Exploring the Contextual Limitations of Angoff Grading Model: The Case of Botswana

Molefhe Obonye Mogapiⁱ
Department of Educational Foundations
University of Botswana
mogapimo@mopipi.ub.bw

Abstract

Grading is a central component of educational assessment because it is through grading that a student's performance is judged as having satisfied the requirement for a given unit or grade level. Accurate determination of the cut-off point between adjacent grades is thus critical in ensuring that a letter grade assigned, and ultimately the certificate given, are a true reflection of a candidate's mastery of the assessed subject matter. Different systems are used to establish cut-off points; each method with its own advantages and disadvantages. In Botswana, the Angoff grading model was adopted following the introduction of criterion referenced testing in 1997. This model was implemented in order to correctly reflect the achievement level of primary school graduates and at the same time, maintain performance standards from year to year. While this model appears to have robust theoretical foundations, its practical application and success are highly constrained by contextual country-specific factors. This paper outlines the observed practical limitations of the Angoff grading system in Botswana.

Keywords: Angoff Grading System, Botswana, Performance Standards, Grading Cut-off Point.

Reference to this paper should be made as follows:

Mogapi, M. O. (2012). Exploring the Contextual Limitations of Angoff Grading Model: The Case of Botswana. *International Journal of Scientific Research in Education*, 5(4), 278-286. Retrieved [DATE] from http://www.ijsre.com

INTRODUCTION/BACKGROUND

The Angoff grading method was implemented following the introduction of the first primary school criterion referenced graduation examination in 1997. At the time, a team of primary school teachers was assembled and trained on how to use Angoff procedures. Each teacher (or judge) read each item and determined the number of minimally competent candidates within each grade band who would be able to answer the item correctly. The proportions produced by each teacher/judge are then averaged to generate cut-off scores for the grades: A, B, C and D.

Although Angoff procedures have a sound theoretical background supported by empirical evidence from a wide variety of fields (Ashby, 2001; Supernaw and Mehvar, 2002; Carlson, Tomkowiak & Stilp, 2009; Stahk, 2009), this paper contends that the accuracy, and as such the validity, of the Angoff system can seriously be constrained by contextual factors prevalent in a particular education system. In Botswana for example, the educational enterprise is in

a state of flux as is evidenced by numerous reforms recently implemented by the government of Botswana in conjunction with other education stakeholders.

The Primary Education Improvement Project [PEIP], which was jointly operated and financed by Botswana Government and United States Agency for International Development [USAID] from 1981 to November 1991, offers a prime example of this partnership strategy. The existence and longevity of this reform program is an indication of the fact that teaching and learning in Botswana schools is at a less than desirable level. As such, a grading model that relies on teachers, who themselves are in need of professional development, to generate cut-off points may results in less than accurate cut-off points. This paper thus contends that grading procedures which may have worked in other educational systems may not be automatically transferable to countries where key educational inputs are still at a formative developmental stage.

Two contextual challenges that stand to compromise the accuracy of Angoff in Botswana are the current status of the teaching profession and learner performance at the primary school level. The factors potential impacting the quality of education at the primary school level are numerous, complex, and interrelated. Two central factors affecting the quality of teaching and learning and consequently on the quality of education as a whole are the teaching force and learner performance. The teaching force at primary school level has been characterized as being of low quality due to low entry requirements (the minimum qualification for students interested in joining the profession at the primary school level has been a primary or junior school certificate) and short training periods:

The present admission requirements into primary teacher training fall into three levels. Firstly, a Junior Certificate fail or Standard 7 with at least two years teaching experience; secondly, a JC pass and one year work experience preferably in teaching; or thirdly, COSC with one year of work experience, preferably in teaching. Candidates holding JC pass account for about 85% to 90% of PTTC entrants while only 4% hold GCE. (Republic of Botswana, 1993, p. 345)

The second factor impacting the quality teaching is the pre-service programme offered at various Teacher Training Colleges. Prospective student teachers enrol in a two year training programme at the end of which they receive a Primary Teaching Certificate or PTC:

The two-year training period further produces an under-trained teacher as it does not afford enough time for the trainee to acquire the necessary knowledge and skills required to make an effective and confident teacher (Republic of Botswana, 1993, p. 345).

The very short programme does not give teacher trainees adequate time to acquire the academic knowledge and pedagogical skills needed to be effective teachers once they are in the classroom. Given that the teacher is the main facilitator in the classroom, the quality of learning and hence achievement levels are directly depended on the quality of the teacher. Though poor learner performance cannot be wholly attributed to quality of teaching, research studies have in fact found a substantial degree of correlation between the two variables (Republic of Botswana, 1997).

The consequence of the two factors discussed above has been the dominance of PTC teachers in primary schools; this phenomenon is reflected in Table 1 below.

Table 1: Standard Four Teachers with Primary Teaching Certificate and Secondary Teaching Certificate (STC) and Pup

| Subject | Primary Te | aching Certificate | Secondar | y Teaching Certificat | e |
|---------------------|--------------|--------------------|--------------|-----------------------|------|
| Mathematics | 5446 | | 2030 | | |
| Setswana English | 5461 5446 | | 2045 2039 | | |
| Lugusu | 3440 | | 2039 | | |
| Average | 5451 | 73% | 2038 | 27% | 7489 |
| | | | | | 100% |

Source: Botswana Examinations Council (2007).

Table 1 indicates that by 2007, the majority or 73% of standard four pupils were being taught by teachers with a Primary Teaching Certificate confirming that primary school education continues to be dominated by teachers with Primary Teaching Certificates.

Learner Performance Levels

Learner achievement is a key indicator of the quality of instructions in schools. Low performance levels among primary school learners in Botswana have been observed as far back as the 1970s.

The first National Commission on Education concluded form a study it administered on Standard Seven 7 Pupils in 1976 that average scores on achievement tests were low and a substantial proportion of children attained unacceptably low scores on their reading and mathematics test. In another survey conducted by the International Association for the Evaluation of Educational Achievement in 1990/91 to test the reading ability of 14 year olds in 32 countries, Botswana ranked last in all types of reading tests (story, textbook and document reading) (Report of the National Commission on Education, 1993).

In a study examining the learning achievement of standard four pupils in Botswana in 1999 by the Examinations Research and Testing Division, the low level of competence in numeracy, and literacy in English and Setswana, was taken to imply that the concern expressed in 1977 over low academic achievement persists today (Government of Botswana, 2000 p. 96).

Table 2: Performance of Standard Four Pupils

| Performance | of pupils | Number | Mean Percentage | Standard Deviation | Maximum | Minimum |
|-------------|-----------|--------|-----------------|--------------------|---------|---------|
| Mathematics | Overall | 6217 | 29.8 | 17.28 | 95 | 0 |
| | Girls | 2878 | 31.95 | 17.31 | 95 | 0 |
| | Boys | 3038 | 28.65 | 17.09 | 92 | 0 |
| Setswana | Overall | 6228 | 45.02 | 16.84 | 96 | 0 |
| | Girls | 2898 | 48.77 | 16.29 | 96 | 0 |
| | Boys | 3062 | 42.16 | 14.48 | 90 | 0 |
| English | Overall | 6171 | 33.53 | 15.58 | 88 | 0 |
| | Girls | 2837 | 36.39 | 15.81 | 88 | 0 |
| | Boys | 2994 | 31.5 | 14.99 | 84 | 0 |

Source: Botswana Examinations Council (2007)

Generally pupils performed poorly in all the subjects with the mean performance being less than 50%. The mean performance for Setswana was 45.02%, for Mathematics was 29.80% and for English was 33.53%. These low means suggest that the pupils found the three tests to be quite difficult. Other studies conducted in Botswana for different grade levels, such as Monitoring of Learning Achievement (MLA, 2001) for Standard Four pupils, Trends in International Mathematics and Science Study (TIMSS, 2003) for Form One students, and Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ II, 2005) for Standard Six pupils also revealed low performance by students in Botswana.

The contextual challenges discussed above have been a concern to the government of the day since the country attained independence in 1966. As a result successive governments have developed policies and programs to try and address the low quality teaching force and poor learner performance. The minimum required qualification for entry into Primary Teacher Training College, for example, was raised to the Cambridge Overseas Schools Certificate [COSC] and the length of the training programme was increased from two years to three years. Diagnostic and remediation assessment were introduced at the primary school level to try to improve classroom assessment practices and hence learner performance. These factors are, however, still a concern today and cannot be ignored during any new standard setting exercises.

Angoff standard setting as practiced in Botswana has relied heavily on primary teachers to set grade cut-off points. The accuracy and validity of such cut-off points is thus open to scrutiny based on the observed constraints related to low quality teaching and poor performance by learners. In this system, individual teachers or judges will set a passing score based on their subject matter knowledge and the status of the learner being graded. If both of these two conditions are at an undesirable level, the standard set will be low since no judge can be expected to set a standard that is over and above what they know and can defend.

LITERATURE REVIEW

According to Stahl (2008), the central issue in any standard setting procedure is generating a defensible categorization of exams on the basis of the subject matter tested as reflected by a set of test items. Firstly, an appropriate measurement scale has to be identified and applied to differentiate exams according to the underlying construct being tested. Cut-off points along the scale then define each category or grade as explained below.

The second condition is placement of a point or points on this scale of measurement using a psychometrically sound procedure. These points are frequently referred to as cut-off points. The points demarcate regions on the scale of measurement that are deemed to be different in terms of the purpose of the test. A point may separate a region of pass from a region of fail. Multiple points may separate regions of insufficient mastery from acceptable performance and separate regions of acceptable performance from mastery. An examinee's performance on a test will place them in one of these regions (Stahl, 2008).

An assessment of literature on standard setting shows that there are many standard setting procedures available but in most cases two methods are generally applied. Standards are generally classed as absolute (criterion-based) or relative (norm-based). An absolute standard determines the pass/fail outcome in terms of how well a candidate performs and he/she is usually judged against an arbitrarily set external standard. Hence it is independent of the performance of the group. A relative standard on the other hand, compares how well the examinee has performed compared to others who took the test and hence the outcome (pass/fail) is dependent on the performance of the group (George, Haque, & Oyebode, 2006). Angoff falls under absolute grading standards as the performance of the candidate is as much as possible independent from or not necessarily influenced by performance of the norm group.

According to Carlson and Tomkowiak (2009), the Angoff standard has three basic elements: conceptualization of the borderline examinees, identification of specific test items, and the usage of experts to estimate whether a borderline examinee will appropriately perform each of the test items. Conceptualization of the borderline examinee or the minimally competent examinee is central as their performance is ultimately used as a cut-off point between two adjacent grades. The minimally competent examinee score is used in this case to identify the lowest possible or acceptable level of performance for any candidate to be classified under that particular grade or labelled as pass or fail in case of dichotomous classification. The score obtained by a minimally competent candidate then becomes the grade boundary between adjacent grades. Candidates scoring below this level are believed to lack sufficient knowledge, skills, or abilities to be certified. Subject-matter experts (SMEs) have to read each test item and then predict how many minimally-qualified candidates would answer the item correctly. The sum of the predicted difficulty values for each item averaged across the judges and items on a test is the recommended Angoff cut score (ALTA Language Services, 2008). Table 3 below illustrates how the procedure works for a single grade. As shown in the table, the minimum score generated is 99.45 out of a total of 150 items in the test. The panel will then have to repeat the procedure in order to arrive at a cut of score for other grades.

Table 3: Angoff Estimated Probabilities

| | | Ang | Across Judges | | | | |
|------------|-----------|-----|---------------|-----|-----|---------|-----------------------|
| | 1 2 3 4 5 | | | | | Average | Standard Deviation |
| Question 1 | .75 | .75 | .80 | .65 | .70 | 0.73 | .057 |
| Question 2 | .65 | .70 | .75 | .65 | .80 | 0.71 | .065 |
| Question 3 | .70 | .65 | .60 | .65 | .65 | 0.65 | .035 |

| Question 4 | .65 | .75 | .65 | .70 | .60 | 0.67 | .057 |
|---------------|--------|--------|--------|-----|-----|-------|-------|
| Question 5 | .55 | .50 | .45 | .65 | .55 | 0.54 | .074 |
| | | | | | | | |
| Question 146 | .80 | .80 | .80 | .70 | .60 | 0.74 | .089 |
| Question 147 | .80 | .75 | .70 | .55 | .65 | 0.69 | .096 |
| Question 148 | .55 | .60 | .65 | .65 | .45 | 0.58 | .084 |
| Question 149 | .65 | .65 | .70 | .75 | .65 | 0.68 | .045 |
| Question 150 | .65 | .70 | .65 | .65 | .55 | 0.64 | .055 |
| Passing score | 101.25 | 102.75 | 101.25 | 99 | 93 | 99.45 | 3.846 |

Retrieved from; http://chrp.com/angoff-panel/

The Angoff grading procedures enjoy wide application cross many diverse fields and the method has proven to be a reliable and credible one for setting a cut-off score for exams. An excerpt from Näsström and Nyström (2008) confirms this observation. According to these authors, the Angoff method is chosen to represent the test-centered methods because in its original version, or in a modified and extended version, it is the most widely used procedure for standard-setting. Furthermore, a modified version of the Angoff method is used regularly as the standard setting procedure for the national tests in mathematics in Sweden (Näsström & Nyström, 2008).

The Angoff procedure has been known to produce meaningful cut-off points especially in cases where there is high inter-judge consistency which produces considerable confidence in the resulting cut-scores (Hambleton & Pitoniak, in Näsström & Nyström, 2008). Cut-off scores are generated systematically as using the Angoff method ensures that the passing grade of a test is determined empirically. Another cited advantage of the Angoff method is that it is fairly easy to employ because it does not require judges to directly observe every student's performance, a process that is very time consuming (Carlson & Tomkowiak, 2009).

Several researchers have, however, documented weaknesses associated with the Angoff standard setting method. The first noted weakness is that 'judges occasionally report feeling that there is no firm basis for the standard that is set, since they are predicting performance as opposed to directly observing examinee performance' (Carlson& Tomkowiak, 2009). Secondly, it has been observed that the Angoff method relies on subjective estimations of the ability of a minimally competent candidate or MCC. SMEs base their decisions on the mental image they form of this candidate and this is not an easy task as explained below.

The way that SMEs are asked to use this estimation is problematic. By definition, SMEs are familiar with the content tested on an examination. By the process described above, they refine their concept of the candidates' abilities. The problem arises when they are asked to combine their content expertise and their conceptualization of the ability of the MCC in ways that are outside their expertise (Stahl, 2008).

This observation resonates with Boursicot and Roberts's comments on the viability of Angoff. These two scholars refer to the idea of a borderline or minimally competent candidate as a 'nebulous concept' and research has shown that often judges find it difficult to accurately define and understand a hypothetical borderline student (George, Haque, & Ovelode, 2006)

The third weakness of this approach is the tendency of judges to use Angoff to produce low cut off scores when compared with other grading methods. This effectively means more candidates are able to access higher grades than would have been the case if a different method was used. For the students participating in the exam generating data for this study, the Angoff method yielded a 100% passing rate (percentage of students who scored above the passing cut-off) compared to lower passing rates for both other standards set. This may imply that the cut-off scores produced by the Angoff method are in fact too lenient (Carlson & Tomkowiak, 2009). Table 4 provides more evidence to support the argument that the Angoff grading method frequently generates lenient cut off points when compared with other methods.

Table 4: Comparison of Angoff and Absolute Grading Procedures

| | Mean Score | Standard Deviation | Angoff | | Arbitrary | – Absolute | Norm | |
|---------------------|---------------|-----------------------|---------|----------|-----------|------------|---------|------|
| | | | Cut-off | Pass | Cut-off | Pass | Cut-off | Pass |
| Pre – Op | 86% | 9.0 | 63 | 100% | 70 | 89 | 77 | 82 |
| Shortness of Breath | 73% | 11.8 | 70 | 69% | 70 | 69 | 61 | 86 |
| Acute Abdomen | 86% | 8.8 | 62 | 100 % | 70 | 93 | 77 | 77 |
| Lower Back Pain | 72% | 13.0 | 52 | 91% | 70 | 58 | 59 | 79 |
| Full Exam | 79% | 7.1 | 62 | 100% | 70- | 88 | 72 | 81 |

Source: Carlson, J. & Tomkowiak, J. (2009).

The Angoff method has wide application and is able to produce good standards. However, the procedure has serious limitations that cannot be ignored. Its theoretical foundations – requiring a panel of judges to form a mental picture of a minimally competent candidate – are hard to prove empirically since different judges would undoubtedly generate different images. Secondly, validity and reliability of the method depends to a large extend on the SME's characteristics such that different groups of judges could produce quite different standards. For example, a panel made up of teachers may produce different cut-off scores from a panel made up of curriculum officers. Care should thus be taken that the standards generated do not depend on the kind of panel selected.

Implementation of Angoff Grading Method in Botswana

Table 5 below shows how the Angoff standard setting method has been used in Botswana to generate cut-off points for five subjects: Setswana, English, Science, Social Studies and Mathematics.

Table 5: Angoff Generated Cut Off Points for five Subjects

| | Holistic Grading Method | | | | Angoff Grading Method | | | | |
|----------|-------------------------|------|------|------|-----------------------|------|------|------|------|
| | | A | В | C | D | A | В | C | D |
| Setswana | | 91 | 64 | 36 | 18 | 81 | 68 | 57 | 41 |
| | | 82 | 57 | 28 | 12 | 72 | 61 | 52 | 35 |
| | Dimensions | 80 | 70 | 40 | 20 | 77 | 68 | 57 | 40 |
| | | 84.3 | 63.7 | 34.7 | 16.7 | 76.7 | 65.7 | 55.3 | 38.7 |
| English | | 74 | 63 | 45 | 32 | 84 | 68 | 56 | 37 |
| | Dimensions | 80 | 64 | 48 | 28 | 80 | 69 | 55 | 33 |
| | | 77 | 63.5 | 46.5 | 30 | 82 | 68.5 | 55.5 | 35 |
| Science | | 80 | 65 | 40 | 15 | 77 | 63 | 45 | 25 |
| | Dimensions | 75 | 60 | 25 | 15 | 82 | 65 | 53 | 35 |
| | | 85 | 70 | 35 | 15 | 80 | 64 | 48 | 32 |
| | | 80 | 65 | 33.3 | 15 | 79.6 | 64 | 48.6 | 30.6 |

| Social | | 80 | 65 | 50 | 35 | 74 | 66 | 54 | 29 |
|-------------|------------|----|------|------|------|----|------|------|------|
| Studies | Dimensions | 85 | 68 | 50 | 37 | 76 | 63 | 51 | 28 |
| | | 87 | 60 | 50 | 35 | 81 | 70 | 58 | 32 |
| | | 84 | 64.3 | 50 | 35.6 | 77 | 66.3 | 54.3 | 26.6 |
| Mathematics | | 82 | 66 | 48 | 28 | 91 | 81 | 63 | 31 |
| | Dimensions | 81 | 67 | 53 | 29 | 91 | 79 | 58 | 28 |
| | | 82 | 67 | 50.5 | 29 | 91 | 80 | 60.5 | 29.5 |

Source: Examinations Research and Testing Grading Document for 1997.

A panel of judges made up of primary school teachers was assembled for each of the five subjects and asked to read each item in the test so as to determine its difficulty level. Each judge then worked independently to generate two separate cut-off points using the Holistic and then Angoff procedure. The scores for each judge were then averaged to arrive at a composite cut-off point for each grade. As shown in the table above, the cut-off points for Grade A that were generated using the Angoff method are generally lower than the corresponding cut-off points arrived at using the Holistic method. The cut-off for an A grade in Setswana Dimension B, for example, is 91% based on the Holistic method and 81% based on the Angoff method. This means that more examinees would access an A grade under Angoff. The situation is reversed at the lowest point on the scale. The minimally competent D student would need to score 18% in the Holistic method but 41% in the Angoff method, a difference of 23 points. A further inconsistency produced by the Angoff method can be observed in the grading of mathematics. Although Angoff cut-off points are higher than Holistic ones, the two procedures seem to have comparable cut-off points for a D grade. With the Angoff method, however, there is a difference of 32 points between a C and D grade as shown by the 31% cut-off for D and 63% for C. This wide gap may point to problems associated with the conceptualisation of a minimally competent candidate

It is appropriate at this juncture to further explain or summarize the factors in Botswana that may limit the effectiveness of the Angoff standard setting procedure.

- a) Conceptualisation of a Minimally Competent Examinee [MCE]: the panel of judges often finds it difficult to form a mental picture of a MCE. Different judges have different conceptualisations resulting in low inter-judge agreement. Alternately, judges might be operating on the basis of an average student within each grade band thus generating a midpoint score rather than a cut-off point. The situation above where cut-off points for grades C and D were set at 61% and 31% respectively serves as a good example. The 61% for C appears to be a midpoint for a C grade rather than a score for a minimally competent C examinee.
- b) Subject Matter Experts [SMEs]: the generation of accurate, and as a result, valid cut-off points depends entirely on the characteristics of the judges or SMEs in the panel. It is not unusual for two different panels to produce different standards for the same examination paper. As noted above, the panel of judges used in Botswana is made up of entirely of primary school teachers. The accuracy of the standards generated would then be influenced by the current subject matter expertise possessed by the judges. Given the low entry levels for teacher trainees and inadequate training programmes, it seems reasonable to expect that judges will set standards that correspond to their current knowledge level. This effectively means setting lower cut off points.
- c) Learner Performance at School Level: corollary to the point above, the judges themselves are not only constrained by their knowledge status, but are also aware of students' low performance levels. The final decision of a judge is thus influenced by an interplay of these two factors: lack of requisite subject matter expertise as a result of recruitment practices and pre-service training on one hand, and low performance levels of primary school learners on the other hand. These two factors will lead judges to lower the cut off point in order to accommodate the learning and teaching limitations experienced in the field.
- d) Content Domain Assessed: a panel of judges that is entirely made up of teachers leads to the enhancement of the effective curriculum at the expense of the official curriculum. The difference between the two is important. The curriculum that is actually implemented by teachers in schools is called the effective curriculum. The effective curriculum consists of those topics and learning objectives that teachers actually teach to students. The effective

curriculum stands in contrast to the desired curriculum. The official or desired curriculum consists of those topics and learning objectives which are found in the national curriculum and which government desires to be taught. The effective and desired curriculums are often overlapping, but are also different. Obviously, when a national curriculum exists and when government invests in educational inputs, it is desirable to have these two curricula be congruent.

e) Test Dimensionality: the introduction of Criterion Referenced Testing following the Revised National Policy on Education recommendations in 1994 encouraged assessment of a broad range of skills. Every examination paper developed was based on a blue print that ensured test items targeted both low order and high order cognitive skills. The Angoff method, however, does not require SMEs to consider the dimensionality of the test when determining cut off points; this apparent oversight has been noted by Ricker (2003). An examinee can meet a performance standard set using an Angoff method either by being minimally competent on all dimensions or areas of a test, or by making up for deficiencies on a given dimension with strengths in other dimensions. Using the Angoff method, judges only decide the probability that a minimally acceptable candidate will answer an item correctly, but they cannot determine that a student must answer a question correctly in order to be considered competent. (Ricker, 2003)

The only way to ensure that national aspirations on basic primary education are achieved is to assess learners on the basis of the official curriculum. This can be done by setting pre-determined cut-off scores that reflect desired national expectations as embodied in reform policies and programmes.

CONCLUSION

The Angoff standard setting model has very solid theoretical background and the model is used in many fields. However, the reliability and validity of standards set using Angoff depend to a large extent on the prevailing contextual conditions specific to a country. In Botswana, the educational sector is still at a formative stage as exemplified by numerous government sponsored reforms. Any grading system adopted should enable generation of standards that captured the spirit of those reforms. This translates into setting high standards that would help in the identification of strengths and weaknesses in the learning and teaching processes. A grading system, such as Angoff, that produces lenient cut off points only helps in maintaining the status quo. In the long run, the aims and objectives of the reforms will not be achieved. The observed limitations of commonly used grading systems such as Angoff should serve as a reminder to education researchers and measurement experts to constantly review, refine, modify and develop procedures that are relevant and have the potential to improve the quality of teaching and learning processes.

REFERENCES

- ALTA Language Services (2008). What is the Angoff Grading Method? Retrieved from: http://www.altalang.com/beyond-words/2008/10/06/what-is-angiff-method/
- Ashsly, David. (2001). The CFP certification examinations process: A discussion of the modified Angoff scoring method. Financial Services Review. 10 92001) 187-195. Retrieved from http://www.stetson.edu/fsr/abstracts/vol_10_num_p187_.pdf
- Botswana Examinations Council (2007). Standard Four Assessment Report. Retrieved from; http://chrp.com/angoff-panel/
- Carlson, J., Tomkowiak, J., & Stilp, Curt. (2009). Using the Angoff Method to Set Defensible Cut-off Scores for Standardized Patient Performance Evaluations in PA Education. The Journal of Physician Assistant Education | 2009 Vol 20 No 1 15. Retrieved from; http://www.paexcelence.org/
- George, S., Haque, S. M., & Oyebode, F. (2006). Standard setting: Comparison of two methods. Department of Psychiatry, University of Birmingham, UK. Retrieved from: http://www.altalang.com/beyond-words/2008/10/06/what-is-angiff-method/
- Näsström, G., & Nyström, P. (2008). A comparison of two diffrent methods for setting performance standards for test with constructed response items. Retieved from: http://www.pareonline.net/pdf/v13n9.pdf
- Republic of Botswana (1993). Report on the National Commission on Education. Gaborone: Government Printers. Republic of Botswana (1997). Education Statistics Report. Gaborone: Central Statistics Office. Republic of Botswana.

- Ministry of Education. (2005). Trends in the international mathematics and science study TIMMS. Examinations Research and Testing Division. Gaborone: Government Printer
- Republic of Botswana. Ministry of Education (2000). Report on the monitoring learning achievement project MLA. Examinations Research and Testing Division. Gaborone: Government Printer.
- Republic of Botswana (2005). The SACMEQ II Project in Botswana: A study of the conditions of schools and the quality of education. Ministry of Education. Retrieved from: http://www.sacmec.org.reports.htm
- Ricker, K. L. (2003). Setting Cut Scores: Critical Review of Angoff and Modified-Angoff Methods, Centre for Research in Applied Measurement and Evaluation. University of Alberta. Retrieved from: http://www2.education.ualberta.ca/educ/psych/crame/psych/files/
- Stahl, J., & Pearson, V. (2009). Using the Angoff Method to Set Defensible Cut-off Scores for Standardized Patient Performance Evaluations. The Journal of Physician Assistant Education | 2009 Vole 20 No 1 23. Retrieved from http://www.iaea2008.cambridgeassessment.org.uk/ca/digital/Assets/154191 stahl.pdf
- Supernaw, R., & Mehvar, R.(2009). Methodology for assessment of competenies and the definition of deficiecies of students an all levels of curiculum. Retrieved from: http://www.archive.ajpe.org/legacy/pdfs/AJ660101.pdf

.....

ⁱ Molefhe Obonye Mogapi is a Lecturer in the Department of Educational Foundations of the University of Botswana. Mogapi can be reached on mogapimo@mopipi.ub.bw .