November 2012

Assessing impact of Aga Khan Academy Mombasa's (AKA, M) outreach programme on Students' Learning Outcomes – an example of English language classrooms

Sadia M. Bhutta
Aga Khan University, Institute for Educational Development, Karachi

Anthony Gioko
Aga Khan Academy, Mombasa

ASSESSING IMPACT OF AGA KHAN ACADEMY MOMBASA'S (AKA, M) OUTREACH PROGRAMME ON STUDENTS' LEARNING OUTCOMES – AN EXAMPLE OF ENGLISH LANGUAGE CLASSROOMS

Sadia M. Bhutta, Aga Khan University Institute for Educational Development

Anthony Gioko, Aga Khan Academy, Mombasa

AKA, M has been offering professional development programmes to English teachers and headteachers as part of its outreach activities since 2009. The Academy invited AKU-IED to develop AKA, M capacity in the area of monitoring and evaluation with a specific focus on outreach programmes. In response to this invitation, a small scale study was carried out to pilot teachers' and students' assessment tools developed as part of the capacity building workshop held in Mombasa, Kenya. The evaluation study aimed to examine the impact of AKA, M outreach programmes on students' (grades 4 to 6) learning outcomes in English language in intervention areas of Mombasa. Three tools were administered to assess students' cognitive (*English Achievement Tests*) and non-cognitive (*English Language Attitude Scale*, *Student-Teacher Relationship Questionnaire*) outcomes. A quasi-experiment was employed for the small-scale study to compare outreach (n=3) and non-outreach (n=3) classes. A sample of 367 students was recruited from the target classes. This paper focuses on pilot results of students' learning outcomes.

In general, students' assessment tools have demonstrated satisfactory psychometric properties. However, a few extremely difficult items can be reviewed before launching the main study. Results of group comparison have revealed that outreach classes have scored higher in achievement tests than their non-outreach counterparts in grade 4 and 6 with a significant difference in the latter ($p<0.05$; $r = 0.16$). Similarly, outreach group has shown a more positive attitude towards English language ($p<0.001$; $r = 0.35$) and has reported more positive student-teacher relationship than non-outreach group ($p<0.05$; $r = 0.13$). Interestingly, those students who have reported more positive attitude towards English language tend to perform better in achievement test ($r = 0.267$; $p<0.001$). Conversely, relationship between students and their English language teacher did not demonstrate a significant relationship with test scores ($p>0.05$).

The study has begun the process of accumulating evidence that the outreach programme evaluated can serve as a model of teacher development within professional development activities of AKA, M. It also provides a modest foundation for a larger and more representative sample study to produce generalizable data.

Introduction

The Aga Khan Academy, a unit of Aga Khan Education Service, Mombasa is part of an integrated network of schools offering an international standard of education from pre-primary to senior-secondary levels with a rigorous academic and leadership experience. One of the distinguishing features of the Aga Khan Academy, Mombasa (AKA, M) is the provision of high quality professional development programmes for educators focusing on subject content acquisition, general pedagogical skills, interactive learning techniques and student-centered

teaching methodologies. These programmes are available both to the Academy's faculty and teachers mainly from government schools[3] with the aim to develop a pool of well-trained teachers regionally. As part of its outreach activities AKA, M has been offering professional development programmes to English teachers since 2009. The programme is designed in three tiers where duration of each tier is 150 hours. *Tier one* aims to improve teachers' own English fluency skills; *tier two* focuses on strengthening teachers' pedagogical skills to teach English language; and *tier three* empower teachers to develop teaching/learning material using low-cost and locally available resources.

In order to develop a monitoring and evaluation framework to assess efficacy of its outreach activities, AKA, M invited the Aga Khan University Institute for Educational Development (AKU-IED) to support it in this regard. The first author represented AKU-IED to help the academy develop an evaluation framework. As part of this collaboration, the representative visited AKA, M in July, 2012 to conduct a 5-day workshop in order to develop a framework for monitoring and evaluation and pilot it on a small sample in English language classrooms in target schools. The workshop was attended by a team of 12 members with diverse backgrounds in teaching and research.

The evaluation study aims to examine the impact of AKA, M outreach professional development programmes on students' (grades 4, 5, 6) learning outcomes (cognitive and non-cognitive) in English language in intervention areas of Mombasa, Kenya. Furthermore, it aims to assess teachers' classroom practices and their views about the school environment. The study is planned in two phases: the *first phase* which has recently been completed aimed to: (i) develop/adapt/adopt assessment tools; and (ii) examine trends of how students/teachers in outreach classes have performed/taught differently from non-outreach classes; *phase two* will be a relatively large-scale impact study. Due to limited sample size, quantitative analysis for data collected from teachers could not be carried out. Hence, this paper focuses on results of students' data.

Impact Study Design

A quasi-experiment design was used for the study where outreach classes (n=3) represented the intervention group while comparison classes (n=3) were recruited from the same school. An inclusion of an outreach class helped measure the progress of the outreach group against a 'soft'comparison group. For the main study comparison classes will be recruited from matching schools.

Sample and Sampling

Three intervention schools were selected for the Mombasa pilot study. Students from outreach (n=184) and non-outreach (n=183) classes participated in the study. The sample represents two systems of schools including government (grades 4 and 6) and private (grade 5). Demographic comparisons revealed that the outreach and non-outreach groups were comparable on most of the aspects (e.g. average age, gender ratio). All workshop participants (n=12) and the facilitator participated in data collection which was completed as part of the workshop in July, 2012.

---

[3]Some private schools have also benefitted from the programme.

Developing/Adapting Impact Assessment Tools

Keeping in mind the nature of the intervention both cognitive and non-cognitive outcomes of students were included in the framework in order to capture their learning in a holistic manner. A summary of students' assessment tools are presented in table 2.

Table 1: A summary of students' assessment tools

Scale	Example item	No. of items
English Achievement Tests	**Content:** Vocabulary, Grammar, Comprehension and Composition **Items:** Multiple Choice Questions (24) and Constructed Response Questions (1)	25 (26 for grade 4)
English achievement test for grade 4, 5 and 6 were adapted from available samples of Zonal English language tests and examples of achievement tests developed and tested in Pakistan.		
English Learning Attitude Scale (ELAS)		
It assesses students' attitude towards English learning in out of school and its importance in life.	*I look forward to my English lessons* *& English helps to get good jobs.*	21
ELAS (Kind, Jones, & Barmby, 2007) was a 5-point Likert scale (1= strongly disagree to 5= strongly agree)		
Students' Views of Teacher-Student Relationships Questionnaire (TSRQ)		
It focusses on student-teacher connectedness, availability of the English teacher and his/her communication skills.	*My English teacher likes me* *& I can go up to my English teacher any time.*	10
Student-Teacher Relationship Questionnaire (TSRQ) was used to elicit students' views on their relationship with English language teachers (Leitão, & Waugh, 2007). It was a 4-point Likert scale (1= not at all to 4= almost always)		

Assessment tools were developed during the AKA, M 5-day workshop through a rigorous multi-step process. Three grade-specific achievements tests were developed and piloted. All tests comprised four sections including: (i) vocabulary; (ii) grammar; (iii) comprehension; and (iv) composition. The first three sections comprised MCQs while the fourth section was designed to give a space for children to write based on pictures (composition). Non-cognitive assessment tools were adapted from other studies and were administered to the whole sample.

Analytical Strategies

A database was developed using Statistical Package for Social Scientists (SPSS) 19. Data for both cognitive and non-cognitive measures were entered at item level. All tests carried equal marks (30); however, the marks distribution across four categories (i.e. vocabulary, grammar, comprehension, composition) in each test was not consistent. Therefore, raw scores were used for the overall average scores; however, it was imperative to use standardised scores for making any sensible comparison at category level. After standardising the scores the mean was set at 500 with a standard deviation (SD) of 100. The score for ELAS negative statements was reversed before computing the mean score. Analyses were undertaken in two steps. Psychometric analysis was carried out to establish reliability and validity of assessment tools. Difficulty indices of test items were computed. In addition, the internal consistency of the students' learning outcomes was assessed using Cronbach's Alpha (Black, 1999; Robson, 2002). A comparison was made for the test scores and students' ratings on non-cognitive measures in order to gauge the difference

between outreach and non-outreach groups. Tests for comparisons were used based on the normality of the data - independent t-test or Mann-Whitney (Field, 2005). Correlations were computed to explore association between cognitive and non-cognitive assessments.

Results

Psychometric Results

A summary of results of Difficulty Indices (*DI*) of achievement tests are presented in table 2. It is evident that there was a wide spectrum of the level of difficulty across grades and a majority of items were in the 'acceptable' range. Altogether there are only five items which were attempted accurately by less than 14% students. These questions will be revised before launching main study.

Table 2: English achievement tests – difficulty index

DI	Less than 0.14		0.15 - 0.30		0.31 - 0.50		0.51 - 0.70		0.71 - 0.90		> 0.90	
Grades (items)	Items	(%)	Items	(%)	Items	(%)	Items	(%)	Items	(%)	Items	(%)
Grade 4 (26)	1	4	2	8	9	35	10	38	4	15	0	0
Grade 5 (25)	2	8	1	4	2	8	5	20	13	52	2	8
Grade 6 (25)	2	8	3	12	9	36	7	28	4	16	0	0
	Extremely difficult		Difficult		Moderatley easy		Easy		Very easy		Extremely easy	

Value of Cronbach's *alpha* (α) for cognitive assessments were within the acceptable limits with the highest being 0.81 (grade 4) and the lowest was 0.60 (grade 5) Similarly, alpha values for all, save one, non-cognitive assessments fell within acceptable range with the highest being 0.83 (ELAS for grade 6) and lowest being 0.64 (TSRQ for grade 4) (Harms, Jacobs & White, 1996; Field, 2005). Grade 6 TSRQ demonstrated a lower level of internal consistency (α <0.6); however, it would be appropriate to retain TSRQ for the main study for which a larger and more diverse sample will be recruited.

Comparisons

English achievement tests: comparison across groups

The total mean test score for this sample was 15.7 (52%; SD = 5.1), while outreach classes scored higher (M=16.1, 53.6%; SD= 5.1) than their non-outreach counterparts (M=15.1, 50.3%; SD= 5.4). However, the difference was not significant [t (365) = 1.216, p>0.05]. As shown in table 3, results have followed the same pattern in grades 4 and 6; however the difference favoured outreach group only in grade 6 (p<0.05) with a small magnitude (r=0.16). Non-outreach group has an edge over outreach group in grade 5; however the difference was not significant (p>0.05).

Table 3: English achievement test – comparison of total score across group by grades

Grades	Outreach	Non-outreach	Significance
Grade 4	15.9 (5.7)	15.3(5.8)	t(140) = 0.561; ns
Grade 5	19.6 (3.6)	20.5(3.0)	t(65) = -1.210; ns

| Grade 6 | 14.7 (4.4) | 13.3(4.3) | t(156) = 2.049; p<0.05 (r = 0.16) |

Figure 1 shows a comparative overview of scaled mean scores across four categories of tests. The outreach group has performed significantly higher in composition (M = 531; SD = 89) than their non-outreach (M = 469; SD = 101) counterparts [t (365) = 6.293; p<0.001] with a medium effect size (r = 0.33). In contrast, non-outreach group has scored slightly higher in vocabulary; however, the difference was not significant (p>0.05).

Figure 1: English achievement test – comparing scores of test categories across groups
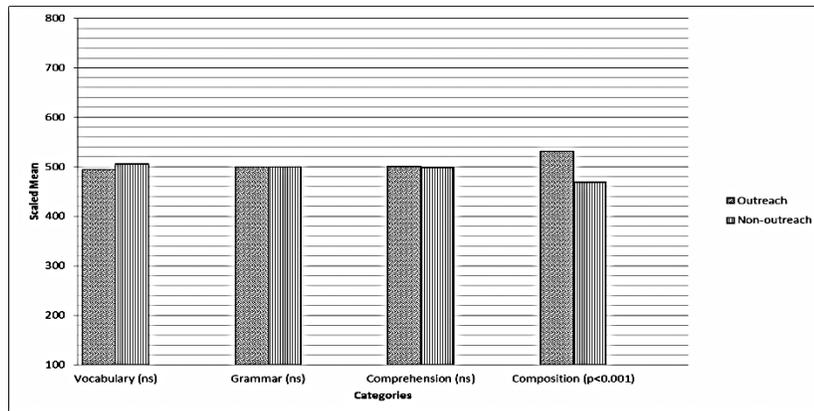

Table 4 presents group differences across grades for each category. Of the twelve comparisons presented in table 4, only four depicted significant differences. Of these differences, three (composition scores for target grades) were in favour of outreach group with a large (grade 4); medium (grade 5) and small (grade 6) effect size. Conversely, grade 4 of non-outreach group has scored higher than outreach group in vocabulary with medium effect size (r=-0.48).

Table 4: English achievement scores – comparing category scores across grades

Vocabulary			
Grades	**Outreach**	**Non-outreach**	**Significance**
Grade 4	574 (103)	600 (93)	t(140) = -1.604; ns
Grade 5	408 (30)	440 (30)	t(65) = -4.360; p<0.001; -0.48
Grade 6	460 (53)	448 (48)	t(156) = 1.389; ns
Grammar			
Grades	**Outreach**	**Non-outreach**	**Significance**
Grade 4	464 (38)	458 (42)	t(140) = 0.986; ns
Grade 5	679 (75)	703 (53)	t(65) = -1.498; ns
Grade 6	454 (32)	454 (29)	t(156) = -0.083; ns
Comprehension			
Grades	**Outreach**	**Non-outreach**	**Significance**
Grade 4	418 (77)	441 (79)	t(140) = -1.817; ns
Grade 5	515 (61)	511 (46)	t(65) = 0.294; ns
Grade 6	570 (81)	545 (103)	t(156) = 1.694; ns

	Composition		
Grades	**Outreach**	**Non-outreach**	**Significance**
Grade 4	499 (63)	405 (74)	t(140) = 8.151; p<0.001 (r= 0.57)
Grade 5	574 (45)	522 (59)	t(65) = 4.059; p<0.001 (r = 0.45)
Grade 6	542 (112)	504 (106)	t(156) = 2.202; p<0.05 (r = 0.17)

Mean (Standard Deviation)

Non-cognitive assessments: comparison across groups

ELAS ratings revealed that students in outreach group have shown a significantly more positive attitude (M = 4.1; SD = 0.5) than their non-outreach (M = 3.7; SD = 0.5) counterparts [Mann-Whitney U = 10096.500; p<0.001] with a medium effect size (r = -0.35). As shown in table 5, results followed the same pattern in grade 4 and 6 with a substantially large magnitude in grade 6 (r= -0.68) and small in grade 4 (r = -0.17). For grade 5 non-outreach group has reported a slightly more positive attitude; however the difference was not significant (p>0.05).

Table 5: Comparing ELAS scores across grades

Grades	**Outreach**	**Non-outreach**	**Significance**
Grade 4	3.9 (0.4)	3.7 (0.6)	Mann-Whitney U = 2024.000; p<0.05 (r = -0.17)
Grade 5	3.9 (0.6)	4.1 (0.4)	Mann-Whitney U = 485.000; ns
Grade 6	4.4 (0.4)	3.6 (0.5)	Mann-Whitney U = 663.500; p<0.01 (r = -0.68)

Mean (Standard Deviation)

TSRQ ratings demonstrated that students in outreach group have experienced a more positive relationship with their teachers (M = 2.9; SD = 0.5) as compared to non-outreach (M = 2.8; SD = 0.5) group [Mann-Whitney U = 14312.500; p<0.05]; however, the magnitude was rather small (r = -0.13). As shown in table 6, results followed the same pattern in grades 5 and 6 with a significant difference only in grade 6 with a medium effect size (r= -0.38). In contrast, for grade 4 the significant difference was in favour of non-outreach group (p<0.05) with a small effect size (r = -0.18).

Table 6: Comparing TSRQ scores across grades

Grades	**Outreach**	**Non-outreach**	**Significance**
Grade 4	3.0 (0.4)	3.2 (0.5)	Mann-Whitney U = 1968.000; p<0.05 (**r = -0.18**)
Grade 5	2.7 (0.5)	2.5 (0.7)	Mann-Whitney U = 413.000; ns
Grade 6	2.9 (0.5)	2.6 (0.3)	Mann-Whitney U = 1768.000; p<0.01 (**r = -0.38**)

Mean (Standard Deviation)

Correlations: cognitive and non-cognitive outcomes

Overall ELAS score has shown a positive association with students' performance (r = 0.267; p<0.001). In other words, those students who reported to have more positive attitude towards English language tend to perform better. Unlike 'more positive attitude' better relationship between teachers and students did not demonstrate a significant association with the overall achievement test scores (p>0.05). As shown in table 8, ELAS has maintained the trend of positive relationship with overall test scores with the relationship being significant for grades 4 and 6. TSRQ followed the trends which were observed for overall scores – no association with test scores.

Table 7: Correlation with overall score of English achievement test

Grades	ELAS	TSRQ
Grade 4	0.27**	0.06
Grade 5	0.07	-0.06
Grade 6	0.41**	0.03
* p<0.05; ** p<0.01		

Conclusion

The newly developed/adapted assessments have managed to 'pick' differences across groups. Although results were not consistent across categories of the achievement test (i.e. vocabulary, grammar, comprehension, composition), more differences were in favour of the outreach group. It is argued that the study has begun the process of accumulating evidence that the outreach programme evaluated can serve as a model of teacher development within professional development activities of the Academy. Evidence is presented for the impact of the language professional development programme on enhanced child outcomes in specific cognitive areas (e.g. composition writing) and selected non-cognitive aspects (e.g. attitude towards English).

In summary, this was the first step that was undertaken to provide a modest foundation for the main study. To illuminate and broaden understanding of the impact of outreach programmes at classroom and student level larger and a more representative sample study is required in order to have sufficient power to produce generalizable data. Furthermore, once the evaluation model is developed by recruiting a larger and more varied sample it can be expanded to add other components for evaluation.

References

Anwar, N. P., & Bhutta, S. M. (2010). *Students' attitude towards science: a study from rural areas of Sindh and Balochistan* (Unpublished research paper). Institute for Educational Development, Aga Khan University, Karachi, Pakistan.

Black, T. (1999). *Doing quantitative research in the social sciences: an integrated approach to research design, measurement and statistics*. London: Sage.

Field, A. (2005). *Discovering statistics using SPSS* (2nd ed.). London: Sage.

Harms, T., Jacobs, E., and White, D.R. (1996). *School Age Care Environment Rating Scale*. New York: Teachers College Press.

Kind, P. M., Jones, K., & Barmby, P. (2007). Developing attitude towards science measures. *International Journal of Science Education*, 29(7), 871-893.

Leitão, N, and Waugh, E. C.(2007). *Students' Views of Teacher-Student Relationships in the Primary School.* A paper presented at the 37th Annual International Educational Research Conference, held by the Australian Association for Research in Education at Fremantle, Western Australia.

Robson, C. (2002). *Real world research: a resource for social scientist and practitioner-researchers* (2nd ed). Oxford: Blackwell.