will play very important roles over the 2 years of the project. The research team will guide the students through all phases of the research program, including research design, data analysis and interpretation, software development and testing, and dissemination of results at conferences and for publication. The students will gain methodological knowledge and organizational skills that allow them to successfully plan and execute their own research projects. In addition, students will be encouraged to base on the findings from this project to pursue their own research interest and potential dissertation topics. In order to achieve these goals, students will be involved in the following research activities:

Participation in Research Meetings. Students will attend regular weekly or biweekly team meetings to discuss current research activities, identify research problems, engage in problem-solving, share recent research articles, and discuss plans for the next step of the research.

Data Analysis and Interpretation. Students will work closely with the principal and co-investigators to use the new classification indexes to evaluate the consistency and accuracy of real diagnostic data. Students will receive extensive training and experience in cognitive diagnostic assessment design, data cleaning, analysis and interpretation.

Software Development and Testing. Supervised and assisted by the principal investigator, students will take an active role in the development and testing of the new computer software. Students will learn how to develop user-friendly software and use real and simulated data to test the software.

Communication of Results. One of our goals is to help students develop effective writing skills. We believe being able to effectively write is an invaluable skill for anyone. All undergraduate and graduate students should leave the university with the ability to effectively write. To accomplish this, the two students who participate in the project will take an active role in the communication of results and the writing of manuscripts. We will help them with the writing of proposals and papers for academic and professional conferences. And the students will have the opportunity to attend these meetings and present research findings. They will be encouraged to take the lead in writing the proposals and journal articles when they feel comfortable. The experience and skills students acquire through the project will directly contribute to the progress of their study and research.

Family name, Given name
Cui, Ying

Funds Requested from SSHRC

For each budget year, estimate as accurately as possible the research costs that you are asking SSHRC to fund through a grant. For each Personnel costs category, enter the number of individuals to be hired and specify the total amount required. For each of the other categories, enter the total amount required.

	Year 1		Year 2	
Personnel costs	No.	Amount	No.	Amount
Student salaries and benefits/Stipends				
Undergraduate				
Masters				
Doctorate	1	15,723	2	32,703
Non-student salaries and benefits/Stipends				
Postdoctoral				
Other				

Travel and subsistence costs	Year 1	Year 2
Applicant/Team member(s)		
Canadian travel	0	2,200
Foreign travel	0	3,200
Students		
Canadian travel	0	2,200
Foreign travel	0	3,200
Other expenses		
Professional/Technical services		
Supplies	1,000	1,000
Non-disposable equipment		
Computer hardware		
Other		
Other expenses (specify)		
Total	16,723	44,503

Budget Justification

1) Student salaries and benefits. Total: **$48,426**

- I propose to have one doctoral student working on the project for 3 terms in Year 1 and two doctoral students in Year 2.
  - In Year 1, a research assistant with strong background in quantitative methodology will be recruited to assist with the data cleaning, analysis and interpretation in real data studies. The student will be paid based on an 8-hour work per week. At the University of Alberta, the rate of pay per term for 2011/2012 for a doctoral student is $5,241 [2010/2011 Schedule of Payment for Graduates, University of Alberta adjusted for an expected raise of 4%]. 4% corresponds approximately to the average wage increase of graduate student assistantships over that last three years. The total for the student at the University of Alberta in Year 1 will be: **15,723**.
  - In Year 2, another research assistant with strong programming skills will be hired to help with the software development and the research assistant from year 1 will work independently to test whether the results from the new software are reliable and accurate using real data and simulated data. Each student will be paid based on an 8-hour work per week. Using the 4% increase, the total for the students at the University of Alberta in Year 2 will be **$32,703**.

2) Travel and Subsistence Costs (Year 2). Total: **$10,800**

- Conference Travel—Applicant/Team Members: Based on past experience, we estimate the average cost of each trip (including conference registration, travel, hotel, and meals) from Edmonton to the USA will be approximately $1,600 and from Edmonton to other Canadian cities, $1,100.
  - *Canadian Society for the Study of Education* (2013, Victoria, BC) 2 people @ $2,200.
  - *National Council on Measurement in Education* (2013, Atlanta, GA) 2 people @ $3,200.

  Total faculty conference travel: **$5,400.**
- Conference Travel—Students:
  - *Canadian Society for the Study of Education* (2013, Victoria, BC) 2 students @ $2,200.
  - *National Council on Measurement in Education* (2013, Atlanta, GA) 2 students @ $3,200.

  Total student conference travel: **$5,400.**

3) Supplies. Total: **$ 2,000**

- The budget includes memory sticks, DVDs and CDs for mass storage of computer programs and data, printer toner, photocopying, reference materials, long distance telephone charges.

Grant total: $61,226

Family name, Given name
Cui, Ying

Funds from Other Sources

You must include all other sources of funding for the proposed research. Indicate whether these funds have been confirmed or not. Where applicable, include (a) the partners' material contributions (e.g. cash and in-kind), and (b) funds you have requested from other sources for proposed research related to this application.

Full organization name Contribution type	Confirmed	Year 1	Year 2	
	☐			
	☐			
	☐			
	☐			
	☐			
	☐			
	☐			
	☐			
	☐			
Total funds from other sources		0	0	

Personal infomation will be stored in the Personal Information Bank for the appropriate program.

Family name, Given name
Cui, Ying

Intended Outcomes of Proposed Activities

Elaborate on the potential benefits and/or outcomes of your proposed research and/or related activities.

Scholarly Benefits

Indicate and rank up to 3 scholarly benefits relevant to your proposal.

Rank	Benefit	If "Other", specify
1	Enhanced research methods	
2	Student training/skill development	
3	Knowledge creation/intellectual outcomes	

Social Benefits

Indicate and rank up to 3 social benefits relevant to your proposal.

Rank	Benefit	If "Other", specify
1	Training and skill development	
2	Social outcomes	
3	Economic outcomes, including enhanced commercialization	

Audiences

Indicate and rank up to 5 potential target audiences relevant to your proposal.

Rank	Audience	If "Other", specify
1	Students	
2	Academic sector/peers, including scholarly associations	
3	Provincial/territorial government	
4	Practitioners/professional associations	
5	International audiences	

Family name, Given name
Cui, Ying

Expected Outcomes Summary

Describe the potential benefits/outcomes (e.g., evolution, effects, potential learning, implications) that could emerge from the proposed research and/or other partnership activities.

The proposed research is aimed to address one of the most fundamental questions in the area of cognitive diagnostic assessments: how consistent and accurate are the classification results produced by a cognitive diagnostic assessment? This research addresses this question by empirical experimentation with two new statistical indexes for validating the results produced by cognitive diagnostic assessments. It will provide researchers and practitioners such as test developers with a useful tool to evaluate the consistency and accuracy of diagnostic results and, therefore, promote the future success of cognitive diagnostic assessment so as to improve the diagnostic value of large-scale assessments in producing more instructionally relevant results that support classroom teaching and learning. Through the better integration of testing, teaching and learning, we can provide higher quality education to Canadian students to equip them with the knowledge and skills they need to succeed in the labour market, which ultimately contributes to the future success of Canada in the global economy and a higher quality of life for Canadians.

We will implement these new indexes into software that will be disseminated broadly to researchers and practitioners to promote more effective use of cognitive diagnostic assessment. This new computing tool will lead to fast and increased application of the new indexes for evaluating cognitive diagnostic assessment results. By putting these indexes into the hands of researchers and practitioners, they can estimate the consistency and accuracy of the classification results with their own diagnostic data. It can also inspire new research that assesses the performances of the new classification consistency and accuracy indexes by conducting simulation or cross validation studies and new research that examines possible ways to improve these indexes.

Outcomes of this research will be disseminated in both academic and applied settings to educational researchers, test developers and policymakers. Interim results will be presented at major conferences such as Canadian Society for the Study of Education (CSSE) and National Council on Measurement in Education (NCME). The final results will be valuable additions to the cognitive diagnostic assessment literature and will be submitted to major refereed journals. For example, one paper will focus on the results from the real data application of the new classification indexes and will be submitted to Applied Measurement in Education. Another paper will focus on the discussion of the impact of different real data problems on the new indexes and how to deal with these problems. This paper will be submitted to Educational Measurement: Issues and Practice. The abstract of our new software will be submitted to Applied Psychological Measurement.

Anticipated avenues for dissemination to practitioners include the new software and its user guide as an online resource and a workshop on how to use the new software to validate the cognitive diagnostic assessment results. Intended audiences are policymakers, test developers and psychometricians from the Council of Ministers of Education, Canada, and different provincial departments of education across Canada.

The proposed research will help initiate a sustained program of research on how to systematically evaluate the reliability and validity of cognitive diagnostic assessments. The results will form the theoretical foundation for a future SSHRC Insight Grant application in which we will propose to conduct external validation studies to further assess the usefulness of the new indexes.

Do not photocopy this page.

Internal use	CID (if known)
696396	181155

Identification

Only the information in the Name section will be made available to selection committee members and external assessors. Citizenship and Statistical and Administrative Information will be used by SSHRC for administrative and statistical purposes only. Filling out the statistical and Administrative Information section is optional.

Name

Family name	Given name	Initials	Title
Cui	Ying		Dr.

Citizenship - Applicants and co-applicants must indicate their citizenship status by checking and answering the applicable questions.

Citizenship status					Have you applied for permanent residency?
◯ Canadian	◉ Permanent resident since (yyyy/mm/dd) 2007/04/06	◯ Other (country)			◯ Yes ◯ No

Statistical and Administrative Information

Birth year	Gender	Permanent postal code in Canada (i.e. K2P1G4)	Correspondence language	Previous contact with SSHRC? (i.e. applicant, assessor, etc.)
1978	◉ F ◯ M	T6G2G5	◉ English ◯ French	◉ Yes ◯ No

Full name used during previous contact, if different from above

Contact Information

The following information will help us to contact you more rapidly. Secondary information will not be released by SSHRC without your express consent.

	Country code	Area code	Number	Extension
Primary telephone number	1	780	4923848	
Secondary telephone number	1	780	4287623	
Primary fax number	1	780	4920001	
Secondary fax number				

Primary E-mail: yc@ualberta.ca

Secondary E-mail:

Personal information will be stored in the Personal Information Bank for the appropriate program.

Checked
Web CV
2011/01/28

Do not photocopy this page.

Family name, Given name

Cui, Ying

Current Address

Use only if you are not affiliated with a department at a Canadian university. (If you are affiliated with a department at a Canadian university, the department's mailing address will be used.) If you wish to use another address, specify it under the Correspondence Address.

Address

CRAME, 6-110 Education North

Faculty of Education

University of Alberta

City/Municipality	Prov. / State	Postal/Zip code
Edmonton	AB	T6G2G5

Country CANADA

Correspondence Address

Complete this section if you wish your correspondence to be sent to an address other than your current address.

Address

City/Municipality	Prov. / State	Postal/Zip code

Country

Temporary Address

If providing a temporary address, phone number and/or E-mail, ensure that you enter the effective dates.

Address

City/Municipality	Prov./ State	

Country

Start date (yyyy/mm/dd)	End date (yyyy/mm/dd)

Temporary E-mail

Permanent Address in CANADA

Address

City/Municipality	Prov./ State	Postal/Zip code

Country

Temporary telephone/fax number

Country code	Area code	Number	Extension

Do not photocopy this page.

Family name, Given name
Cui, Ying

Research Expertise (optional)

The information provided in this section refers to your own research expertise, not to a research proposal. Filling out the following 4 sections is optional. This page will not be seen by selection committee members and external assessors. This section will be used for planning and evaluating programs, producing statistics, and selecting external assessors and committee members.

Areas of Research

Indicate and rank up to 3 areas of research that best correspond to your research interests as well as areas where your research interests would apply. Duplicate entries are not permitted.

Rank	Code	Area
1	140	Education
2	111	Children
3	382	Youth

Temporal Periods

If applicable, indicate up to 2 historical periods covered by your research interests.

From	To
Year BC AD	Year BC AD
_____ ◯ ◯	_____ ◯ ◯
_____ ◯ ◯	_____ ◯ ◯

Geographical Regions

If applicable, indicate and rank up to 3 geographical regions covered by your research interests. Duplicate entries are not permitted.

Rank	Code	Region
1		
2		
3		

Countries

If applicable, indicate and rank up to 5 countries covered by your research interests. Duplicate entries are not permitted.

Rank	Code	Countries	Prov./ State
1			
2			
3			
4			
5			

Curriculum Vitae

Family name, Given name: Cui, Ying

Language Proficiency

	Read	Write	Speak	Comprehend aurally
English	X	X	X	X
French				

Other languages: Chinese

Work Experience

List the positions, academic and non-academic, you have held beginning with the current position and all previous positions in reverse chronological order, based on the start year.

Current position: Assistant Professor
Start date (yyyy/mm): 2007/7

Org. code: 1480111
Full organization name: University of Alberta

Department/Division name: Educational Psychology

Position type: ○ Tenured ○ Non-tenure ◉ Tenure-track ○ Non-academic
Employment status: ◉ Full-time ○ Part-time ○ Non-salaried ○ Leave of absence

Position: Statistical Consultant
Start date (yyyy/mm): 2005
End date (yyyy/mm): 2007

Org. code: 1480111
Full organization name: University of Alberta

Department/Division name: Educational Psychology

Position: Research Assistant
Start date (yyyy/mm): 2003
End date (yyyy/mm): 2007

Org. code: 1480111
Full organization name: University of Alberta

Department/division name: Educational Psychology

Position: Research Assistant
Start date (yyyy/mm): 2000
End date (yyyy/mm): 2003

Org. code: 9574104
Full organization name: Renmin University of China

Department/Division name: Department of Statistics

Family name, Given name
Cui, Ying

Academic Background

List up to 5 degrees, beginning with the highest degree first and all others in reverse chronological order, based on the start date.

Degree type	Degree name	Start date (yyyy/mm)	Expected date (yyyy/mm)	Awarded date (yyyy/mm)
Doctorate	Doctor of Philosophy	2003/09		2007/11

Disc. code	Discipline	Did SSHRC support enable you to get this degree?
61232	Measurement and Evaluation	◯ Yes ◉ No

Org. code	Organization
1480111	University of Alberta

Country: CANADA

Degree type	Degree name	Start date (yyyy/mm)	Expected date (yyyy/mm)	Awarded date (yyyy/mm)
Master's	Master of Science	2000/09		2003/06

Disc. code	Discipline	Did SSHRC support enable you to get this degree?
99999	Statistics	◯ Yes ◉ No

Org. code	Organization
9574104	Renmin University of China

Country: CHINA

Degree type	Degree name	Start date (yyyy/mm)	Expected date (yyyy/mm)	Awarded date (yyyy/mm)
BA Gen.	Bachelor of Science	1996/09		2000/06

Disc. code	Discipline	Did SSHRC support enable you to get this degree?
99999	Statistics	◯ Yes ◉ No

Org. code	Organization
9574094	Qingdao Technological University

Country: CHINA

Degree type	Degree name	Start date (yyyy/mm)	Expected date (yyyy/mm)	Awarded date (yyyy/mm)

Disc. code	Discipline	Did SSHRC support enable you to get this degree?
		◯ Yes ◯ No

Org. code	Organization

Country:

Degree type	Degree name	Start date (yyyy/mm)	Expected date (yyyy/mm)	Awarded date (yyyy/mm)

Disc. code	Discipline	Did SSHRC support enable you to get this degree?
		◯ Yes ◯ No

Org. code	Organization

Country:

Personal information will be stored in the Personal Information Bank for the appropriate program.

Family name, Given name

Cui, Ying

Credentials

List up to 6 licences, professional designations, awards and distinctions you have received and feel would be the most pertinent to the adjudication of your application. List them in reverse chronological order, based on the year awarded.

Category	Name	Source or Country	Duration (Months)	Value / Year awarded
Graduate Scholarship	Provost Doctoral Entrance Award	CANADA	12	$8,000 2004

Research Expertise

The information provided in this section refers to your own research expertise, not to a research proposal.

Keywords

List keywords that best describe your areas of research expertise. Separate keywords with a semicolon.

Measurement and Evaluation; Statistics; Cognition

Disciplines

Indicate and rank up to 5 disciplines that best correspond to your research interests. Duplicate entries are not permitted.

Rank	Code	Discipline	If Other, specify
1	61222	Educational Psychology	
2	63020	Psychometrics	
3	99999	Other	Statistics
4			
5			

Family name, Given name
Cui, Ying

Funded Research

List up to 8 grants or contracts you have received from SSHRC or other sources. List them in reverse chronological order, based on the year awarded. If you are not the applicant (principal investigator), specify that persons' name.

Org. code	Full name of funding organization	Year awarded (yyyy)	Total amount (CAN$)
3010747	Canadian Institutes of Health Research	2010	$94,303

Role: Co-applicant — Completion status: ☐ Complete

Project title: Neurobehavioural outcomes of children with prenatal alcohol exposure and fetal alcohol spectrum disorder

Applicant's family name	Applicant's given name	Initials
Rasmussen	Carmen	

Org. code	Full name of funding organization	Year awarded (yyyy)	Total amount (CAN$)
1	University of Alberta	2009	$10,000

Role: Applicant — Completion status: ☐ Complete

Project title: Developing an Item Selection Algorithm for Computerized Adaptive-Attribute Testing

Applicant's family name	Applicant's given name	Initials

Org. code	Full name of funding organization	Year awarded (yyyy)	Total amount (CAN$)
1	University of Alberta	2009	$49,800

Role: Co-applicant — Completion status: ☐ Complete

Project title: Microbiology and Molecular Biology for Environmental Engineers

Applicant's family name	Applicant's given name	Initials
Liu	Yang	

Org. code	Full name of funding organization	Year awarded (yyyy)	Total amount (CAN$)
1	University of Alberta	2007	$6,994

Role: Applicant — Completion status: ☒ Complete

Project title: Person-fit Analysis for Cognitive Diagnostic Assessment

Applicant's family name	Applicant's given name	Initials

Family name, Given name: Cui, Ying

Funded Research (cont'd)

Org. code	Full name of funding organization	Year awarded (yyyy)	Total amount (CAN$)
1	College Entrance Examination Board	2007	$156,000

Role: Co-applicant — Completion status: [X] Complete

Project title: Using Assessment Engineering Principles to Promote Diagnostic Inferences about Student Performance on the PSAT/NMSQT

Applicant's family name	Applicant's given name	Initials
Gierl	Mark	J

Org. code	Full name of funding organization	Year awarded (yyyy)	Total amount (CAN$)
1	College Entrance Examination Board	2005	$150,000

Role: Research Assistant — Completion status: [X] Complete

Project title: Identifying Cognitive Dimensions that Affect Student Performance on the New SAT

Applicant's family name	Applicant's given name	Initials
Gierl	Mark	J.

Org. code	Full name of funding organization	Year awarded (yyyy)	Total amount (CAN$)
3010325	Social Sciences and Humanities Research Council of Canada	2003	$59,547

Role: Research Assistant — Completion status: [X] Complete

Project title: Using Multivariate Statistics to Identify the Relationship between Profiles of Students' Academic Skills and the Underlying Trait Dimensions Measured in Large-scale Assessments

Applicant's family name	Applicant's given name	Initials
Leighton	Jacqueline	P.

Org. code	Full name of funding organization	Year awarded (yyyy)	Total amount (CAN$)

Role: — Completion status: [ ] Complete

Project title:

Applicant's family name	Applicant's given name	Initials

1. Research Contributions over the Last Six Years (2005-2011)

Refereed contributions

(R) Cui, Y. & Leighton, J. P. (2009). The hierarchy consistency index: A person-fit statistic for cognitive diagnostic assessment. *Journal of Educational Measurement, 46,* 429-449. **[Primary Author]**

(R) Gierl, M. J. & Cui, Y. (2008). Defining characteristics of diagnostic classification models and the problem of retrofitting in cognitive diagnostic assessment. *Measurement: Interdisciplinary Research and Perspectives, 6, 263-268.* **[Secondary Author]**

(R) Gierl, M. J., Cui, Y., & Hunka, S. (2008). Using connectionist models to evaluate examinees' response patterns on tests. *Journal of Modern Applied Statistical Methods, 7,* 234-245. **[Equal Author]**

(R) Gierl, M. J., Cui, Y., & Zhou, J. (2009). Attribute reliability in cognitive diagnostic assessment. *Journal of Educational Measurement, 46(3),* 293-313. **[Secondary Author]**

(R) Gierl, M. J., Zheng, Y., & Cui, Y. (2008). Using the attribute hierarchy method to identify and interpret differential group performance on tests. *Journal of Educational Measurement, 45,* 65-89. **[Secondary Author]**

(R) Klein, S., Guiltner, V., Sollereder, P, & Cui, Y. (2011). Relationships between fine-motor, visual-motor, and visual perception scores and handwriting legibility and speed. *Physical and Occupational Therapy in Pediatrics,* Early online: 1-12, 2011**.** **[Secondary Author]**

*(R) Leighton, J. P., Cui. Y., & Cor, M. C. (2009). Testing expert-based and student-based cognitive models: An application of the attribute hierarchy method and hierarchy consistency index. *Journal of Applied Measurement in Education, 22(3),* 229-254. **[Equal Author]**

*(R) Leighton, J. P., Gokiert, R. J., & Cui, Y. (2007). Using exploratory and confirmatory methods to identify the cognitive dimensions in a large-scale science assessment. *International Journal of Testing, 7(2),* 141-189. **[Secondary Author]**

(R) Li, J. C**.,** Chan, W. & Cui. Y. (2010). Bootstrap standard error and confidence intervals for indirectly range-restricted correlation: A Monte Carlo simulation study. *Journal of British Mathematical and Statistical Psychology.* Advanced online publication. Doi:10.1348/2044-8317.002007. **[Equal Author]**

(R) Olson, K., Hayduk, L., Cree, M., Cui, Y., Quan, H., Hanson, J., Lawlor, P., & Strasser, F. (2008). The changing causal foundations of cancer-related symptom clustering during the final month of palliative care: A longitudinal study. *BMC Medical Research Methodology, 8* (36), 11. **[Secondary Author]**

(R) Wang, S., Cui, Y., & Parrila, R. (2011). Examining the effectiveness of peer-mediated and video-modeling social skill interventions for children with autism spectrum disorders: A meta-analysis in single-case research using HLM. *Research in Autism Spectrum Disorders, 5,* 562-569. **[Equal Author]**

(R) Zheng, Y., Gierl, M. J., & Cui, Y. (2010). Using Cochran's Z statistic to test the kernel-Smoothed IRF differences between focal and reference groups. *Educational and Psychological Measurement, 70,* 541-556. **[Equal Author]**

Other Refereed Contributions

(R) Cui, Y. & Gierl, M. J. (2011). *Pattern recognition techniques for cognitive diagnostic assessment scoring and reporting: An artificial neural network approach.* Paper to be presented at the 2011 annual meeting of the National Council on Measurement in Education, New Orleans, LA.

(R) Cui, Y. & Gierl, M. J. (2009). *Validating a cognitive theory of algebra learning: The union of cognitive psychology and educational measurement.* Paper presented at the 2009 annual meeting of the Canadian Society for Studies in Education, Ottawa, ON.

(R) Cui, Y. Gierl, M. J., & Leighton, J. P. (2009). *Estimation the attribute hierarchy method with Mathematica*. Paper presented at the 2009 annual meeting of the American Educational Research Association (AERA), San Diego, CA.

(R) Cui, Y., & Leighton, J. P (2008). *The Hierarchy consistency index: evaluating person fit for cognitive diagnostic assessment.* Paper presented at the 2008 annual meeting of the National Council on Measurement in Education (NCME), New York, NY.

(R) Cui, Y., Leighton, J. P., Gierl, M. J., & Hunka, S. (2006). *The hierarchical consistency index: A person-fit statistic for the attribute hierarchical method.* Paper presented at the 2006 annual meeting of the National Council on Measurement in Education (NCME), San Francisco, CA.

(R) Cui, Y., Leighton, J. P., & Zheng, Y. (2006). *Simulation studies for evaluating the performance of the two classification methods in the AHM.* Poster presented at the 2006 annual meeting of the National Council on Measurement in Education (NCME), San Francisco, CA.

(R) Cui, Y., & Leighton, J. P. (2005). *Using linear discriminant analysis and logistic regression analysis to identify student's skills Measured by the SAIP science 99 written assessment.* Paper presented at the annual meeting of the Canadian Society for Studies in Education, Hamilton, On.

(R) Cui, Y., & Leighton, J. P. (2005). *A simulation approach for determining the number of factors in the SAIP science 99 written assessment.* Poster presented at the annual meeting of the Canadian Society for Studies in Education, Hamilton, On.

(R) Cui, Y. & Li, J. C. (2010). *Building diagnostic computerized adaptive testing: Issues and challenges.* Paper presented at the annual meeting of the Canadian Society for Studies in Education, Montreal, QC.

(R) Cui, Y. & Li, J.C. (2011). *Evaluating the performance of different procedures of constructing confidence interval for coefficient alpha: A simulation study.* Paper to be presented at the 2011 annual meeting of the American Educational Research Association, New Orleans, LA.

(R) Cui, Y., Roberts, M. R., & Gotzmann, A. (2010). *Evaluating statistical reasoning of college students in the social and health sciences with diagnostic assessment.* Paper presented at the 2010 annual meeting of the American Educational Research Association (AERA), Denver, CL.

(R) Gierl, M.J., Cui, Y., & Hunka, S. (2007). *Using connectionist models to evaluate examinees' response patterns on tests using the Attribute Hierarchy Method.* Paper presented at the 2007 annual meeting of the National Council on Measurement in Education (NCME), Chicago, IL.

*(R) Gokiert, R. J., Leighton, J. P., & Cui, Y. (2005). *Investigating the statistical and cognitive dimensions of large-scale science assessments.* Paper presented at the American Educational Research Association Annual Meeting, Montreal, QC.

*(R) Leighton, J. P., Cui, Y., Cor, K. M. (2008). *Testing expert-based and student-based cognitive models: An application of the attribute hierarchy method and hierarchical consistency index.* Paper presented at the 2008 annual meeting of the National Council on Measurement in Education (NCME), New York, NY.

*(R) Leighton, J. P., Heffernan, C., Cor, K. M., Gokiert, R. J., & Cui, Y. (2008). *An experimental test of student verbal reports and expert teacher evaluations for revising achievement items.* Paper presented at the 2008 annual meeting of the American Educational Research Association (AERA), New York, NY.

(R) Wang, S., & Cui, Y. (2009). *Examining the effectiveness of different social skill interventions for*

children with Autistic Spectrum Disorder: A meta-analysis. Paper presented at the 2009 annual meeting of the Canadian Society for Studies in Education, Ottawa, ON.

(R) Zhou, J., Gierl, M. J., & Cui, Y. (2007). *Computerized attribute-adaptive testing: A new computerized adaptive testing approach incorporating cognitive psychology.* Paper presented at Graduate Management Admission Council (GMAC) Conference on Computerized Adaptive Testing, Minneapolis, MN.

(R) Zhou, J., Gierl, M. J., & Cui, Y. (2009). *Attribute reliability in cognitive diagnostic assessment.* Paper presented at the 2009 annual meeting of the National Council on Measurement in Education (NCME), San Diego, CA.

Non-Refereed Contributions

Cui, Y. (2010). Discriminant analysis. *Encyclopedia of Research Design.* Thousand Oaks, CA: Sage Publications.

Cui, Y. (2009*). Constructing and validating a diagnostic test in the domain of statistical hypothesis testing.* Invited talk at the Statistical and Applied Mathematical Sciences Institute (SAMSI) Summer Program on Psychometrics, Research Park Triangle, NC.

Cui, Y. (2009). *Evaluating person fit for cognitive diagnostic assessment with the hierarchy consistency index.* Invited talk at the Statistical and Applied Mathematical Sciences Institute (SAMSI) Summer Program on Psychometrics, Research Park Triangle, NC.

Cui, Y. (2009). *Pattern recognition techniques for cognitive diagnostic assessment: An artificial neural network approach.* Invited talk at the Statistical and Applied Mathematical Sciences Institute (SAMSI) Summer Program on Psychometrics, Research Park Triangle, NC.

Cui, Y. (2009). *From theory to practice: What will it take for cognitive theory to have a real impact on large-scale assessment.* Invited talk at Southwest University, China.

Forthcoming Contributions

(R) Canel-Çınarbaş, D., Cui, Y., & Lauridsen, E. (in press). Cross-cultural validation of Beck Depression Inventory-II across U.S. and Turkish sample. *Measurement and Evaluation in Counseling and Development.* **[Equal Author]**

(R) Cui, Y. & Chang, H. (under review). Impact of the quality of item bank on the performance of computerized adaptive testing for cognitive diagnosis: A simulation study. *Applied Psychological Measurement.* **[Primary Author]**

(R) Cui, Y. & Gierl, M. J. (under review). Pattern recognition techniques for cognitive diagnostic assessment scoring and reporting: An artificial neural network approach. *Journal of Classification*. **[Primary Author]**

(R) Cui, Y., Gierl, M. J., & Chang, H. (under review). Estimating classification consistency and accuracy for cognitive diagnostic assessment. *Journal of Educational and Behavioral Statistics.* **[Primary Author]**

(R) Cui, Y. & Li, J.C. (under review). Evaluating the performance of different procedures of constructing confidence interval for coefficient alpha: A simulation study. *Journal of British Mathematical and Statistical Psychology.* **[Primary Author]**

(R) Cui, Y. & Roberts, M. R. (revised and resubmit). A person-fit study for cognitive diagnostic assessment. *Applied Measurement in Education.* **[Primary Author]**

(R) Hayduk, L., Olson, K., Quan, H., Cree, M., & Cui, Y. (in press). Temporal changes in the causal foundations of palliative-care symptoms. *Quality of Life.* **[Secondary Author]**

(R) Leighton, J. P., Heffern, C., Cor, M. K., Gokiert, R. J., & Cui, Y. (in press). An experimental test of student verbal reports and expert teacher evaluations for revising achievement items. *Applied*

Measurement in Education. **[Secondary Author]**

(R) Li, J. C**.,** Cui. Y., Gierl, M.J., & Chan, W. (under review). A more accurate confidence interval for Hunter and Schmidt's (2004) meta-analysis of indirectly range-restricted correlation. *Psychological Methods.* **[Equal Author]**

(R) Li, J. C**.,** Cui. Y., Gierl, M.J., & Chan, W. (under review). Bootstrap standard error and onfidence intervals for the reliability coefficient corrected for range restriction: a simulation study. *Multivariate Behavior Research.* **[Equal Author]**

(R) Olson, K., Rogers, W. T., Cui, Y., Cree, M., Baracos, V., Rust, T., Mellott, I., Johnson, L., Macmillan, K., Bonville, N. (in press). Development and psychometric testing of the adaptive capacity index, an instrument to measure adaptive capacity in individuals with advanced cancer. *International Journal of Nursing Studies.* **[Secondary Author]**

Theses

Cui, Y. (2007). *The hierarchical consistency index: A person-fit statistic for the attribute hierarchical method.* Unpublished Doctoral Dissertation. University of Alberta, Department of Educational Psychology.

Cui. Y. (2003*). Using the Bayesian Network approach to investigate consumer satisfaction index model.* Unpublished Master's Thesis. Renmin University, Department of Statistics.

2. Other Research Contributions

Reviewer Activities

Guest Reviewer for American Educational Research Association, Canadian Journal of Education, Educational Measurement: Issues and Practice, Healthcare Policy, Journal of Classification, Journal of Educational Measurement, Journal of Technology, Learning, and Assessment, National Council on Measurement in Education.

Grant Reviewer for Social Sciences and Humanities Research Council.

Consulting Activities

I am currently supervising the statistical consulting services provided by the Centre for Research in Applied Measurement and Evaluation (CRAME) at University of Alberta. Consulting services are provided to both students and faculty throughout the University. Our clients come from diverse departments including Secondary Education, Pediatrics, Psychology, Nursing, Physical Therapy, Physical Education, and Rehabilitation Medicine in addition to Educational Psychology.

4. Career Interruptions and Special Circumstances

Maternity Leave (May, 2010-Nov. 2010).

5. Contributions to Training

I started my academic career on July 1, 2007 in the Department of Educational Psychology at University of Alberta. Currently I am supervising 2 doctoral students. I have served on 10 supervisory committees (2 Master thesis and 8 doctoral dissertations) and 4 examination committees (1 Master thesis and 3 doctoral dissertations). Students are from a variety of departments, including Educational Psychology, Statistics, and Occupational Therapy. My goal in working with graduate students and research assistants is to help them develop research competence by guiding them through all phases of the research program, including research design and instrument development, data collection, statistical analysis and interpretation, and dissemination of results at conferences and for publication. Students and research assistants participate in weekly or bi-weekly research meetings to discuss current research activities, recent research articles, and plans for future studies. I have coauthored 6 conferences presentations and 7 research papers with my students and research assistants.