

Copyright is retained by the first or sole author, who grants right of first publication to the **EDUCATION POLICY ANALYSIS ARCHIVES. EPAA** is a project of the [Education Policy Studies Laboratory](#).

Articles appearing in **EPAA** are abstracted in the *Current Index to Journals in Education* by the [ERIC Clearinghouse on Assessment and Evaluation](#) and are permanently archived in *Resources in Education*.

Volume 12 Number 29

June 28, 2004

ISSN 1068-2341

High-Stakes Testing in the Warm Heart of Africa: The Challenges and Successes of the Malawi National Examinations Board

Elias Chakwera
University of Massachusetts Amherst and Domasi College

Dafter Khembo
**University of Massachusetts Amherst and Malawi National
Examinations Board**

Stephen G. Sireci
University of Massachusetts Amherst

Citation: Chakwera, E., Khembo, D., Sireci, S., (2004, June 28). High-Stakes Testing in the Warm Heart of Africa: The Challenges and Successes of the Malawi National Examinations Analysis Archives, 12(29). Retrieved [Date] from <http://epaa.asu.edu/epaa/v12n29/>.

Abstract

In the United States, tests are held to high standards of quality. In developing countries such as Malawi, psychometricians must deal with these same high standards as well as several additional pressures such as widespread cheating, test administration difficulties due to challenging landscapes and poor resources, difficulties in reliably scoring performance assessments, and extreme scrutiny from political parties and the popular press. The

purposes of this paper are to (a) familiarize the measurement community in the US about Malawi's assessment programs, (b) discuss some of the unique challenges inherent in such a program, (c) compare testing conditions and test administration formats between Malawi and the US, and (d) provide suggestions for improving large-scale testing in countries such as the US and Malawi. By learning how a small country instituted and supports its current testing programs, a broader perspective on resolving current measurement problems throughout the world will emerge.

Malawi is a small landlocked country in Africa, south of the Equator covering an area of 118, 484 square kilometers of which 20% is water. The country is bordered to the North and North-East by the Republic of Tanzania and to the East, South and South-West by the Republic of Mozambique. The Republic of Zambia forms the Western border.

Malawi gained independence from Britain in 1964 and operated under one-party state until 1994 when a multiparty government was elected. The population of Malawi is estimated at 11 million people. About 46% of the population consists of children and youth less than 15 years of age.

The literacy level is estimated at 40% of the adult population (29% female and 48% male). Because of poor levels of literacy, there has been rampant poverty. This situation prompted the new government to introduce Free Primary Education (FPE) in 1994 as a tool for alleviating poverty. This innovation received an overwhelming support from the public in that the enrolment in primary schools increased from 1.8 million to about 3 million pupils. This meant an increased demand for resources that support learning, including assessment.

In this paper, we provide an overview of the national testing systems that support FPE in Malawi. The psychometric, logistic, and political factors affecting this system are discussed, as are the similarities and differences between educational testing in Malawi and in the United States. We begin with a description of the Malawi National Examinations Board.

A Brief History of the Malawi National Examinations Board

In 1969, the Malawi parliament enacted a law that created the Malawi Certificate Examination Board (MCE Board). This Board was charged with the responsibility of developing and administering the Malawi Certificate of Education (MCE) examination in conjunction with the Associated Examining Board (AEB) of the UK. The first such examination was administered in 1972. Prior to 1972, school leavers in Malawi were taking the Cambridge Overseas School Leaving Examination from the UK.

Seven years later, the MCE Board became the Malawi Certificate Examinations and Testing Board (MCE and TB). The MCE and TB continued to administer the MCE examinations with the AEB until 1989 when the handover was completed.

Following an evaluation of examinations in Malawi in 1984, it was decided that all public examinations should be developed and administered by one central authority. Consequently, in 1987, parliament approved legislation merging the examinations section of the Ministry of Education with the MCE and TB, thus forming the Malawi National Examinations Board (MANEB), which currently operates the major educational testing programs in Malawi. In addition, MANEB took over the responsibility of developing and administering Teacher Certificate Examinations and Craft Examinations for technical schools.

Malawi's Education System

The Malawi education system consists of three levels: primary, secondary, and tertiary. The primary education level is an eight-year cycle running from Standards (grades) 1 to 8. Standard 8 is an equivalent of Grade 8 in the United States. At the end of Standard 8 pupils take the Primary School Leaving Certificate Examination (PSLCE).

Secondary education lasts for four years, running from Form 1 to Form 4, which are the equivalents of Grades 9 to 12 in the U.S. Two national examinations are administered at this level: the Junior Certificate Examination (JCE) at the end of junior secondary in Form 2 and the Malawi School Certificate Examination (MSCE) at the end of senior secondary in Form 4.

Tertiary education is usually four years particularly at the University level, although there are other tertiary educational institutions that offer courses and programs for less than four years. Teacher training for primary school teachers is usually two years, while technical training may last for four years or less depending on the field of specialization. Access to tertiary education is still very limited because of scarcity of places at that level.

A description of these major educational testing programs follows. Table 1 presents a brief summary of these programs that includes the grade at which they are administered, the purpose of the test, as well as the number of students sitting for each test and the passing percentages for the most recent administration on which data are available.

| Exam | Grade Administered | Purpose | # Examinees 2001 | % Pass 2001 |
|-------------|---------------------------|-----------------------------------------------|-------------------------|--------------------|
| PSLCE | 8 | secondary school entrance | 161,786 | 26.33* |
| JCE | 10 | 10th grade exit; basic employment certificate | 82,530 | 57.21 |
| MSCE | 12 | High school exit; postsecondary admissions | 61,856 | 18.01 |

*This represents proportion of PSLCE examinees selected to secondary school.

Major Testing Programs in Malawi

MANEB develops and administers three major national school examinations: PSLCE, JCE and MSCE. A brief description of these examination programs is provided below.

The PSLCE

PSLCE terminates the primary cycle. Its results are used for certification and selection into Form 1 of the secondary education. The results are reported in letter grades A-F, where A denotes excellent performance and F a fail. Five subjects are offered at this level. These are English, Mathematics, Primary Science, Chichewa (a local language), and Social Studies. For selection purposes, students are ranked within their districts. Each district is allocated a certain number of Form 1 places in national secondary schools. The district quota depends on the proportion of candidates in the district in relation to the national total. The remaining candidates are considered for places in District Secondary Schools and Community Day Secondary Schools (CDSSs). At each selection level, boys and girls are considered separately to ensure gender equity (i.e., within-group norming). A single merit list would result in boys getting a disproportionate number of secondary school places, since they generally perform better than girls.

For many people, the certification aspect of the PSLCE is not as important as its selection function, because the certificate can no longer be used for employment purposes as is the case with MSCE. Therefore the pupils are under pressure to perform well enough to be selected into secondary education. A longitudinal sampling of the numbers of students taking the PSLCE and the numbers of students passing it, are presented in Table 2. As these data show, the demand for secondary education has always outstripped the available places.

| Year | PSLCE Entry | # Passing | % Selected |
|-------------|--------------------|------------------|-------------------|
| 1977 | 47,317 | 4,854 | 10.3 |
| 1987 | 95,631 | 6,894 | 7.2 |
| 1997 | 128,379 | 9,170 | 7.1 |
| 2001 | 161,786 | 42,600 | *26.3 |

*CDDs were instituted in 1998. This figure includes the 5% of students who went to national secondary schools and the 21% who went to CDSS's.

Formal education for those who fail to get into national secondary school

effectively stops at Standard 8. Before 1999, some pupils received secondary school tuition through Distance Education Centers (DECs), which have since been turned into Community Day Secondary Schools (CDSS). The increased transition rate in 2001 is a reflection of the inclusion of students who go to CDSSs. However, the quality of learning at CDSSs is considered inferior to the conventional secondary schools in terms of materials and number and quality of teachers in these schools. Private secondary schools also provide secondary education but are too expensive for most parents, and the majority of them are not well resourced. Until now the competition is high for places in the conventional secondary schools where the government subsidizes tuition and the schools are better resourced than the private and CDSSs. This is what makes PSLCE a high-stakes examination: it determines one's opportunity for higher, better and affordable education.

JCE

The JCE is administered after two years of secondary education. Originally this examination was meant to assess skills and knowledge leading to gainful employment and further education in senior secondary school. Twenty-two subjects are offered for this examination. Candidates must pass at least six of them including English to qualify for a certificate and proceed to Form 3 (11th grade). The examination results are shown in Table 3.

| Table 3 JCE results | | | |
|--------------------------------|--------------|-----------------|-----------------|
| Year | Entry | #passing | %passing |
| 1998 | 74,122 | 51,878 | 70.0 |
| 1999 | 69,148 | 63,133 | 91.3 |
| 2001 | 82,530 | 47,218 | 57.21 |

In 1997 the government phased out the JCE as a minimum requirement for entry into civil service. However, due to intense job competition, it is still used as a hiring criterion for some blue-collar jobs. Furthermore, a student must pass the JCE before sitting for the MSCE. The competition for Form 3 places is no longer stiff since there are equal numbers of places in junior and senior secondary sections, and so all who pass the JCE are automatically promoted to Form 3.

MSCE

MSCE, which is equivalent to High School Diploma in the US, is administered at the end of secondary education. The examination results are used for certification (i.e., certifying successful completion of secondary education) and selection into the university and other tertiary institutions. A total of 21 subjects are offered at this level. Each subject is graded on a nine-point scale using the following standards:

Grades 1-2 for distinction;

3-6 for credit;
 7-8 for general pass; and
 9 for fail.

To qualify for a certificate a candidate must pass at least any combination of six subjects including English, and one of the grades must at least be a credit pass. An MSCE certificate can also be awarded if a candidate passes five subjects including English, and three of which are at least credit passes.

The grading process for the MSCE makes the following assumptions:

- The examinations are equivalent across years in terms of difficulty level, content covered, and skills examined;
- The test administration conditions are uniform from year to year;
- The student cohorts taking the examination each year are randomly equivalent.

For candidates to be considered for selection into the university, they must have earned credit or distinction on at least six exams, one of which must be English. A pass with at least credit grade in English ensures that the candidates have adequate communication skills to fully participate in college lectures. Currently, the University of Malawi admits only 0.3% of the secondary school leavers, which illustrates the stiff competition. In addition, the MSCE certificate has become the minimum qualification for gainful employment. Because of these two functions – selection and certification – there is a lot of pressure on the students to pass the examination, making it extremely high-stakes.

Over the last decade there have been declining trends in pass rates on the MSCE as shown in Table 4. This trend has been attributed to several factors. One factor is indiscipline due to some student’s misinterpretation of their newly found democracy, human rights, and freedom (Malunga, et al. 2000). For example, Kuthemba-Mwale, Hauya and Tizifa (1996) observed “a general lack of interest among students to do academic school work.” Other factors that contribute to poor examination results include increasing number of students without a corresponding increase in instructional resources, and inadequate and under-qualified teachers. As Malunga et al. reported, there are only 4,998 secondary school teachers instead of 12,000 required by the system. In addition, it was also observed that 67.2% of the teachers were under-qualified.

| Year | MSCE Entry | #passing | Pass (%) |
|-------------|-------------------|-----------------|-----------------|
| 1992 | 10753 | 5653 | 44.4 |
| 1993 | 13254 | 7123 | 46.7 |
| 1994 | 16264 | 7871 | 43.1 |
| 1995* | 23219 | 7421 | 29.4 |
| 1996 | 24213 | 8036 | 30.7 |

| | | | |
|------|-------|-------|------|
| 1997 | 26543 | 6740 | 23.6 |
| 1998 | 35438 | 6329 | 17.9 |
| 1999 | 36732 | 5536 | 14.3 |
| 2001 | 61856 | 11143 | 18.0 |

*Private secondary schools opened in 1995, which may explain the large increase in number of students tested in this year. Other increases are harder to explain, but may be due in part to students who failed the exam in previous years sitting again for the exam.

Practical, Political, and Psychometric Issues Confronted by MANEB

MANEB administers the three national examinations described above besides other responsibilities such as development and administration of Malawi Craft Examinations and teacher certification examinations. In all these examinations the numbers of candidates and examination centers have been increasing every year. This has resulted in a number of administrative challenges, which we describe next.

Dealing with Limited Resources

From the three tables above, it is apparent that MANEB's volume of work and expenditure increase every year. MANEB's source of funding is largely government subvention, whose annual increase does not match the increased costs of administering examinations, especially in view of increasing inflation.

The major areas of expenditure with regard to increasing number of examination centers that are scattered throughout the country are delivery and collection of examination materials to and from all centers, scoring, and invigilation (supervision of examinations in the centers). In explaining the delay in releasing the 2001 JCE and MSCE examination results, MANEB's Executive Director made reference to inadequate funding as a major cause of some of the problems facing examination administration (The Nation, 2002a). For instance, the scoring process was disrupted by persistent strikes by the scorers, who were demanding more money from MANEB (The Nation, 2002b). It is now being proposed that examinations should be administered much earlier during the school year to allow adequate time for processing the results. The likely consequence of this proposal is that it will be difficult for schools to adequately cover the syllabi before the examinations are administered.

As a way of reducing costs due to delivery and collection MANEB has introduced Examination Distribution Points from where surrounding schools come to collect examination materials for their schools on the daily basis.

Examination Security Concerns

One of the major concerns regarding security of examinations is leakage. In

some centers examination envelopes have been intentionally opened before the specified time, and contents exposed for the benefit of candidates. In extreme cases the prematurely exposed examination papers have been duplicated and sold to the candidates. Such a practice led to the cancellation of the 2000 examinations, and another set of examination papers had to be developed and distributed. In an attempt to deal with this problem MANEB established Examination Distribution Centers for storage of examination materials which are guarded by police officers.

Another area of concern is cheating. Cheating takes place in many forms including impersonation, giving extra time, substitution where a candidate's script is replaced by one prepared by a more competent person, referring to books, copying from each other, copying from a common source, teachers dictating answers to the class, etc. As a way of curbing cheating during examinations, MANEB carries out spot checks during examinations, but these are done to a limited extent due to shortage of personnel, vehicles, and finances. MANEB also provides civic education to the general public about the dangers of examination malpractice, since in some cases cheating involves the general public. Sometimes MANEB applies sanctions such as nullification of results, deregistration of examination centers, withholding results, and prosecuting the culprits if examinations regulations have been infringed. In addition, MANEB uses external invigilation system whereby a teacher from a different school invigilates examinations. The headteacher of the schools is the overall supervisor of examinations at the school. MANEB also prints examination papers outside the country to curb possibility of leakage originating from MANEB offices.

Dealing with Public and Political Pressure

MANEB works under considerable pressure because of the high-stakes nature of its examinations. There are many groups that directly influence the way in which MANEB operates. For example, The Ministry of Education, which directs all MANEB's activities, requires timely release of examination results so that the school calendar is not disturbed. For MANEB to administer all examinations and process the results within a single school year, means that some exams must be administered well before the end of the school year to allow time for processing the results. Consequently, the examinations are likely to test material that has not yet been covered in classes. This causes anxiety to both the examinees and their teachers.

Another significant problem faced by MANEB is cash flow. MANEB's cooperating partners such as invigilators, supervisors, and scorers, want to be paid promptly for the work they do. For some time MANEB has not been able to make prompt payments due to unavailability of funds. This has soured the relationship between MANEB and its partners who sometimes wait for up to two years before they are paid.

Another problem to be dealt with by MANEB is score challenges. When examinees and their guardians do not agree with the examination results, they request a re-scoring of the exam. Given that the majority of MANEB exams involve constructed-response items that are scored subjectively, score

challenges and re-scoring of exams is time-consuming and expensive.

Like many educational testing programs in the United States, MANEB is also a target for criticism in the popular press. Quite often, the press reports on the tension between MANEB and its cooperating partners, and they often highlight the negative aspects. For example, commenting on the delay in releasing the 2001 examination results, *The Nation* newspaper reported:

MANEB and its parent ministry should take responsibility for the inconvenience that has been created and take necessary remedial action. Is it really impossible to conduct incident-free examinations whose results are released in good time? We believe it is possible and MANEB can only justify its existence by doing no less. (The Nation, 2002a)

In 1999, the negative coverage of examination results in the press prompted the State President to institute a commission of inquiry into the causes of poor MSCE results (Malunga et al. 2000). The opposition parties took advantage of the poor results to criticize government education policies.

Measurement Issues

Curricular Validity and Teaching to the Test

The examinations in Malawi are so important that they have assumed a “gate-keeping” role in the system. Because of this importance, the examinations exert considerable influence on what goes in schools. Although the curriculum has generally incorporated issues of the cognitive, psychomotor, and affective domains, examinations mainly focus on the cognitive domain. With so much emphasis on passing examinations it is not surprising that the instruction has become examination oriented. Thus, curricular validity of Malawian exams is a contentious issue.

MANEB is aware of the demands of the curriculum, but is unable to meet them because of inadequate resources. In some subjects the number of examination papers was reduced and practical-work in some subjects was scaled down to cut down costs. For example, assessment by project method (Note 1) had to be discontinued. This resulted in a mismatch between the examinations and course objectives because only selected parts of the curriculum are assessed, and therefore taught.

When projects were removed from assessment to cut down costs on project inspection and scoring, the schools no longer felt the need to teach by project method, even though it remained an important part of the national curriculum. If teachers are to cover the whole curriculum, then examinations must cover the curriculum. By not covering some parts of the curriculum, the examinations limit the scope of instruction.

Scoring Free-Response Items

Most MANEB exams use free-response items. Because of the large numbers of candidates the items are scored only once, with 10% re-scored by the Chief

Examiners who supervise the scoring exercise. This raises the problem of examination reliability. For each of the three major exams, over 800,000 student papers are scored. Scoring each exam takes about four weeks and involves considerable human and monetary resources.

The move by MANEB towards objective assessment using multiple-choice examinations was met with strong resistance from the general public who felt that the multiple-choice examinations would dilute the education quality. As a way of improving the reliability of the examination scores, MANEB put in place a number of measures such as training of scorers, pre-scoring exercise, standardization of scoring, script checking, and data entry verification. All these measures are meant to ensure that no errors are made during scoring of scripts and processing of examination results. However, even with this rigorous error-searching process, some errors go undetected and are discovered at the re-scoring stage, and only if such a request is made. The candidates' requests for re-scoring are attended to only on payment of a re-scoring fee.

High-Stakes Educational Tests in Malawi and the U.S.: Similarities and Differences

The preceding sections outlined the major issues confronted by measurement professionals in Malawi. Interestingly, most of these issues are policy-oriented or deal with the practical problems involved in test administration. Many of these issues are also confronted by measurement professionals in the U.S., but U.S. psychometricians appear to be more focused on technical issues, particularly those related to the reliability and validity of test scores. In this section, we discuss the similarities and differences between testing in Malawi and the U.S. with respect to both psychometric and educational policy issues.

Testing and Educational Reform: An Important Area of Commonality

It is interesting to note that educational reform movements in both Malawi and the U.S. use standardized tests as the primary mechanism for accountability and certification goals. Almost all states within the U.S. have a state-mandated assessment system (Linn, 2000), which is used to evaluate school districts, schools, teachers, and students. In many states, such as Massachusetts, state-mandated tests are also used (a) to encourage teachers to align their instruction with state curriculum frameworks and (b) for certification functions such as granting high school diplomas.

The Malawi national examination system has also been at the heart of its educational reform movement. For many schools where instructional resources are scarce or nonexistent, the syllabi associated with MANEB tests represent significant instructional resources for teachers. However, it is interesting to note that the reform movement in Malawi is a national movement, instituted by national laws, and the tests are developed and administered by a national testing agency. This situation is quite different than the U.S., where efforts to create nationally mandated tests continually fail. States want the authority to decide what is taught and what is tested and so even efforts to institute the

Voluntary National Test have been met with resistance. Only tests associated with the National Assessment of Educational Progress (NAEP) have been accepted by the states, perhaps because no student-level data are reported, the effect of these tests on state curricula is minimal, and there are absolutely no stakes at all for the students who take them. Thus, although both countries use tests as the primary data source in their educational accountability and certification systems, the difference between “local” and national control is striking.

It is also interesting to note that “teaching to the test” is seen as a significant issue in both countries. In both Malawi and the U.S., there are critics who see mandated testing as a weakening of the curriculum, while others praise this practice as an effective means for improving instruction. It appears that the use of high stakes tests to improve classroom instruction has supporters and detractors on both continents.

High Stakes Versus Really High Stakes

High stakes testing receives a great deal of attention in the popular press and educational policy journals within the U.S. The two most common issues are the appropriateness of admissions tests for making postsecondary admissions decisions and the appropriateness of using standardized tests for awarding high school diplomas. Relatively poor performance on the SAT or ACT can certainly inhibit a student’s chances of getting accepted into a postsecondary institution, particularly the institution of her or his choice. Also, not receiving a high school diploma due to failing an exit exam also has serious, negative consequences for students in the U.S. But the stakes associated with the PSLCE and the MSCE in Malawi are much higher. With respect to postsecondary admissions in the U.S., the community college system is available to students who cannot get into a four-year college and most of these schools have open enrollment policies that do not require admissions test scores. More importantly, there is not a huge discrepancy between the number of seats available for postsecondary education and the number of students who seek it. With respect to high school graduation and postsecondary admission, the U.S. offers a multitude of well-paying jobs that do not require high school or college degrees. Furthermore, second-chance programs such as the Tests of General Educational Development and those found in adult basic education provide opportunities for adults who did not complete high school to earn a high school diploma later in life and continue their education.

The situation in Malawi is very different. Students who do not pass the PSLCE do not even make it into secondary school. Even for those who do pass, there are limited spaces in the national secondary schools and CDSSs. Last year, only about 5% of primary school students were placed into the coveted national secondary schools and only 20% more were placed into the lower quality CDSSs. Most of the other 75% of students will never have the opportunity to pass the JCE and MSCE and be able to compete for the best jobs in Malawi. For many Malawians, passing the JCE and the MSCE makes the difference between a life of self-sufficiency and a life of poverty. Passing the MSCE makes numerous career options possible that cannot be attained through other routes. For example, the national government requires an MSCE certificate for civil

service employment. Therefore, “high-stakes testing” has a more pronounced meaning in Malawi. The national educational tests are the sole criteria for academic certification and the stakes associated with the tests include starvation versus prosperity.

Equity Issues in Assessment

In the U.S., the equity issues associated with educational testing most commonly involve ensuring or evaluating test fairness with respect to (a) racial, ethnic, or linguistic minority groups; (b) females and males; and (c) individuals with disabilities. In Malawi, only sex differences in test performance receive significant attention by researchers, politicians, and the popular press.

At first, an outsider may think equity issues associated with ethnicity are not relevant to Malawi because all citizens are African. However, although the ethnic composition of the country is much more homogeneous than the U.S., there are still significant differences with respect to tribal origins, religion, and language. However, test bias with respect to these groups has not been extensively studied. Systematic study of differences across linguistic groups also has not been conducted, which is unfortunate since most Malawians primarily speak Chichewa, even though English is its official language (two other languages, Tumbuka and Yao, are also the native tongue for hundreds of thousands of Malawians).

The issue of accommodating tests for individuals with disabilities has received much less attention in Malawi than in the U.S. There is no acknowledgement of students with learning disabilities in the educational system and so granting extended time on tests to such students, which is common in the U.S., is not even on the radar screen. However, MANEB does make Braille tests available to students with visual disabilities and provides 1/6 additional time for such students to take the tests.

In the U.S., equity issues are at the forefront of educational assessment policy debates. When achievement differences are found across racial/ethnic groups on educational tests, researchers, lawmakers, and policy analysts are often divided about what should be done. Claims of test bias against minority groups have led to the abandonment of some educational tests, but affirmative action practices (e.g., using different standards for selecting minority and non-minority candidates) have not stood up to legal scrutiny (Green & Sireci, 1999; Sireci & Green, 2000). These policy issues have led psychometricians and other researchers to focus much of their research on issues of adverse impact and test bias. For example, studies of differential predictive validity and differential item functioning are common in the U.S., but are practically non-existent in Malawi.

An interesting difference between Malawi and the U.S. with respect to equity in assessment is the way they handle sex differences on educational tests. In the U.S., within-group norming (Note 2) practices have been outlawed for organizations that receive federal funds, which include virtually all accredited educational institutions and all governmental agencies. Thus, adjusting for performance differences across males and females is not conducted. In Malawi,

performance differences between females and males on educational tests are more pronounced. Furthermore, the proportions of females at secondary and postsecondary schools are well below that of males. Thus, colleges and universities struggle to admit qualified females. To address educational opportunity differences across the sexes, Malawi secondary schools, colleges, and universities rank females and males separately so that the highest-ranking females will be accepted over males that may have scored higher on a test. Thus, given the same equity issue, the two countries made completely opposite policy decisions. This difference stems not so much from philosophical differences in assessment or admissions equity, but from differences in the numbers of women remaining in school after the primary grades.

Use of Item Formats

As described above, MANEB exams use predominantly constructed-response items. Multiple-choice items are used on some exams, but the public perception is that such items dumb down the curriculum and are not effective for measuring important academic knowledge and skills. These criticisms have also been raised in the U.S., but the psychometric community has worked hard to educate the public about the benefits of multiple-choice items (e.g., increasing score reliability and content coverage, measuring higher-level skills, reduced scoring costs and reduced testing time) as well as the limitations of constructed-response items (Note 3) (lack of content coverage, task specificity, reduced reliability, higher scoring costs). In the U.S., the majority of educational tests use either only multiple-choice items, or a combination of multiple-choice and constructed response items. These practices reflect a desire to ensure adequate levels of score reliability and content validity while keeping down scoring costs. In Malawi, construct representation is emphasized at the expense of score reliability, testing time, score reporting time, and scoring costs.

Computer-Based Testing

Another striking difference between educational assessment in Malawi and the U.S. is the amount of attention paid to computer-based testing (CBT). In the U.S. almost all testing programs are moving towards computerized administration of their tests or are considering the use of computers in improving their assessment systems (Zenisky & Sireci, 2002). Conferences within the educational measurement community feature programs that are dominated with CBT issues such as computerized-adaptive testing, innovative item types, and automated scoring of constructed-response items. These topics are not receiving considerable attention in Malawi, primarily due to the lack of computer resources within the country.

Measurement Community

Another huge difference between the U.S. and Malawi is the presence of a significant educational measurement community. In the U.S. there are thousands of measurement professionals who meet and interact regularly. For example, there are approximately 3,000 members of the National Council on Measurement in Education (NCME) and even more members of the

Measurement, Evaluation, and Statistics Division of the American Educational Research Association (AERA). The Psychometric Society and the measurement and statistics division of the American Psychological Association (APA) also provide national forums for measurement professionals. In Malawi, the measurement community is much younger and much smaller. For example, there is no Malawi equivalent of the *Standards for Educational and Psychological Testing* (AERA, APA, NCME, 1999). However, in 2001, a grant from USAID to the University of Massachusetts Amherst (UMASS) established a program to build educational measurement expertise within Malawi. Currently, nine measurement professionals from Malawi are receiving doctoral or master's degrees in psychometrics from UMASS and an educational measurement program is being reinforced at the University of Malawi. This development should bring measurement practices and research areas across the two countries closer together in the future.

Technical Versus Practical Measurement Issues

In addition to differences in the attention paid to differential predictive validity and differential item functioning, there are also significant differences between Malawi and the U.S. with respect to the issues that receive the most attention as part of the normal operating procedures of testing agencies. For example, in the U.S., procedures for scaling educational tests are widely researched and item response theory (IRT) is a common procedure for scaling educational tests. Furthermore, tests administered in different years are typically equated onto a common scale to ensure differences in test difficulty are taken into account when monitoring student progress and awarding credentials. In Malawi, IRT is not used at all, and tests are not equated across years. Instead, different test forms are assumed to be equivalent in content and difficulty. Another significant difference is in procedures used to set standards on educational tests. In the U.S., standard setting is one of the busiest areas of research and new methods appear continuously (e.g., Cizek, 2000). In Malawi, standard setting is conducted in a less systematic fashion drawing from subjective estimates of test difficulty and student cohort differences.

Due to the higher stakes, more limited resources, and limited technical expertise, the measurement issues that get the most attention in Malawi are more logistical. Reducing cheating is a significant issue, since it is widespread and it represents a significant threat to the validity of exam scores. Developing, administering, and scoring the exams essentially exhausts the personnel and financial resources of MANEB and so there is little time or resources to conduct research on test validation.

Conclusions: Testing Collegiality Around the World

This paper illustrates how different countries deal with common measurement issues, as well as those that are unique to their own situation. Many of the practical problems in measurement are universal and so much can be learned from what other countries are doing. For example, the U.S. can learn from Malawi about successful implementation of large-scale performance assessment and about alternative strategies for achieving equity in test-based

admissions decisions. Malawi can learn technical measurement solutions to problems such as scaling, equating, standard setting, and item and test bias research.

By building international collegiality within the measurement community we will be better positioned to help each other tackle our significant measurement problems. For example, there is much that could be done in both countries to build computerized systems for test delivery that could reduce cheating and test administration costs. Also, measurement programs in the U.S. could do more to reach out and train professionals in Malawi. This expertise could then be extended to other countries in Africa through the measurement program at the University of Malawi and through similar programs that could be developed in other countries. Quality educational systems need quality assessments. Through the process of building measurement expertise in developing countries such as Malawi, we can help these countries improve their educational systems.

Notes

This research was entirely collaborative and the order of the authors is alphabetical.

Correspondence concerning this article should be addressed to Stephen G. Sireci, Center for Educational Assessment, School of Education, University of Massachusetts, Amherst, MA 01003-4140. E-mail correspondence may be sent to Sireci@acad.umass.edu.

1. Assessment by project refers to embedded assessments where students complete hands-on projects throughout the school year that were graded by MANEB. One assessment that was cancelled was an agricultural project where MANEB officials visited farm sites and evaluated students' agricultural projects.
2. In within-group norming, candidates within a group (say, male or female) are rank-ordered with respect to everyone else in the group. Then, the ranks of the candidates are treated as if they were interchangeable. For example, the highest-ranking female would be considered equivalent to the highest-ranking male, even if their scores on a test were very different.
3. See Dunbar, Koretz, & Hoover, 1991; Linn & Burton, 1994; Wainer & Thissen, 1998, for empirical studies of the advantages and disadvantages of these different item formats.

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.
- Chimwenje, C. (1995). *Secondary School Curriculum Review in Malawi: Implications for Assessment*. Paper presented at the 13th annual conference of the Association for Educational Assessment in Africa, Pretoria, South Africa.
- Chimwenje, C., & Khembo, D. J. (1994). *The Role of Examinations in Educational Change: Malawi's experience with Examination Reform*. Paper presented at the 12th annual conference of the

Association for Educational Assessment in Africa, Accra, Ghana.

- Cizek, C. J. (Ed.) *Standard setting: Concepts, methods, and perspectives*. Mahwah, NJ: Lawrence Erlbaum.
- Dunbar, S. B., Koretz, D. M., & Hoover, H. D. (1991). Quality control in the development and use of performance assessments. *Applied Measurement in Education*, 4, 289-303.
- Green, P. C., & Sireci, S. G. (1999). Legal and psychometric issues in testing students with disabilities. *Journal of Special Education Leadership*, 12(2), 21-29.
- Kuthemba Mwale, J. B., Hauya, R., & Tizifa, J. (1996). *Secondary School Discipline Study*, Zomba, Centre for Educational Research and Training.
- Linn, R. L. (2000). Assessments and accountability. *Educational Researcher*, 29(2), 4-16.
- Linn, R. L., & Burton, E. (1994). Performance-based assessment: Implications of task specificity. *Educational Measurement: Issues and Practice*, 13(1), 5-8, 15.
- Malunga, L. B. et al. (2000). *Presidential Commission of Inquiry into the Malawi School Certificate of Education (MSCE) Examination Results*, Lilongwe.
- Ministry of Education. (1998). *Education Statistics*, Lilongwe.
- Ministry of Education. (1999). *Education Statistics*, Lilongwe.
- Sireci, S. G., & Green, P. C. (2000). Legal and psychometric criteria for evaluating teacher certification tests. *Educational Measurement: Issues and Practice*, 19(1), 22-31, 34.
- The Nation (2002a, February 13). *Delay in exam results create more problems*.
- The Nation (2002b, Mmarch 1). *Release of exam results still cloudy*.
- Zenisky, A. L., & Sireci, S. G. (2002). Technological innovations in large-scale assessment. *Applied Measurement in Education*, 15, 337-362.

About the Authors

Elias W. J. Chakwera is Deputy Principal of Domasi College of Education in Malawi, where he teaches courses in Testing, Measurement and Evaluation. He is also a doctoral student at the University of Massachusetts Amherst. His areas of expertise and interest include evaluation of educational assessments, teacher education and development, and training teachers through distance education. His recent research activities include studies on content validity, test score generalizability, consequential validity, and teacher upgrading through distance education in Malawi.

Dafter J. Khembo is currently a doctoral student in Testing & Measurement at the University of Massachusetts Amherst. He received his B.Ed from the University of Malawi in 1986 and MA (Education) from the University of London Institute of Education in 1991. He is presently an employee of the Malawi National Examinations Board where he works as a Research & Test Development Officer. His areas of interest include: standard setting, differential item functioning, and test score equating.

Stephen G. Sireci is Associate Professor of Education and Co-Director of the Center for Educational Assessment at the University of Massachusetts Amherst, USA. His areas of expertise include educational test development and the evaluation of educational assessments. His most recent research activities

include evaluating content validity, test bias, differential item functioning, and the comparability of different language versions of tests and questionnaires. His vita can be accessed at <http://www-unix.oit.umass.edu/~sireci>.

The World Wide Web address for the *Education Policy Analysis Archives* is
epaa.asu.edu

Editor: Gene V Glass, Arizona State University

Production Assistant: Chris Murrell, Arizona State University

General questions about appropriateness of topics or particular articles may be addressed to the Editor, [Gene V Glass, glass@asu.edu](mailto:glass@asu.edu) or reach him at College of Education, Arizona State University, Tempe, AZ 85287-2411. The Commentary Editor is Casey D. Cobb: casey.cobb@unh.edu.

EPAA Editorial Board

[Michael W. Apple](#)
University of Wisconsin

[Greg Camilli](#)
Rutgers University

[Sherman Dorn](#)
University of South Florida

[Gustavo E. Fischman](#)
Arizona State University

[Thomas F. Green](#)
Syracuse University

[Craig B. Howley](#)
Appalachia Educational Laboratory

[Patricia Fey Jarvis](#)
Seattle, Washington

[Benjamin Levin](#)
University of Manitoba

[Les McLean](#)
University of Toronto

[Michele Moses](#)
Arizona State University

[Anthony G. Rud Jr.](#)
Purdue University

[Michael Scriven](#)
University of Auckland

[Robert E. Stake](#)
University of Illinois—UC

[David C. Berliner](#)
Arizona State University

[Linda Darling-Hammond](#)
Stanford University

[Mark E. Fetler](#)
California Commission on Teacher
Credentialing

[Richard Garlikov](#)
Birmingham, Alabama

[Aimee Howley](#)
Ohio University

[William Hunter](#)
University of Ontario Institute of
Technology

[Daniel Kallós](#)
Umeå University

[Thomas Mauhs-Pugh](#)
Green Mountain College

[Heinrich Mintrop](#)
University of California, Los Angeles

[Gary Orfield](#)
Harvard University

[Jay Paredes Scribner](#)
University of Missouri

[Lorrie A. Shepard](#)
University of Colorado, Boulder

[Kevin Welner](#)
University of Colorado, Boulder

EPAA Spanish & Portuguese Language Editorial Board

Associate Editors

Gustavo E. Fischman
Arizona State University
&

Pablo Gentili
Laboratório de Políticas Públicas
Universidade do Estado do Rio de Janeiro

Founding Associate Editor for Spanish Language (1998—2003)
Roberto Rodríguez Gómez
Universidad Nacional Autónoma de México

Argentina

- Alejandra Birgin
Ministerio de Educación, Argentina
Email: abirgin@me.gov.ar
- Mónica Pini
Universidad Nacional de San Martín, Argentina
Email: mopinos@hotmail.com,
- Mariano Narodowski
Universidad Torcuato Di Tella, Argentina
Email:
- Daniel Suarez
Laboratorio de Políticas Públicas-Universidad de Buenos Aires,
Argentina
Email: daniel@lpp-buenosaires.net
- Marcela Mollis (1998—2003)
Universidad de Buenos Aires

Brasil

- Gaudêncio Frigotto
Professor da Faculdade de Educação e do Programa de
Pós-Graduação em Educação da Universidade Federal Fluminense,
Brasil
Email: gfrigotto@globo.com
- Vanilda Paiva
Email: vppaiva@terra.com.br
- Lilian do Valle
Universidade Estadual do Rio de Janeiro, Brasil
Email: lvalle@infolink.com.br
- Romualdo Portella do Oliveira
Universidade de São Paulo, Brasil

Email: romualdo@usp.br

- Roberto Leher
Universidade Estadual do Rio de Janeiro, Brasil
Email: rleher@uol.com.br
- Dalila Andrade de Oliveira
Universidade Federal de Minas Gerais, Belo Horizonte, Brasil
Email: dalila@fae.ufmg.br
- Nilma Limo Gomes
Universidade Federal de Minas Gerais, Belo Horizonte
Email: nilmagomes@uol.com.br
- Iolanda de Oliveira
Faculdade de Educação da Universidade Federal Fluminense, Brasil
Email: iolanda.eustaquio@globo.com
- Walter Kohan
Universidade Estadual do Rio de Janeiro, Brasil
Email: walterko@uol.com.br
- [María Beatriz Luce](#) (1998—2003)
Universidad Federal de Rio Grande do Sul-UFRGS
- [Simon Schwartzman](#) (1998—2003)
American Institutes for Resesarch—Brazil

Canadá

- [Daniel Schugurensky](#)
Ontario Institute for Studies in Education, University of Toronto, Canada
Email: dschugurensky@oise.utoronto.ca

Chile

- Claudio Almonacid Avila
Universidad Metropolitana de Ciencias de la Educación, Chile
Email: caa@rdc.cl
- María Loreto Egaña
Programa Interdisciplinario de Investigación en Educación (PIIE), Chile
Email: legana@academia.cl

España

- José Gimeno Sacristán
Catedrático en el Departamento de Didáctica y Organización Escolar de la Universidad de Valencia, España
Email: Jose.Gimeno@uv.es
- Mariano Fernández Enguita
Catedrático de Sociología en la Universidad de Salamanca. España
Email: enguita@usal.es
- Miguel Pereira
Catedrático Universidad de Granada, España
Email: mpereyra@aulae.es
- [Jurjo Torres Santomé](#)
Universidad de A Coruña
Email: jurjo@udc.es
- Angel Ignacio Pérez Gómez
Universidad de Málaga

Email: aiperez@uma.es

- [J. Félix Angulo Rasco](#) (1998—2003)
Universidad de Cádiz
- [José Contreras Domingo](#) (1998—2003)
Universitat de Barcelona

México

- Hugo Aboites
Universidad Autónoma Metropolitana-Xochimilco, México
Email: aavh4435@cueyatl.uam.mx
- Susan Street
Centro de Investigaciones y Estudios Superiores en Antropología Social
Occidente, Guadalajara, México
Email: slsn@mail.udg.mx
- [Adrián Acosta](#)
Universidad de Guadalajara
Email: adrianacosta@compuserve.com
- [Teresa Bracho](#)
Centro de Investigación y Docencia Económica-CIDE
Email: bracho dis1.cide.mx
- [Alejandro Canales](#)
Universidad Nacional Autónoma de México
Email: canalesa@servidor.unam.mx
- [Rollin Kent](#)
Universidad Autónoma de Puebla. Puebla, México
Email: rkent@puebla.megared.net.mx
- Javier Mendoza Rojas (1998—2003)
Universidad Nacional Autónoma de México
- [Humberto Muñoz García](#) (1998—2003)
Universidad Nacional Autónoma de México

Perú

- Sigfredo Chiroque
Instituto de Pedagogía Popular, Perú
Email: pedagogia@chavin.rcp.net.pe
- Grover Pango
Coordinador General del Foro Latinoamericano de Políticas Educativas,
Perú
Email: grover-eduforo@terra.com.pe

Portugal

- Antonio Teodoro
Director da Licenciatura de Ciências da Educação e do Mestrado
Universidade Lusófona de Humanidades e Tecnologias, Lisboa,
Portugal
Email: a.teodoro@netvisao.pt

USA

- Pia Lindquist Wong

California State University, Sacramento, California

Email: wongp@csus.edu

- **Nelly P. Stromquist**
University of Southern California, Los Angeles, California
Email: nellystromquist@juno.com
- **Diana Rhoten**
Social Science Research Council, New York, New York
Email: rhoten@ssrc.org
- **Daniel C. Levy**
University at Albany, SUNY, Albany, New York
Email: Dlevy@uamail.albany.edu
- **Ursula Casanova**
Arizona State University, Tempe, Arizona
Email: casanova@asu.edu
- **Erwin Epstein**
Loyola University, Chicago, Illinois
Email: eepei@wpo.it.luc.edu
- **Carlos A. Torres**
University of California, Los Angeles
Email: torres@gseisucla.edu
- **Josué González (1998—2003)**
Arizona State University, Tempe, Arizona

EPAA is published by the Education Policy Studies
Laboratory, Arizona State University