

SECONDARY SCHOOL
EXTERNAL EXAMINATION
SYSTEMS

Reliability, Robustness and Resilience

Edited by
Barend Vlaardingerbroek and Neil Taylor
Introduction by Stephen P. Heyneman



C A M B R I A
P R E S S

AMHERST, NEW YORK

Copyright 2009 Barend Vlaardingerbroek and Neil Taylor

All rights reserved

Printed in the United States of America

No part of this publication may be reproduced, stored in or introduced into a retrieval system, or transmitted, in any form, or by any means (electronic, mechanical, photocopying, recording, or otherwise), without the prior permission of the publisher.

Requests for permission should be directed to:

permissions@cambriapress.com, or mailed to:

Cambria Press

20 Northpointe Parkway, Suite 188

Amherst, NY 14228

Library of Congress Cataloging-in-Publication Data

Secondary school external examination systems: reliability, robustness and resilience / edited by Barend Vlaardingerbroek & Neil Taylor; with an introduction by Stephen P. Heyneman.

p. cm.

Includes bibliographical references and index.

ISBN 978-1-60497-600-7 (alk. paper)

1. High schools—Examinations—Cross-cultural studies. 2. Universities and colleges—Entrance examinations—Cross-cultural studies. I. Vlaardingerbroek, Barend. II. Taylor, N. S. (Neil S.) III. Title.

LB3060.28.S43 2009

373.126'2—dc22

2009000679

CHAPTER 1

INTRODUCTION: THE IMPORTANCE OF EXTERNAL EXAMINATIONS IN EDUCATION*

Stephen P. Heyneman

PREAMBLE

Every nation must select those who are certified as graduating successfully from schooling and those who may proceed to a university. In some places, such as Vietnam, this is done in two stages with different examinations. In other places, such as in New Zealand, the credit units established in secondary school may be sufficient to qualify for admission to a university. This book provides a comprehensive review of the testing processes in 19 national and state entities.

*This chapter is based in part on S. P. Heyneman and J. Fremer's 'The Importance of Fair Education Selection', Report for the World Bank, June 2008.

This introductory chapter reviews the reasons why these various systems matter. It points out that external examinations are necessary as a quality assurance mechanism and supply important feedback on curricular effectiveness and teaching. External examinations may also fulfil an important certification function by allowing those who complete them to advertise that completion. The term *secondary school graduate*, when supported by an external examination, has lifelong meaning.

This chapter also points out that selection is a normal process of nation building and that standardized assessment of one kind or another is inevitable. It makes the point that the most important distinction is not between systems that do or do not have selection tests, but between those circumstances in which test items are well designed and the testing system is secure, and other circumstances in which the tests are poorly designed and the testing system may be corrupt. It provides information on the characteristics in which 'education corruption' occurs and provides illustrations of the consequences.

WHY EXAMINATIONS ARE IMPORTANT

Since the time of Plato, it has generally been recognized that a key ingredient of a nation/state is how it chooses its leaders for technical, commercial, and political functions. It is understood that a nation would not prosper for long if leaders were chosen through ascriptive criteria, that is, on the basis of the characteristics with which they were born—race, gender or social status—or on the basis of inheritance (such as being the son of a king), military power (through force), or through political purchase by the wealthy. For modern nations to prosper, they must choose, so far as possible, future leaders on the basis of their personal potential for achievement.

On the other hand, fair competition to be a leader may contradict the natural tendency for families to protect and advantage their own children and relatives. Every parent would wish success for his/her own child; every group would hope for the success of children from that particular group. How can open and fair competition be balanced by the natural desire to advantage one's own children? This is one function of public schooling.

State-sponsored schooling providing opposing influences can be managed by nations to 'refresh' the source used by nations to measure the sacrifice attempted to measure the sacrifice there is a serious bias in the selection that developing countries could improve (GNP)/capita by five percentage points. Partnership upon merit as opposed to grades, the economic benefit to develop on the basis of merit would be three from a reduction in Organisation for Development (OECD) trade restriction on

Success in schooling is one of the necessary for modern leadership. Although through experience, just good fortune, success in schooling is considered to be a criterion of legitimacy.

But what if schooling itself is not to be believed, perhaps correctly, that the social group over others? What if the merit of teachers on student performance trust the criteria used as a means of the process of schooling had been corrupted?

In a democracy, the public often has a sense of the education system. If the selection system to be fair, it might be in commerce, science, military, or politics. If the system of leadership through privilege and merit system cannot be trusted, it may cause social cohesion, a principal ingredient of

WHY SELECTION IS NECESSARY

Some educators argue that selection is necessary for the classroom. They come

State-sponsored schooling provides the mechanism through which opposing influences can be managed. It is the common instrument used by nations to 'refresh' the sources of leadership. Economists have attempted to measure the sacrifice to economic growth that occurs if there is a serious bias in the selection of leaders. It has been estimated that developing countries could improve their Gross National Product (GNP)/capita by five percentage points if they were to base their leadership upon merit as opposed to gender or social status. By some estimates, the economic benefit to developing countries of choosing leaders on the basis of merit would be three times more than the benefit accruing from a reduction in Organisation for Economic Co-operation and Development (OECD) trade restriction on imports.

Success in schooling is one of the few characteristics believed necessary for modern leadership. Although it is possible for leaders to emerge through experience, just good fortune, or military might, regardless, success in schooling is considered to be a *sine qua non* as an essential criterion of legitimacy.

But what if schooling itself is not fair? What if the public comes to believe, perhaps correctly, that the provision of schooling favors one social group over others? What if the public does not trust in the judgment of teachers on student performance? What if the public does not trust the criteria used as a means of selection? What would happen if the process of schooling had been corrupted?

In a democracy, the public often takes an active interest in the fairness of the education system. If the public does not believe the education system to be fair, it might be said that current leaders—whether in commerce, science, military, or politics—had acquired their positions of leadership through privilege rather than achievement. If the school system cannot be trusted, it may detract from a nation's sense of social cohesion, a principal ingredient of all successful modern societies.

WHY SELECTION IS NECESSARY

Some educators argue that selection tests are biased and distort what is taught in the classroom. They complain that tests determine the nature of

pedagogy, that they emphasize rote memorization of uninterpreted and simplistic information. They characterize tests as being 'academic' and divorced from daily life. They accuse tests of being biased toward high achievers, leaving out all others. In sum, they argue that standardized selection tests do a disservice to public education. Their recommendation might be to have developing nations jettison selection tests altogether. This view fails to recognize the necessity of selection. With compulsory education comes greater participation in higher education. Even where attendance rates approach 80%, such as in Finland and Korea, selection must occur. Educational opportunity is not shaped like a rectangle but like a pyramid. If 'elite' refers to those able to enter selective training, then the education system in every nation has elite characteristics. The purposes of selection may vary, but selection for leadership is universal among those purposes. The question, therefore, is not whether a system selects a few to proceed, since all nations must select; rather, the question is how that selection is made.

Certain selection techniques are said to be 'better' than others. Essay questions and oral examinations are said to be superior to multiple-choice questions on the grounds that there is more opportunity for creative feedback, more 'data points' for observation, and more subtle means to demonstrate one's competence. When divorced from the context of their application, however, the discussion of the superiority of one testing technique over another is spurious. Both oral examinations and essay questions are more subject to subjective judgement because both are more difficult to standardize. Standardization—the ability of the test designer to create test-taking circumstances that are as nearly identical as possible—is an essential characteristic of any selection examination considered fair. Oral examinations leave open the possibility of bribery and corruption. Both oral examinations and essay questions are more labor intensive and hence both are considerably more costly—and cost does count: The ideal test design should not come at the expense of what is economically and administratively feasible.

Three considerations help determine the choice of appropriate selection techniques: the level of available resources, the logistical challenges

for achieving test security, and the (e.g., whether questions are to be taken by a large population in Sweden is about the test designed for the Swedish system 50 times the resources as the test designed for the Swedish system. An appropriate system chosen for Sweden would individually design and grade selection tests. An appropriate selection system for China would be machine scorable, whether multiple-choice responses, or in performance description and so forth. The key ingredient is efficiency given the level of available resources.

Since the Second World War, the examination system in China has changed radically in contrast to the former Soviet Union and other nations that have not kept pace. Today, as the Chinese Unified State Examination (USE) is implemented in the former Soviet Union, each faculty would administer examinations in mathematics and physics, examinations would be taken only in person at the institution. This system of selection was unfair because those who could not afford to travel. The effect was to reduce the opportunity. The effect was to reduce the opportunity for every institution. This could delay entry into the labor market for thousands of young people by the state to be unimportant. It were often designed by elderly faculty members who had changes in the labor market. The examinations were not designed to date skills and designed tests the standardized.

for achieving test security, and the level of desired public accountability (e.g., whether questions are to be made public). The size of the test-taking population in Sweden is about that of a Beijing suburb. Moreover, the test designed for the Swedish student may have behind it as much as 50 times the resources as the test designed for the Chinese student. The appropriate system chosen for Sweden might be to have each teacher individually design and grade selection examinations. The most appropriate selection system for China would have to be more standardized and machine scorable, whether measured in the proportion of correct responses, or in performance descriptors such as letter grades A, B, C, and so forth. The key ingredient is to construct the best selection test given the level of available resources and expected accountability.

Since the Second World War, the technology of examination administration has changed radically in OECD countries, but in many parts of the former Soviet Union and other parts of the world, these technologies have not kept pace. Today, as the chapter on Russia points out, the new Unified State Examination (USE) is in place. But until recently in the former Soviet Union, each faculty within each higher education institution would administer examinations independently. Even in mathematics and physics, examinations would primarily be delivered orally. They could only be taken in person at the faculty where they were designed. This system of selection was unfair, inefficient, and of low quality. It was unfair because those who could not travel to the examination site had less opportunity. The effect was to limit access to candidates who could afford to travel. It was inefficient because students must take a different examination for every institution to which they applied. Since they could not do this at a single sitting, they must wait for a next test-taking occasion. This could delay entry by a year or more. The opportunity cost for thousands of young people for this waiting period was thought by the state to be unimportant. It was of low quality because questions were often designed by elderly faculty members, isolated from recent changes in the labor market. They placed high weightings on out-of-date skills and designed tests the administration of which could not be standardized.

The heterogeneity of external examination systems described in this volume is impressive. The Matura examinations in Hungary are designed by subject panels, as in France, and contain both oral and written sections. Potential corruption of the oral examinations is minimized by using certified professional assessors. The General Graduating Certificate in Vietnam is a comprehensive assessment of the secondary school performance and qualifies the student who wishes to enter university to sit for a separate university entrance examination.

Schools in some nations, particularly private schools, rely on internationally designed examinations such as the International Baccalaureate (IB) and the Cambridge International Examinations (CIE). These examinations may be considered to provide a level playing field across countries. However, they may be useful only for those who intend to study outside their country of origin when universities utilize local performance standards.

One key distinction among nations is not whether the selection test is multiple-choice or free response, but whether whatever technique chosen can be corrupted. How selection is managed is deeply important for maintaining an equality of education opportunity. Tests that are centrally scored can still be corrupted by leaks. In some parts of South Asia, questions are privately sold to high-paying candidates before the test is administered. As faculty salaries decline in value and higher education institutions seek alternative sources of income, bribery surrounding the admissions process can become a matter of routine. Candidates may even know how much a 'pass' will cost and be expected to bring the cash ahead of time.

FREQUENCY OF CORRUPTION

How common is corruption in education? There are two ways to estimate corruption frequency. One is to ask professionals to rank countries by the perception of corruption. For instance, Transparency International publishes a ranking of countries by education corruption parallel to its ranking of more general corruption. The second is to ask university

students to estimate the likelihood of admitting the circumstances in which corruption occurs. When Bulgarian students were asked to estimate the likelihood of admitting the circumstances in which corruption occurs, 79% had heard of such corruption and hence 79% had heard of such corruption. In Serbia, 16% had heard of such corruption and hence 16% had heard of such corruption. At the same time, 33% of the faculty would change an admission if one were available. In the Kyrgyz Republic, the Kyrgyz Technical University reported high bribery scores were recorded in the country. The exception was the Faculty of Economics, a small, independent liberal arts college in the United States with a strong sanction against the possibility of corruption. There, only 16% of the faculty was 'bribeable'.

Which fields are more likely to be corrupted? A study of the University of Michigan (accredited in Turkey) found that in his institution and commissioned a study to investigate the question about whether faculty members were more likely to accept a bribe in response from one department than another. The study found a higher demand—Economics, Business Administration. When a study was published, faculty expected that the level of corruption on corruption. The anticipation of a study on corruption precipitated a decline in corruption. The study found that in Economics where the level of corruption decreased.

Is educational corruption limited to the United States? Evidence from the United States shows that cheating on tests is pervasive and that 75% of the students admit to some

students to estimate the likelihood that their faculty are corrupt or to admit the circumstances in which they have personally participated in corruption. When Bulgarian students were asked about corruption, 21% said they had 'never heard' of any way to gaining admission illegally and hence 79% had heard of such ways. In Croatia, the figure for never having heard about corruption was 40%; in Moldova, it was 21%; and in Serbia, 16%. At the same time, 33% of the Serbian students thought that faculty would change an admissions score if bribed. Forty percent of the students in Moldova would use an illegal method of gaining admission if one were available. In the Kyrgyz Republic, 68% of the students at the Kyrgyz Technical University described their university as 'briable'. High bribery scores were recorded for every government university in the country. The exception was the American University of Central Asia, a small, independent liberal arts college in Bishkek accredited in the United States with a strong sanctionary infrastructure to counteract the possibility of corruption. There, only 5% of the students thought the college was 'briable'.

Which fields are more likely to be open to bribes, and is the propensity to accept bribes constant over time? The rector of the Kazakh Turkish University (accredited in Turkey) wanted to evaluate the quality of his institution and commissioned student surveys in 2001 and 2005. On the question about whether faculty would accept bribes, the differences in response from one department to another were pronounced. Faculty were more likely to accept a bribe in those departments which were in higher demand—Economics, Business and Law. Once the first survey was published, faculty expected the next survey to contain a question on corruption. The anticipation of this new accountability measure precipitated a decline in corruption. The exceptions were in Law and in Economics where the level of perceived bribery increased rather than decreased.

Is educational corruption limited to developing countries? Evidently not. Evidence from the United States would suggest that the propensity to cheat on tests is pervasive and growing. On most university campuses, 75% of the students admit to some form of cheating.

COSTS AND CONSEQUENCES OF A CORRUPT EXAMINATION SYSTEM

The process of academic selection is the linchpin of any education system and overall national cohesion. It represents the essence of the public good. If the system is corrupt, or widely believed to be corrupt, little else in the education system can be successful. Inattention to corruption in selection will place other aspects of a nation's economic and social ambitions at risk.

If there is significant perceived corruption in education, higher education becomes a less effective means to acquire a high income. In Africa, the marginal effect of higher education on the probability of securing a high income falls by about 70% if the education system is perceived to be corrupt. In Western Europe, the marginal effect of higher education on the probability of gaining a high income falls by about 25% if accompanied by a perception of education corruption. Education corruption adversely affects the relative ability of higher education to keep people out of poverty. Within the Commonwealth of Independent States (CIS), corruption increases the probability that those with higher education become poor. When there is significant corruption, higher education is also a less effective means to reduce poverty.

When the education system is corrupt, private firms are forced to establish additional sorting mechanisms to determine the degree to which job candidates are competent. The use of these sorting devices imposes additional costs on firms and is related to the uncertainty attached to the quality of education attained by students. Students from highly corrupt universities are not considered for technical and professional private sector jobs and even some government jobs, and are screened out of jobs in foreign enterprises. If the best alternative for students from corrupt educational backgrounds is employment in the government sector, the loss of earnings is significant. If students from corrupt educational programs sort into government jobs with the potential for bribes, the private income costs of corruption are reduced, but the social costs remain.

In 2006 interviews with faculty at the Technical State University (TSU) in Georgia, East Kazakhstan, and at the American University in the Kyrgyz Republic. Bribery and corruption were discussed by a senior professor at TSU laid out what he saw as the state of his university (transcribed verbatim).

Admissions were a way to reduce corruption inside the university, corruption was worse in Law and Business than in other fields. To be corrupt in math and physics, the student's mind would study math or physics there is no way to be motivated students. Medicine is the most corrupt. But it is not only because of the high cost of education, the prestige, the diploma.

Why is corruption so prevalent? One reason is the quality of today's students. One profes

We have to change this mentality. It took a hundred years to develop the Soviet Union in ten years. Students still think like the Soviets; many of them have jobs depend on their obtaining a diploma. University focused not on learning. Students don't have to earn their diploma. They have to do is have one. This is a result of five years of our market econ

Will replacing the old oral examination system with a new standardized examination system in Georgia and Kyrgyzstan, the answer is yes.

The new [examination] system is better. I am angry when asked to change the system. The requests depends on how strong the pressure. Before the new sy

In 2006 interviews with faculty were conducted at Tbilisi State University (TSU) in Georgia, East Kazakhstan State University in Kazakhstan, and at the American University of Central Asia in Bishkek, Kyrgyz Republic. Bribery and corruption were a common focus. One senior professor at TSU laid out what he considered the recent history at his university (transcribed verbatim):

Admissions were a way to make money, huge money. But once inside the university, corruption depended upon the department. It was worse in Law and Business and Economics. It makes no sense to be corrupt in math and physics. (Why?). No one in his right mind would study math or physics if he is corrupt. If you go into math or physics there is no work, no jobs, so we get only highly motivated students. Medicine was the first faculty to become corrupt. But it is not only because of demand; there is also the question of prestige, the diploma. (Heyneman, 2007)

Why is corruption so prevalent? One answer has to do with the mentality of today's students. One professor said,

We have to change this mentality. Western countries have had two hundred years to develop their market economies. We have had ten years. Students still think of employment as they did under the Soviets; many of them have already been promised jobs. These jobs depend on their obtaining a diploma. So they come to university focused not on learning but on obtaining a diploma. They don't have to earn their diploma in order to be employed; all they have to do is have one. This is because these are the first formative years of our market economy.

Will replacing the old oral examinations designed by each faculty with a new standardized examination system reduce corruption? In Georgia and Kyrgyzstan, the answer is yes. In the words of another interviewee,

The new [examination] system helps. Students have learned that I am angry when asked to give them a favor. How often I get requests depends on how strict I am. If I am more liberal, I get more pressure. Before the new system, the teacher was overwhelmed

with special requests. The new system helps me to be more objective. It is very difficult to not take account of friends and family, but the result will be a [nation of] semi-professionals.

Do faculty understand the implications of education corruption? In some cases, they do. One professor put it this way:

Corruption is bad. It is clear even for a fool that we are producing not very smart people. Corruption will affect our economy. If we produce a foolish agriculturalist, and he chooses a bad crop, a bad seed, the result will affect all of us. We have suffered before from famine. We can again. There are implications of corruption in which it becomes an internal threat to our economy. This is true for every discipline, chemistry, physics, and mathematics. That is why foreign companies are using all kinds of screening and testing devices for their new employees. Our local companies should do the same thing; and these screening devices should be independent.

Although it was widely acknowledged that the problems of corruption are pervasive, there is also evidence that in some instances, these problems are on the decline. All three countries had recently instituted standardized, computer-scored admissions examinations with the explicit purpose of addressing the problems of corruption in university admissions. One professor at TSU assessed the results in this way:

Before national [standardized] exams corruption was very high and moral pressure [to change grades] was also. The departments which are more prestigious are the worst in terms of corruption: Law, Business, International Relations. Because of the exams however, many things have changed. Students from rural areas and from poor homes are more numerous. When bribery was necessary to enter the university, these students had no chance to enter. Today because of the examinations, they do. This is very new and very good.

SUMMARY

Standardized external examinations serve many important quality-assurance functions. Examinations used for monitoring and for selection

are necessary to all modern nations. can be implemented that exacerbate cohesion. To design examination system opposite effect. Good examination system can help choose talent fairly, it can and it can give confidence in the government system.

are necessary to all modern nations. There are many ways in which these can be implemented that exacerbate social tension and detract from social cohesion. To design examination systems well, however, may have the opposite effect. Good examination design can lead to social cohesion, it can help choose talent fairly, it can engender trust in public institutions, and it can give confidence in the general conduct of the public school system.

BIBLIOGRAPHY

- Greaney, V., & Kellaghan, T. (1995). *Equity issues in public examinations in developing countries*. Washington, DC: World Bank.
- Heyneman, S. P. (1979). Why impoverished children do well in Ugandan schools. *Comparative Education*, 15(2), 175–185.
- Heyneman, S. P. (1987). Uses of examinations in developing countries: Selection, research and education sector management. *International Journal of Educational Development*, 7(4), 251–263.
- Heyneman, S. P. (2000). From the party/state to multi-ethnic democracy: Education and social cohesion in Europe and the Central Asia Region. *Educational Evaluation and Policy Analysis*, 21(4), 173–191.
- Heyneman, S. P. (2004). Education and corruption. *International Journal of Education Development*, 24(6), 638–648.
- Heyneman, S. P. (2007, September). Three universities in Georgia, Kazakhstan and Kyrgyzstan: The struggle against corruption and for social cohesion. *UNESCO Prospects*, 3, 305–318.
- Heyneman, S. P., Anderson, K. A., & Nuraliyeva, N. (2008). The cost of corruption in higher education. *Comparative Education Review*, 51(2), 1–25.
- Heyneman, S. P., & Fagerlind, I (Eds.). (1988, January). *University examinations and standardized testing* (Technical Paper No. 78). Washington, DC: World Bank.
- Heyneman, S. P., & Loxley, W. (1983). The effect of primary school quality on academic achievement across twenty-nine high- and low-income countries. *American Journal of Sociology*, 88(6), 1162–1194.
- Heyneman, S. P., & Ransom, A. (1990). Using examinations to improve the quality of education. *Educational Policy*, 4(3), 177–192.
- Kirmani, N. (1984). Effects of income inequality on educational attainment in developing countries. *International Journal of Educational Development* (Vol. 4). Washington, DC: IMF.
- Klitgard, R. (1986). *Elitism and meritocracy*. Baltimore: John Hopkins University Press.
- Piñera, S., & Selowsky, M. (1981). Education, income inequality, and the misallocation of resources. *Development Economics*, 8(1), 111–133.
- Plomp, T., & Voogt, J. (1995). Curriculum change in modern Russian education. In *Curriculum Change in Education in Russia* (pp. 1–10). London: Falmer Publishers.
- World Bank (1995). *Russia: Education and social cohesion*. Washington, DC: World Bank.
- Xueqin, J. (2001, 14 September). *The Chronicle of Higher Education*.

- Kirmani, N. (1984). Effects of increased market access on exports of developing countries. *International Monetary Fund Staff Paper No. 31* (Vol. 4). Washington, DC: IMF.
- Klitgard, R. (1986). *Elitism and meritocracy in developing countries*. Baltimore: John Hopkins University Press.
- Piñera, S., & Selowsky, M. (1981). The optimal ability-education mix and the misallocation of resources within education. *Journal of Development Economics*, 8(1), 111-131.
- Plomp, T., & Voogt, J. (1995). Curriculum standards and assessment in modern Russian education. In F. J. H. Mertens (Ed.), *Reflections on Education in Russia* (pp. 97-113). Antwerp, Belgium: Acco Publishers.
- World Bank (1995). *Russia: Education in the transition*. Washington, DC: World Bank.
- Xueqin, J. (2001, 14 September). Corrupt admissions alleged in China. *The Chronicle of Higher Education*, p. A51.