A COMPARISON OF THE ANALYTICAL HIERARCHY PROCESS AND SUBJECT PAIRS ANALYSIS WHEN RANKING SUBJECT DIFFICULTY

Pauline Masemola
Umalusi Council
Pauline.Masemola@umalusi.org.za

Nthabeleng Lepota
Umalusi Council
Nthabeleng.Lepota@umalusi.org.za

Abstract

This paper describes the application of the Analytical Hierarchy Process (AHP) and Subject Pairs Analysis (SPA) for evaluating subjects' difficulty. The study assesses the difficulty of National Senior Certificate (NSC) subjects for the Grade 12 exit examination and aimed at providing information on the actual performance of learners in different subjects. The study further intended to determine subjects which were consistently more difficult relative to other subjects. The study used data provided by the Department of Basic Education (DBE) on learner performance in the Grade 12 NSC examinations during the 2014 to 2018 academic years. The study was quantitative in nature and used R and SAS software for the statistical analysis. Two techniques, namely AHP and SPA were compared in the analysis of data to identify subjects that were consistently difficult across the years.

Keywords: *Subject, National Senior Certificate, Difficulty, Performance, Pairs, Quantitative*

Introduction and Background

Research often reports on the perceived difficulty of subjects such as Mathematics and Physical Sciences (Anyagh et al, 2018). The perception about the difficulty of these subjects has led to the decline in the interest of young people in pursuing Mathematics and Science related careers (Osborne et al, 2010). Studies such as Mji and Makgato (2003), determined the effect of performance of learners in Mathematics and Physical Science by looking into the teaching strategies, content knowledge, motivation, laboratory use, and non-completion of the syllabus in a year. Many factors might influence the performance of learners in Mathematics and Physical Sciences. However, in this paper the objective is to use statistical methods to rank subjects according to their difficulty. Coe et al, (2008) highlighted several statistical methods that could be used to rank subjects which include the Subject Pair Analysis, Latent trait models, Reference tests and value-added models. However, in this study only the Subject Pair Analysis (SPA) and the Analytical Hierarch Process (AHP) are used. Currently, Umalusi[1] uses the AHP method to rank subjects according to difficulty. Such information is shared with the Assessment Standard Committee (ASC) of Umalusi to assist in making evidence-based decisions relating to the adjustments of mark distributions during the standardization (grading) process.

Umalusi is, amongst others, mandated to quality assure exit point assessments for qualifications in schools, TVET Colleges and for Adults (GETC). The quality assurance processes include the moderation of question papers; verification of School-Based Assessment (SBA); monitoring of the conduct of examinations; the verification of marking; and the standardization of marks. This quality assurance process is done to obtain equivalence of the standard of qualifications by using a norm-referencing method.

[1] Umalusi is a council that sets and monitors standards for general and further education and training in South Africa in accordance with the General and Further Education and Training Quality Assurance Act of 2008, as amended

A norm, also referred to as the historic average is calculated by using the average scores of the previous 3 – 5 years' examinations sittings. During standardisation, the learner marks are adjusted marginally towards the historical average/norm. The objective of the process of adjusting the learner distribution towards the historical average/norm is to ensure that standards are comparable across different years ensuring that there is consistency in mark distribution from one year to the next. This process is followed to ensure that a cohort of learners is not advantaged or disadvantaged by extraneous factors other than their knowledge of the subject, abilities and their aptitude.

During the standardisation process, assessment bodies are given an opportunity to propose their adjustment for subjects and to highlight intervention strategies that have been implemented and that might have affected learner performance during the year. Also, question paper moderators presents reports on content coverage and cognitive levels of the examination papers. This information is shared to assist in providing additional information about factors that might have influenced learner performance other than the learners' subject knowledge, abilities and aptitude. The challenge with this information is that it does not provide information about the actual learner performance, but highlights factors that might have influenced learner performance. As a result, information on AHP, which ranks subjects according to difficulty, is needed. The AHP focuses on the actual performance of learners in

different subjects and as a result one is able to judge the difficulty of a subject compared to other subjects.

Literature review

There are several statistical methods used for ranking subject difficulty such as Latent trait models, Subject pair analysis, Reference tests and value-added models. In this study not all these methods could not be explored. Only the Subject Pair Analysis and Analytical Hierarchy Process methods were used to compare subject difficulty. The reason for the application of the two methods is because the data used is not stored on item level, therefore none of the other methods could be applied. This study was conceptualized to compare the outcomes of the two methods when subject difficulty is compared to determine which subjects are consistently easy or difficult across years.

As previously indicated, the technique called the Analytical Hierarchy Process (AHP) and Subject Pair Analysis (SPA) were used for the analysis of the data. The AHP is a tool for dealing with complex decision-making, and may aid the decision maker to set priorities and make the best decision (Saaty and Katz, 1990). By reducing complex decisions to a series of pairwise comparisons, and then synthesizing the results, the AHP helps to capture both subjective and objective aspects of a decision. In addition, the AHP incorporates a useful technique for checking the consistency of the decision maker's evaluations, thus reducing the bias in the decision making process. The ratio scales are derived from the principal Eigen vectors and the consistency index is derived from the principal Eigen value. This process applies a pair-wise comparison. Pairwise comparison generally is any process of comparing entities in pairs to judge which of each entity is preferred, or has a greater amount of some quantitative property, or whether or not the two entities are identical.

The SPA method considers candidates who have taken a pair of subjects. That is a candidate must have done both subjects being compared to see in which subject they have performed better. The difference between the mean grades achieved by the same candidates in each subject are calculated (Coe, 2007). If they typically achieve better grades in one over the other, then one subject is easier than the other. A more widely used variation is to compute an average difference in the grades achieved in the two subjects. These methods could be described as 'interval' approaches since any such average will be sensitive to the sizes of the gaps between grades, not just to their order. The conventional way to do this is to convert examination levels or grades into a numerical scale using consecutive integer values (e.g. at, 0 – 29 = 1; 30 – 39 = 2; 40 – 49 = 3; 50 – 59 = 4; 60 – 69= 5; 70 – 79 = 6; 80 – 100 =7). For each candidate who took a particular pair of subjects, the difference between their grade in Subject 1 and Subject 2 is calculated. The mean of these differences across all candidates is a measure of the difference in difficulty of the pair, in grade units (Coe et al, 2008).

Purpose of the study

The purpose of this study is to use Analytical Hierarchy Process (AHP) and Subject Pairs Analysis (SPA) for evaluating subjects' difficulties. This is done to provide information on the actual performance of learners in different subjects. Such information assists the Assessment Standards Committee, a committee that is responsible for standardizing examination results after examinations have been written, in making a decision on the kind of adjustment to be applied on a particular subject based on the performance of learners.

Material and methods

This study was quantitative in nature and used secondary data sourced from the Department of Basic Education. Data management was accomplished by the use of SAS, R and Excel.

Samples and variables

The sample was made up of learners who sat for the following six subjects: Mathematics, Physical Sciences, Life Sciences, Geography, History and English First Additional Language written for the 2014 to 2018 Grade 12, National Senior Certificate (NSC) examinations. The reason for considering only learners who have written all six subjects was that the two methods used for ranking subject difficulty uses the notion of common learners in subjects. The dataset contained learner records on the performance of learners in different subjects in Grade 12 from public schools in South Africa.

Research design and methods

Research methods are divided into three categories namely: qualitative methods, wherein the inquirer often makes knowledge claims based primarily on constructivist perspectives; quantitative methods, those that the investigator primarily uses post-positivist claims for developing knowledge; and mixed methods, those that the researcher tends to base knowledge claims on pragmatic grounds (Creswell, 2003).

This study adopted the use of quantitative methods because statistical methods were used in the analysis (Muijs, 2011). Quantitative methods emphasise objective measurements and the statistical, mathematical, or numerical analysis of data collected through polls, questionnaires, and surveys, or by manipulating pre-existing statistical data using computational techniques (Tolmie et al., 2011). This study followed an ex-post-facto design. That is, the researcher did not affect the outcome, but only observed the data without changing it (Sullivan and Krieger, 2001).

Results and Interpretation

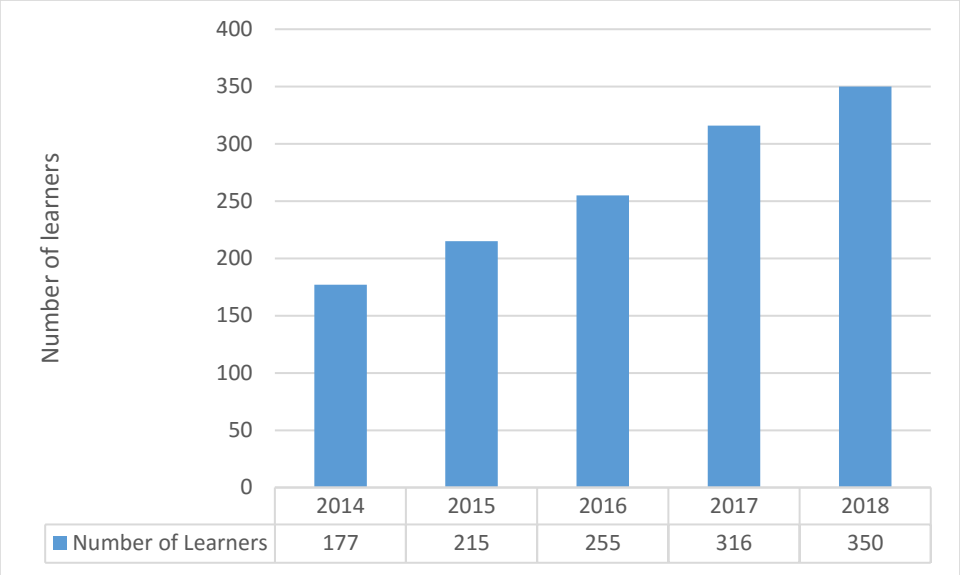
The first part of the results describes the number of learners who took a combination of the six subjects of interest. The second part highlights the outcomes of the application of the AHP and the SPA methods.

Descriptive statistics

In this study, the descriptive and inferential statistics were considered. Descriptive statistics consists of procedures used to summarise and describe important characteristics of a set of observations (Stamatis, 2012). Inferential statistics is made up of procedures that are used to make an inference about the population characteristics observed from a sample (Mendenhall et al., 2013).

Graph 1 shows the number of common candidates who wrote the six subjects ranked in terms of difficulty. The number of candidates enrolled for the combination of subjects is increasing over time. In 2014, there were 177 learners which increased to 350 in 2018.

Graph 1: Number of candidates across used in the study (2014-2018)



Inferential statistics

This section presents the results of the Subject Pair Analysis and Analytical Hierarchy Process. The results from the two methods are compared. In particular, the difficulty levels of the subjects are ranked. In both methods, six subjects namely: English First Additional Language, Geography, History, Life Sciences, Mathematics and Physical Sciences were ranked in terms of difficulty. The two methods are characterised by common examinees for the ranked subjects. That is, this methods depend on the comparisons among the results achieved by the same candidates in different subjects.

Table 3.1 presents the average grades of candidates who wrote six subjects in the year 2014. For the ranking of these subjects, all candidates who have taken a combination of these subjects were considered. The difference between mean grades achieved by the same candidates in each subject is calculated. If the candidates achieve better grades in one subject above another, then one subject is ranked as easier than the other. Table 3.1 shows the rankings of subjects achieved in each grade. The marks for the NSC are ordinal and range from zero to hundred. However, there are levels of achievement associated with these categories represented by values. The average grades represents the ranking of the subjects in terms of their difficulties. The higher the average grade, the easier the subject is compared to others. However, if the average grade has a lower ranking then the subject is interpreted as more difficult compared to others. In the year 2014, History and English First Additional Language were the easiest subjects. However, Mathematics and Physical Sciences were difficult.

Table 3.1: 2014 Subject Pairs analysis

Category	Value	English FAL	Geography	History	Life Science	Mathematics	Physical Science
0-29	**1**	0.01	0.07	0.02	0.19	0.67	0.60
30-39	**2**	0.10	0.38	0.19	0.51	0.31	0.35
40-49	**3**	0.64	0.73	0.54	0.49	0.17	0.34
50-59	**4**	1.27	1.20	0.99	0.70	0.25	0.18
60-69	**5**	1.50	0.73	1.30	0.42	0.20	0.20
70-79	**6**	0.58	0.17	0.88	0.58	0.00	0.10
80-100	**7**	0.12	0.12	0.32	0.24	0.12	0.04

Average Grade	**4.21**	**3.41**	**4.25**	**3.13**	**1.71**	**1.81**

Table 3.2 gives the subject pair analysis for 2015. A trend was observed that the same subjects that were easier and difficult respectively in 2014 maintained their position in the ranking of subject difficulty in 2015. The average grades for Mathematics and Physical Sciences were 1.60 and 1.71, respectively. Whereas, for History and English FAL the average grades were 4.25 and 4.21, respectively.

Table 3.2: 2015 Subject Pairs analysis

Category	Value	English FAL	Geography	History	Life Science	Mathematics	Physical Science
0-29	**1**	0.00	0.18	0.08	0.17	0.68	0.60
30-39	**2**	0.13	0.38	0.28	0.48	0.32	0.40
40-49	**3**	0.81	1.14	0.74	0.81	0.25	0.39
50-59	**4**	1.36	0.61	1.06	0.69	0.17	0.13
60-69	**5**	1.07	0.37	0.88	0.49	0.09	0.09
70-79	**6**	0.64	0.11	0.39	0.25	0.06	0.06
80-100	**7**	0.03	0.03	0.20	0.03	0.03	0.03
Average Grade		**4.04**	**2.83**	**3.63**	**2.93**	**1.60**	**1.71**

Table 3.3 presents the subject pairs analysis for candidates who wrote Grade 12 in 2016. Of the six subjects ranked, History and English FAL were the easiest subjects with the highest rankings of 4.31 and 4.15, respectively. However, Mathematics and Physical Sciences were difficult with the lowest rankings of 1.70 and 1.88, respectively.

Table 3.3: 2016 Subject Pairs analysis

Category	Value	English FAL	Geography	History	Life Science	Mathematics	Physical Science
0-29	**1**	0.00	0.23	0.05	0.20	0.68	0.58
30-39	**2**	0.08	0.60	0.16	0.56	0.26	0.38
40-49	**3**	0.65	0.78	0.56	0.69	0.28	0.29
50-59	**4**	1.62	0.56	0.81	0.56	0.19	0.24
60-69	**5**	1.17	0.30	1.13	0.55	0.11	0.24
70-79	**6**	0.51	0.07	1.12	0.19	0.14	0.12
80-100	**7**	0.12	0.04	0.49	0.02	0.04	0.04
Average Grade		**4.15**	**2.57**	**4.31**	**2.78**	**1.70**	**1.88**

Table 3.4 indicates the ranking of the six subject pairs for 2017. The same trend is observed for subjects that rank easy or difficult. In 2017, the average grades were 4.66 and 4.40 for History and English FAL, respectively. Whereas, Physical Sciences was the most difficult subject with a ranking of 2.00. Physical Sciences was followed by Mathematics with a ranking of 2.03.

Table 3.4: 2017 Subject Pairs analysis

Category	Value	English FAL	Geography	History	Life Science	Mathematics	Physical Science
0-29	**1**	0.01	0.23	0.03	0.19	0.52	0.49
30-39	**2**	0.07	0.59	0.10	0.42	0.40	0.48
40-49	**3**	0.47	0.91	0.38	0.63	0.37	0.42
50-59	**4**	1.28	0.53	0.73	0.94	0.38	0.35
60-69	**5**	1.66	0.14	1.52	0.54	0.21	0.11
70-79	**6**	0.80	0.06	1.29	0.25	0.11	0.06
80-100	**7**	0.11	0.00	0.60	0.09	0.04	0.09
Average Grade		**4.40**	**2.46**	**4.66**	**3.04**	**2.03**	**2.00**

In 2018, the comparisons of the subject difficulty revealed that History followed by English FAL were the easiest subjects with the highest rankings of 4.45 and 4.12, respectively. Mathematics was the most difficult subject with the ranking of 1.95. This was followed by Physical Sciences with a ranking of 2.14. Comparing Mathematics and Physical Sciences, Mathematics appeared to be more difficult with the difference of 0.19 points. Table 3.5 makes this information more explicit.

Table 3.5: 2018 Subject Pairs analysis

Category	Value	English FAL	Geography	History	Life Science	Mathematics	Physical Science
0-29	**1**	0.00	0.23	0.02	0.21	0.54	0.50
30-39	**2**	0.13	0.63	0.12	0.57	0.41	0.41
40-49	**3**	0.63	0.68	0.52	0.61	0.31	0.33
50-59	**4**	1.51	0.61	0.97	0.65	0.37	0.32
60-69	**5**	1.20	0.30	1.31	0.44	0.23	0.33
70-79	**6**	0.53	0.10	1.20	0.27	0.09	0.22
80-100	**7**	0.12	0.00	0.30	0.04	0.02	0.04
Average Grade		**4.12**	**2.55**	**4.45**	**2.80**	**1.95**	**2.14**

Table 3.6 summarises the results of the Analytical Hierarchy Process analysis for the six subjects across five years. In 2014 and 2018, History was the easiest subject followed by English FAL. However, in 2015 to 2017, English FAL was easier than History. Mathematics followed by Physical Sciences were the most difficult subjects across the five years. When using the Subject Pairs Analysis method for the 2017 cohort, Physical Sciences was more difficult than Mathematics. However, using the AHP method, Mathematics was more difficult than Physical Sciences.

It should be noted that the statistical significant difference between the two methods was also tested to check if there are any differences between the two methods and the difference was insignificant. As a result, either of the Analytical Hierarchy Process or the Subject Pairs Analysis could be used when ranking subjects according to difficulty.

Table 3.6: Analytical Hierarchy Process from 2014-2018

Subjects	2014	2015	2016	2017	2018
Life Science	0.161	0.162	0.157	0.161	0.156
English FAL	0.210	0.240	0.227	0.226	0.214

Geography	0.178	0.174	0.155	0.142	0.144
History	0.213	0.204	0.220	0.225	0.232
Mathematics	0.117	0.107	0.107	0.114	0.117
Physical Sciences	0.121	0.113	0.134	0.132	0.138

Conclusions

This study was aimed at comparing indices generated by the Subject Pairs Analysis method and the Analytical Hierarchy Process method when ranking subject difficulty between learners taking a common set of subjects. The target sample comprised of learners who sat for all the six subjects namely: Mathematics, Physical Sciences, Life Sciences, Geography, History and English First Additional Language written for Grade 12, National Senior Certificate examinations in 2014 to 2018.

The Subject Pairs Analysis and the Analytical Hierarchy Process were applied on the same set of data. That is, applied on learners who took the same pair of subjects in a given year. The two methods are used when ranking subject difficulty and use the phenomenon of common candidates. Both methods revealed that English First Additional Language and History were the easiest subjects across different years. However, Mathematics and Physical Sciences were consistently difficult across the years.

The results reported by Forrest and Vickerman (1982) indicates that despite some variation, the trend was that languages and chemistry, physics and mathematics were found to be harder than other subjects. Although in South Africa, Physical Sciences is a combination of Physics and Chemistry, This study confirms the findings observed by Forrest and Vickerman (1982) because Physical Sciences and Mathematics were found to be more difficult than other subjects. However, in their study language was also harder. This is contrary to the finding of this study as English FAL was observed to be easier than other subjects. It should be noted that in the South African context, languages are offered at three levels. The first level is at a home language level which should be cognitive more demanding. Languages are also offered on a first additional language level, and second additional level. English is the language of learning and teaching.

Shortcomings of the study

There are several statistical methods that could be used to rank subject difficulty. Coe et al, (2008) highlighted methods such as Latent trait models, Reference tests, value-added models etc. These methods requires an analysis on item level. However, the data used in this study was captured at paper level and as a result, item analysis was not possible. Ideally, more statistical methods should have been explored.

References

Anyagh, P. I., Hanmane, O. and Abah, J. A. (2018). Secondary School Students' Perception of Teachers' Attitude towards Learning in Mathematics in Wukari Metropolis, Taraba State, Nigeria. *International Journal of Research and Review*, 75(5).

Coe, R. Searle, J. Barky, P. Jones, K. and Huggins, S. (2008). *Relative Difficulty of examinations in different subjects.*

Creswell, J. (2003). *Research design: Qualitative, Quantitative and Mixed Methods Approaches Leadership and Management.* SAGE, London, fourth edition.

Mendenhall, W., Beaver, R., and Beaver, B. (2013). *Introduction to Probability and Statistics.* Brooks or Cole: Cengage Learning, United States of America, fourteenth edition.

Mji, M and Makgato, M (2006). Factors associated with high school learners' poor performance: a spotlight on mathematics and physical science. *South African Journal of Education,* 26(2)253–266.

Muijs, D. (2011). *Doing Quantitative Research in Education with SPSS.* SAGE Publication, London, second edition.

Osbourne, J., Simon, S. and Collins, S (2003) Attitudes towards science: A review of the literature and its implications. *International Journal of Science Education,* 25(9)1049-1079.

Osbourne, J., Simon, S. and Collins, S (2010). Attitude towards science: a review of literature and its implications. *International Journal of Science Education*.

Saaty, L.T. and Katz, J.M. (1990). How to make a decision: The Analytical Hierarchy Process. *European Journal of Operational Research,* 48(9 - 26).

Stamatis, D. (2012). *Essential Statistical Concepts for the Quality Professional.* Taylor and Francis Group, Norway.

Sullivan, J. and Krieger, G. (2001). *Clinical Environmental Health and Toxic Exposures.* Lippincott Williams and Wilkins, Philadelphia, second edition.

Tolmie, A., Muijs, D., and McAteer, E. (2011). *Quantitative Methods in Educational and Social Research Using SPSS.* McGraw-Hill Education, United Kingdom.