Available online at www.sciencedirect.com
ScienceDirect
International Journal of Educational Development 28 (2008) 118–131
International Journal of EDUCATIONAL DEVELOPMENT
www.elsevier.com/locate/ijedudev

Evaluating an improved quality preschool program in rural Bangladesh

Anna C. Moore[a], Sadika Akhter[b], Frances E. Aboud[a,c,*]

[a]*ICDDR, B: Centre for Health and Population Research, Dhaka, Bangladesh*
[b]*Plan International Bangladesh, Bangladesh*
[c]*Department of Psychology, 1205 Dr. Penfield Avenue, McGill University, Montreal, Canada H3A 1B1*

Abstract

An important goal of education in developing countries is to implement and improve early childhood education. A pre–post intervention–control design was used to compare a piloted-revised versus a regular preschool program offered by an organization in rural Bangladesh. After 7 months in operation, the quality of the piloted-revised program was higher than the regular program, though the regular program had also improved. Children attending pilot preschools made greater gains than children attending regular preschools on most outcome measures. Action research was conducted alongside the quantitative evaluation to study the process of the implementation and to identify areas for further improvement.

Keywords: Preschool; Bangladesh; Education development; Early childhood education; ECERS

1. Evaluating an improved quality preschool program in rural Bangladesh

Evidence from developed and developing countries shows that children who attend preschool are generally better prepared for formal schooling than children without preschool experience (Engle et al., 2007). The positive effects of preschool education are most marked for those children from less advantaged backgrounds, even after controlling for family selection factors. For these children, preschools can provide extra social and language stimulation that is essential during the early years of life (EFA Global Monitoring Report, 2005; Peisner-Feinberg, 2004). Very few studies from developing countries examine cognitive outcomes in relation to quality of the program. A recent study in Bangladesh found that cognitive and school readiness outcomes were correlated with quality scores, but quality was low (Aboud, 2006). The present study evaluated the implementation of a low-cost improved program in a subset of these rural preschools in comparison with the regular preschools. Changes in the quality of the program and child outcomes were emphasized.

A few studies from developing countries have examined cognitive and school readiness outcomes, but assessment of quality is rare. The Turkish Early Enrichment Project showed that children who

*Corresponding author. Department of Psychology, 1205 Dr. Penfield Avenue, McGill University, Montreal, Canada H3A 1B1.
E-mail address: frances.aboud@mcgill.ca (F.E. Aboud).

0738-0593/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ijedudev.2007.05.003

attended preschools performed better on cognitive measures than children in custodial or home care at a 4-year follow-up. Children also performed better in primary school if during preschool their mothers received weekly training and materials for literacy stimulation, a biweekly support program to help with parenting skills, and regular visits from trained community mothers (Kagitcibasi et al., 2001). In countries such as Bangladesh where many mothers are illiterate, interventions that rely on parents to provide literacy stimulation are not feasible; the preschool must provide this type of input.

In Taiwo and Tyolo's (2002) study in Botswana, children with preschool experience performed better than their peers without preschool experience on English language, mathematics and science during the first few weeks in Grade 1. Another study conducted in Nepal showed that children who attended preschool were more likely to enroll in primary school, had higher levels of school readiness, better school attendance and obtained better results in end of year examinations (Bartlett et al., 2003). Neither of these studies included information about the quality of the preschools and there is no information about the costs of implementing these programs.

1.1. Preschool quality

The quality of the preschool program mediates the positive benefits of attending preschool; put simply, better quality child care is generally related to better cognitive and social outcomes for children (e.g., Sylva et al., 2006). Although it has not yet been possible to establish a causal relationship between quality and child outcomes (NICHD, 2003), most experts are satisfied that the evidence base from many correlational studies is sufficiently convincing to advocate for the provision of high-quality care (Love et al., 2003; Myers, 2004; Peisner-Feinberg, 2004).

Yet what do we mean by quality and what does this mean for preschools in resource-poor countries? Few studies in developing countries have assessed quality in an evaluative sense (see Isley, 2001). Exceptions include a recent study by Aboud (2006) which used the Early Childhood Environment Rating Scale-Revised (ECERS-R), modified for the cultural context of Bangladesh, to evaluate program quality in preschools run by Plan International Bangladesh, a non-governmental organization (NGO) in rural Bangladesh. Although the use of international measures of quality in resource-poor settings has been criticized for inappropriately applying western standards to a cultural context that may value different attributes and skills (Myers, 2004; Prochner, 2002), Aboud (2006) argues that in this instance it was appropriate as the aims of the preschool program were consistent with the qualities measured by the ECERS-R. Furthermore, the activity- and program-related quality scores of these preschools predicted how well children performed on cognitive and school readiness tests.

The Bangladeshi preschools evaluated on the ECERS-R obtained an overall mean quality score within the 'adequate range', although there was variation amongst the preschools and the subscales (Aboud, 2006). Preschools generally performed well on subscales concerned with literacy, mathematics and interpersonal interaction but they performed poorly on subscales concerned with activities and program structure—two key aspects of any preschool program. This explained why preschoolers had acquired so many more school readiness skills than controls, but only small benefits in reasoning and vocabulary. Consequently, implemented changes focused on materials available for play, a more hands-on approach to language and math, and responsive non-instructional talk from the teacher.

1.2. Identifying and evaluating changes

It is hardly surprising that the overall quality of the preschools in Aboud's (2006) study was lower than that reported for child care in developed countries (e.g. Love et al., 2003). In Bangladesh, like most other developing countries, teachers are para-professionals with minimal training and education; learning materials such as books and blocks are limited and class sizes are large. But despite these limitations, Plan International Bangladesh adopted certain recommendations to increase the amount of stimulation children received through materials, activities and instruction. This is particularly important in Bangladesh where children receive little stimulation from play materials, books, or conversations before entering school (UNICEF, 2001).

With support from consultants with expertize in education and child development, local program staff operationalized broad-based recommendations into specific changes to learning materials, teaching

methods and curriculum. The program's previous ECERS profile was used to identify specific qualities to change and how to make improvements.

Curriculum changes in this half-day program gave more prominence to language and literacy: Instead of one daily period reciting rhymes or reading one book a month, there was to be daily story reading with several new stories each week, along with the rhymes. The number of storybooks was increased from 10 to 50, and teachers were taught how to read and talk about stories in an engaging manner rather than requiring memorization. To encourage more free verbal expression from the children, there was to be a Morning News session and a Journal-Writing and Drawing period. This was in addition to the regular literacy class where children learned to write the alphabet. Similarly, the recommendation was to encourage non-verbal reasoning by giving children a math bag with materials, such as matchsticks, buttons and string, to be used for exercises in mathematical reasoning. This was in addition to the use of only three materials (sticks, stones and seeds), the blackboard and exercise books used in the regular program. Free rather than assigned playtime was to take place daily for 40 min and teachers were taught how to use individualized talk to encourage children to expand their repertoires and verbalize their ideas and actions. More play materials were provided (e.g. group games such as memory, imaginative play materials beyond the household theme, puzzles, multi-sized colored blocks), to be periodically rotated. Children were not taught how to draw simple figures as in the regular program, but rather encouraged to draw as they wished and eventually to create scenes. The format of learning was also altered, with a greater emphasis on working in small groups or pairs rather than continuously as one large group. Reasoning was to be emphasized rather than rote repetition and memorization. Technical Officers responsible for training preschool teachers in each district were given additional training that emphasized the need for child-centered learning and particularly the use of non-instructional talk when addressing individual children. Technical Officers then trained the teachers. Non-pilot or regular schools were given some of the additional play materials such as blocks (not puzzles, games or pretend play) and 20 rather than 50 storybooks. Otherwise they kept to the regular program.

Our goal was to evaluate the revised program in a small number of preschools prior to scale-up. The revised program was piloted in 10 preschools at a cost of approximately US$35 extra per school per year, including the extra materials and teacher training. Six of the revised (hereafter called Pilot) preschools and six regular preschools were evaluated in March at the start of the school year, just prior to the introduction of the new program, and again 7 months later in November. We evaluated improvements to quality using the Activities and Program Structure subscales of the ECERS, along with improvements in children's cognitive functioning using standardized and school readiness tests, and social development in terms of play. It was predicted that, compared to the regular preschools, the pilot preschools would show greater improvements from pre- to post-test in quality, and children in these schools would show greater gains in cognitive and social outcomes.

1.3. Action research to follow the process of change

Rather than relying on one single measure or indicator, Myers (2004) supports the use of both qualitative as well as quantitative methods when evaluating preschool education. At the request of Plan, quantitative outcome data on child skills and program quality to be assessed by an independent investigator were combined with process oriented in-house action research. Action research was appropriate as it explicitly seeks to bring about change and improvement and has been used extensively in education (e.g. Asimeng-Boahene, 2004; Sahaesewiyon, 2004; Wijesundera, 2002) and healthcare settings (Meyer, 2005).

Tomal (2003) defines action research as "a systematic process of solving [educational] problems and making improvements". It is a dynamic and cyclical process based on planning, acting, observing and reflecting. The "bottom-up" process is intended to be conducted and owned by those involved, in this case, the local staff of Plan and the teachers themselves (Kember, 2002). This contrasts with other commonly used qualitative methods such as in-depth interviews with key informants and focus group discussions with teachers. Here, we needed to focus on materials, activities and instruction, so observations were essential, as were the inputs of local staff at different levels. Consequently, action research was the method chosen to study the process of implementation.

In summary, quantitative methods were used to examine changes in the quality of the preschool

program and child outcomes in comparison with the regular program; qualitative methods were used to describe how teachers and children were implementing the new materials and activities.

2. Method

2.1. Study design

School quality and child outcome data were obtained using an intervention–control pre–post design. The performance of children in pilot preschools was compared with children in regular preschools delivered by the same NGO, namely Plan Bangladesh. Preschools were selected for the intervention and then regular controls matched for ECERS quality scored the previous year were selected. The children and preschools were pretested before the program started and again 7 months later. Approval of the protocol was provided by the Research Review Committee and the Ethics Review Committee of the ICDDR, B.

2.2. Study population, recruitment and sample

Of the five districts where the organization works, three that were part of the original evaluation were selected for this study, namely Gazipur, Chirirbandar and Jaldhaka (Aboud, 2006). Two preschools from each district were selected to pilot the new program and two matched preschools selected as controls. Overall 12 schools were included in the evaluation, six regular schools and six pilot schools. The pilot and regular schools were geographically distant to avoid contamination, but teachers were supervised by the same Technical Officers. In each school, 15 out of 25–30 children were randomly selected by the research assistants from the class lists, for a total of 186. If a selected child was not present that day, then the 16th randomly selected child was substituted. In the final analysis, only children between 5 and 6.5 years of age at posttest were included in the analyses because they were the appropriate age for preschool (i.e., preprimary schooling) in Bangladesh. Consent was obtained from mothers on behalf of their children. All mothers agreed to participate.

2.3. Measurement

2.3.1. Sociodemographic and nutritional status

Mothers reported on the household members, their age, their child's age, sex, educational attainment and occupation, and the family's religion. Economic status was assessed with questions about the ownership of 11 assets commonly included in the Bangladesh Health and Demographic Surveys (e.g. table, chair, wardrobe, bed, watch, latrine, bicycle, tube well, radio, electricity and television), ownership of a homestead and of land for production, and household income per month. The sum of all assets correlated highly ($p<.01$) with income ($r=.37$), owning land for production ($r=.37$), mother's education ($r=.42$) and father's education ($r=.52$), and was used as the family economic indicator.

Children were weighed on a Uniscale and heights were taken with a meter stick following the usual guidelines concerning head angle and body posture (World Health Organization, 1983). These were converted to height-for-age and weight-for-height z-scores using CDC 2000 guidelines. The child's age at pretest was determined from an immunization or birth registration card if these were available, otherwise from parental report with the help of a Bangla calendar and notable events.

2.3.2. Measurement of preschool quality

The ECERS-R (Harms et al., 1998) is an observational measure that assesses the quality of the program offered in terms of seven subscales. Items from two of the subscales, namely Activities and Program Structure, were used for this study for four reasons: (i) they are the core of a preschool program, (ii) they had been rated lowest in the original evaluation, (iii) they had correlated most highly with children's cognitive scores (Aboud, 2006) and (iv) they had been the focus of recommended changes in the new improved preschool program. Three items were excluded from the Activities/Program Structure subscales because there were no TVs, videos or computers and little attempt to promote diversity or make provisions for children with disabilities. One item, informal use of language, was added to the Activities subscale from the Language and Reasoning subscale. The scale underwent two iterations of translation and back-translation into Bangla. Qualitative items were defined quantitatively for this context, e.g., enough blocks meant 20 per child so 40 were enough for two children, enough space meant 1.5 m^2 per child, a variety of water toys meant five differently shaped objects, and some books meant 10. Inter-rater reliability and predictive validity were very good (Aboud, 2006).

2.3.3. Measurement of child outcomes

Cognitive development was measured with three WPPSI-III (2002) subtests appropriate for children in the 4–7.25 year age range, and a School Readiness test. The WPPSI tests included Vocabulary, Matrix Reasoning and Block Design and were scored on a range 0–19. The tests, along with standard instructions for administration and scoring, had been translated into Bangla and certain items modified for the setting (see Aboud, 2006). Social development was assessed with the Play Observation Scale (Rubin, 2003). Inter-tester reliabilities were good in the previous year's study using the same research assistants. Details on the measures are as follows:

1. Vocabulary assesses children's knowledge of words and their ability to express the word's meaning. For the 25 items, answers were scored out of 0, 1 or 2 according to the item, for a maximum of 43. Scores standardized for age and ranging from 0 to 19 were used in analyses.
2. Matrix Reasoning assesses visual, analytic reasoning in the completion of patterns and analogies. There are 29 items; 3 practice items provide children the rationale for choosing one out of 4 or 5 options. The maximum score was 29; the age standardized score out of 19 was analyzed.
3. Block Design was used to assess visual-spatial, analytic reasoning in the completion of red and white patterns with the use of small cubes. This is the only measure not used previously in Bangladesh, so it will be described in greater detail. Twenty items are presented; the first 10 use red and/or white blocks and the last 10 use red/white bicolor blocks where two sides have both red and white divided along the diagonal and the other sides are either red or white. Items require progressively more difficult patterns. On early items the tester makes a model to be copied by the child, and on later items the model is presented as a two-dimensional picture. Two trials for each of the first 6 items are allowed if the child errs on the first. Time limits of 30, 60 and 90 s were not strictly adhered to. Items are scored as 0, 1, or 2 on the first 6 items and 0 or 2 on the last 14. Although not a part of the standard administration, 8 practice trials were first given using different patterns. This was found to be necessary because a preliminary sample of children showed significant improvement on a second administration of the test, as a result of increased familiarity with the use of blocks to create patterns. Inter-tester reliability, where tests were administered by two assistants several days apart, was equivalent to the other cognitive tests at $r(13) = .60$, $p < .03$.
4. A School Readiness Test developed for Bangladesh was used to assess skills similar to other school readiness tests, namely colors, shapes, letters, numbers, math concepts and nature/health (Aboud, 2006). The maximum score was 30. Performance on this test correlated significantly ($p < .001$) with all age standardized WPPSI measures at pretest and posttest.
5. Social development was measured within the context of free play using the Play Observation Measure (Rubin, 2003). This observational measure assesses three levels of sociability (solitary, parallel and interactive) for four cognitive levels of play (functional, constructive, dramatic and games with rules). Additional play codes include: unoccupied, onlooker, reading, peer conversation, adult conversation and aggression. Children were observed during a 40-min period of free play on 2 separate days. Each observer was responsible for seven children, observing each for 10 s and recording the play category before moving on to the next child, and finally back to the first again for another round. This way, 20 10-s play episodes were observed and coded per day for each child. The number of 10-s units during which the child was engaged in each of the play codes was tallied and expressed as a percent of the total. For example, if the child was working on a puzzle for 10 of the 40 units, he/she received a score of 25% for the category of solitary constructive play.

2.4. Procedure

Six research assistants were trained and experienced in using these measures over a 6-month period before the study started. Refresher training was provided prior to pretest and posttest. All research assistants had university degrees. They were blind to the pilot and regular status of the preschools. Pairs of research assistants spent 2 days in each village collecting the interview and observational data. The pretest data were collected during March 2004; measures were repeated at the posttest from mid-November to mid-December 2004.

Action research was conducted in all 10 pilot preschools, including the six from which quantitative data were collected, between July and October 2004. The key method of study was observation of classroom activities. The observations were done by Technical Officers from the field, whose responsibilities are program development and supervision within the sector of early childhood care and development. They observed preschools not under their supervision in order to avoid conflict of interest. Technical Officers designed checklists to help them structure their observations. The checklists consisted of a series of different questions relating to what and how materials were being used, and to teacher and child behaviors related to the changes.

In addition, a research coordinator from Plan (S.A.) visited each pilot preschool once a month for a full morning to record continuous unstructured observations and her impressions of how the children and teachers were responding to the new teaching materials, activities and instruction. Her field notes included descriptions, direct quotations and comments. These notes were discussed, elaborated and corrected with the Technical Officers, thereby receiving a degree of inter-observer reliability. From these, we identified commonly seen improvements and problems in material use, activities and instruction, along with specific examples. They were then used to assist in our interpretation of the quantitative findings, and in discussions with Plan on the form of a final program.

Every month the Country Office Coordinator of early childhood programs held a workshop with the Technical Officers and the teachers to discuss their observations, and develop strategies for change. Unfortunately no full record was kept of these meetings. A few decisions were made to change course in the middle, but most decisions were delayed until all data were available.

2.5. Method of analysis of quantitative data

Of the original 186 pretested children, 28 (Gazipur $n = 13$; Chirirbandar $n = 11$; Jaldhaka $n = 4$) could not be located at the posttest. The reasons for this were varied; six of the children had dropped out of preschool and one child had moved to another school. The remaining children were simply not available on the day of testing because they were visiting relatives or were sick. Children who could not be located were split evenly between pilot and regular preschools. Twenty children were found to be outside the age range of 5–6.5 years at posttest and so were excluded. This resulted in a final sample size of 138 children (67 from regular preschools; 71 from pilot preschools). Frequencies and mean scores for sociodemographic and nutritional variables were calculated for the two groups of children to identify possible differences. Then the regular and pilot preschools were compared on ECERS-R scores. Finally, child outcomes for the two programs were compared using an analysis of covariance (ANCOVA). Group (regular, pilot) was a between-subjects factor and Time (pretest, posttest) was a within-subject factor. The outcomes were Vocabulary, Matrix Reasoning, Block Design and School Readiness as well as social and cognitive levels of play. Additional analyses were also conducted to examine associations between quality and child outcomes.

3. Results

3.1. Description of sample

Table 1 shows the frequency distribution for the categorized social and economic data for pilot and regular groups. Table 2 provides the means along with *t*-test comparisons of pilot and control groups for continuous variables and nutritional status. At pretest there were no differences between the groups on child age, nutritional status, assets, or parental education. Over half of mothers had received no education at all and 88% were housewives. Most fathers were farmers, merchants or wage laborers. Other SES indicators revealed findings typical of rural Bangladeshi families and children; 20% were stunted at pretest (height for age, $z < -2.0$) and one-third were wasted (weight for height, $z < -2.0$). The sample appeared to be representative of families who participate in Plan's early childhood programs (Aboud, 2006).

Table 3 shows the means for the Activities subscale and Program Structure subscale of the ECERS-R. A Group (regular, pilot) × Time (pretest, posttest) ANOVA was conducted on each subscale. For both subscales the ANOVA yielded a significant main effect of time: Activities, $F(1, 10) = 20.81$, $p = .001$, and Program, $F(1, 10) = 20.25$, $p = .001$. Thus, both types of schools showed improvements in quality over the intervention period. A planned paired *t*-test comparing pretest with posttest scores showed a significant improvement

Table 1
Frequency distribution for categorized sociodemographic and nutritional status variables

	Regular ($n = 67$)		Pilot ($n = 71$)		Total ($n = 138$)	
	No.	%	No.	%	No.	%
Sex						
Male	26	38.8	42	59.2	68	49.3
Female	41	61.2	29	40.8	70	50.7
Mother education						
None	41	61.2	39	54.9	80	58.0
Primary	18	26.9	17	23.9	35	25.4
Secondary +	8	11.9	15	21.1	23	16.7
Father education						
None	35	52.2	41	57.7	76	55.1
Primary	14	20.9	12	16.9	26	18.8
Secondary +	17	25.4	17	23.9	34	24.6
Father occupation						
Farmer	26	38.8	19	26.6	45	32.6
Merchant	16	23.9	26	36.6	42	30.4
Wage labourer	10	14.9	18	25.4	28	20.3
Own home	65	97	66	93	131	94.9
Own land	39	58.2	34	47.9	73	52.8
Religion						
Muslim	45	67.2	58	81.7	103	74.6
Hindu	22	32.8	13	18.3	35	25.4
Height for age						
$z < -2.0$	9	13.4	18	25.4	27	19.6
$-2.0 \leqslant z < -1.0$	33	49.3	33	46.5	66	47.8
$-1.0 \leqslant z < +2$	25	37.3	20	28.2	45	32.6
Weight for height						
$z < -2.0$	17	25.4	26	36.6	43	31.2
$-2.0 \leqslant z < -1.0$	29	43.3	26	36.6	55	39.8
$-1.0 \leqslant z < +2$	21	31.3	19	26.8	40	29.0

Table 2
Means (SD) and t values comparing regular and pilot groups on social, economic and nutritional status variables

	Regular	Pilot	$t(136)$	p
Child age (months)	65.0 (4.4)	64.6 (4.8)	.57	ns
Mother education	2.4 (3.7)	2.8 (3.7)	.64	ns
Father education	3.6 (4.5)	3.0 (4.2)	.77	ns
Household members	5.8 (1.3)	5.3 (1.5)	.37	ns
Assets	5.6 (2.9)	5.5 (2.7)	.19	ns
Height/age				
Pretest	−1.3 (.9)	−1.5 (1.0)	.93	ns
Posttest	−1.4 (.9)	−1.6 (.9)	1.09	ns
Weight/height				
Pretest	−1.4 (.9)	−1.7 (1.3)	1.49	ns
Posttest	−1.5 (1.0)	−1.8 (1.2)	1.37	ns

for the pilot preschools on Activities, $t = 5.86$, $p = .002$, and Program, $t = 10.04$, $p = <.0001$. There was no significant difference for the regular preschools, Activities, $t = 1.81$, $p = .13$, Program, $t = 1.51$, $p = .19$. This indicates that although both types of school improved, the improvement was greatest amongst the pilot schools. For comparison purposes, Table 3 also provides mean ECERS-R scores obtained the previous November for 22 preschools of which these 12 are a subset. The quality of pilot and regular schools was observably higher than those obtained in the original evaluation when mean scores for these same subscales were under 3 (Aboud, 2006); a score of 5 indicates a very good quality program. The possibility that 'contamination' was responsible for improvements in the regular preschools will be discussed later.

3.2. Cognitive and social outcomes

To identify variables that required covarying, correlations of the WPPSI and School Readiness scores with SES and nutrition variables were performed (see Table 4). Standardized Vocabulary and Matrix scores correlated negatively with age indicating that with age children declined in relation to age norms. As expected, they correlated positively with mother's and father's education and assets. Vocabulary, Matrix Reasoning and School Readiness correlated positively with height for age. The four covariates in subsequent analyses on measures of cognition and social play were child's age, height for age, assets and mother's education. Although these did not differ between groups, they were covaried for comparability with other published studies (e.g., Aboud, 2006).

Analyses were conducted to test the prediction that children attending pilot preschools would make greater gains on cognitive and social outcomes than children at regular preschools. A MANCOVA was conducted on the four cognitive measures (Vocabulary, Matrix Reasoning, Block Design and School Readiness). This yielded a significant Group × Time interaction, $F\ (1,\ 130) = 12.61$, $p = .001$, and a significant Group × Time × Cognitive Test interaction, $F\ (3,\ 130) = 3.24$, $p = .02$.

ANCOVA's run separately on each cognitive test yielded significant Group × Time interactions for Matrix Reasoning, Block Design and School Readiness but not Vocabulary. The means for both groups at pretest and posttest are presented in Table 5, along with the F values for the significant

Table 3
Means (SD) of regular and pilot preschools on the Early Childhood Environment Rating Scale-Revised (ECERS-R) along with previous year's evaluation

	Nov. 2003	Regular		Pilot	
	Posttest	Pretest	Posttest	Pretest	Posttest
Activities total	2.5 (.5)	3.3 (1.1)	3.9 (.5)	3.5 (1.1)	4.7 (.6)
Informal use of language	3.9 (1.2)	4.0 (.0)	3.8 (1.6)	3.5 (1.2)	5.3 (1.2)
Fine Motor	2.2 (1.4)	2.3 (1.5)	4.3 (1.9)	3.2 (1.2)	5.7 (.5)
Art	2.8 (.9)	3.2 (.4)	4.0 (1.5)	3.3 (.5)	4.3 (1.8)
Music	2.2 (.9)	3.0 (.0)	3.2 (.4)	3.2 (.4)	3.8 (.8)
Blocks	3.0 (.9)	4.8 (.8)	4.3 (.8)	5.0 (.6)	4.7 (.5)
Sand/water	3.9 (1.2)	4.7 (.8)	4.7 (.8)	5.3 (1.5)	5.0 (.0)
Dramatic play	3.5 (.9)	3.7 (.8)	3.8 (.4)	3.5 (.5)	4.0 (.0)
Nature/science	1.5 (.9)	2.3 (.5)	3.5 (1.2)	2.5 (.6)	4.3 (.5)
Math/number	3.7 (.8)	1.7 (1.0)	3.5 (1.2)	2.0 (.9)	5.0 (.6)
Program structure total	2.64 (.5)	3.4 (2.1)	4.7 (.9)	3.7 (2.4)	6.5 (.5)
Schedule	3.4 (.8)	4.2 (2.4)	4.8 (2.2)	4.5 (2.2)	6.5 (.5)
Free play	3.5 (.8)	5.0 (.9)	5.5 (1.9)	5.5 (.8)	7.0 (.0)
Group time	2.6 (1.2)	1.0 (.0)	3.8 (1.5)	1.0 (.0)	6.0 (1.3)

Table 4
Correlations between pretest cognitive scores and child socio-demographic and nutritional variables

	Vocabulary	Matrices	Block	Readiness
Gender	−.03	.18*	−.21*	.01
Age	−.35**	−.47**	−.08	.06
Mother education	.27**	.29**	.09	.20*
Father education	.19*	.26**	.17	.30**
Assets	.15	.21*	.10	.24**
Height for age	.30**	.19*	.03	.28**
Weight for height	−.04	.01	−.01	−.01

*$p<.05$, ** $p<.01$.
Note: WPPSI scores for Vocabulary, Matrices and Block Design were age standardized.
For gender, boys = 1, girls = 2.

interactions. For Matrix Reasoning, Block Design and School Readiness, the pilot group made greater improvement over the intervention period compared to the regular group but effect sizes were small (between .04 and .08). On Vocabulary, there was overall no significant improvement between pretest and posttest for either group.

Repeated measures ANCOVA's were conducted on play using Group (regular, pilot) as a between-subjects factor and Time (pretest, posttest) along with the three levels of sociability during play (solitary, parallel and interactive) as within-subject factors. Table 6 shows the statistics. There was a main effect of Group, $F(1, 132) = 16.07$, $p < .0001$, indicating a greater proportion of time spent in play as opposed to non-play activities by the pilot children but there was no interaction with time. There was also no Group × Time × Sociability interaction. However, examination of the means indicates that the pilot children showed an increase in interactive play whereas the control group did not, though the difference was not reliable. Interactive play is the most socially sophisticated type of play (Rubin et al., 1983). As expected, solitary and parallel play declined over time.

A similar ANCOVA was conducted using the four levels of cognitive play as within-subject factors (functional, constructive, dramatic and games) (see Table 6). Again this yielded a main effect of Group, $F(1, 132) = 15.45$, $p = .001$, indicating more play overall in the pilot group, but no interactions between Group × Time or Group × Time × Cognitive Play. There was a sizeable increase in games for the pilot group over the intervention period, indicating that the new math and word games offered in the pilot schools were popular. Regular children; however, increased their time with books during play. Both groups of children showed less unoccupied and onlooker behavior over time. There were no main effects or interactions for conversation with adults or peers.

This procedure was repeated, introducing one at a time, the factors of gender (male, female) and

Table 5
Means (SD) and ANCOVA statistics on cognitive indicators of regular and pilot children.

Indicator	Regular		Pilot		Interaction		
	Pretest	Posttest	Pretest	Posttest	Source	F (1, 132)	p
Vocabulary	7.4 (1.3)	7.2 (1.4)	8.0 (1.5)	8.1 (2.3)	G × T	.24	ns
Matrices	4.9 (2.0)	4.0 (1.5)	4.2 (1.7)	4.4 (1.5)	G × T	9.32	.003
Block Design	3.6 (1.7)	4.9 (2.2)	3.7 (1.9)	6.3 (2.5)	G × T	7.73	.006
Readiness	14.8 (4.3)	21.6 (3.8)	14.2 (4.5)	22.9 (5.0)	G × T	6.69	.01

Note: Vocabulary, Matrices and Block scores can range from 0 to 19; Readiness 0–30.
G × T refers to Group × Time interaction.

Table 6
Means (SD) of regular and pilot children on all play categories as a percentage of the total play units observed at pretest and posttest and ANCOVA statistics

Indicator	Regular		Pilot		Interaction		
	Pretest	Posttest	Pretest	Posttest	Source	F (1, 132)	p
Social	52.8	45.2	58.2	54.1			
Solitary	25.6 (14.5)	21.4 (16.7)	33.0 (15.5)	27.3 (13.9)	G × T × S	.90	ns
Parallel	10.5 (9.6)	7.2 (8.1)	8.9 (9.7)	6.3 (7.7)			
Interactive	16.7 (13.8)	16.6 (12.7)	16.3 (12.6)	20.5 (14.3)			
Cognitive							
Functional	16.2 (12.7)	11.9 (12.1)	20.2 (14.5)	13.2 (11.1)	G × T × C	.63	ns
Constructive	18.6 (18.6)	15.5 (12.8)	21.8 (17.9)	19.5 (15.9)			
Dramatic	17.0 (16.9)	15.4 (12.9)	15.7 (12.0)	15.4 (15.2)			
Games	.71 (2.6)	2.0 (6.3)	1.1 (3.5)	6.3 (9.9)			
Other							
Reading	3.4 (6.5)	6.6 (10.1)	4.4 (6.3)	3.0 (5.0)	T × G	7.88	.006
Onlooker	15.5 (9.1)	7.1 (4.6)	11.0 (7.6)	5.9 (4.5)	T × G	4.91	.03
Unoccupied	11.0 (11.8)	5.0 (5.0)	8.7 (7.7)	1.4 (2.6)	T × G	.16	ns
Conversation with adult	4.5 (6.2)	2.4 (3.1)	4.2 (4.3)	3.4 (4.2)	T × G	1.63	ns
Conversation with peers	11.3 (7.8)	7.4 (5.4)	12.0 (8.6)	8.1 (7.8)	T × G	.01	ns

Note: G = group, T = time, S = sociability, C = cognitive level.

nutritional status (below, above the median). Of all the social and cognitive outcomes, there were no significant interactions for either gender or nutritional status. These results indicate that neither sex nor nutritional status modified the benefit obtained from the pilot intervention.

Finally, the mean score for each preschool's cognitive test was correlated with the school's quality score. At posttest, but not pretest, performance on School Readiness correlated significantly with Activities, $r\ (11) = .69$, $p = .01$ and Program Structure, $r\ (11) = .63$, $p = .03$. Thus, the two subscales appear to be valid in terms of being significantly related to school readiness skills.

3.3. Action research

Some of the findings from the action research are summarized below under headings relating to three parts of the curriculum (story telling, math and free play) and teacher–child communication. To help inform year-end decisions about aspects of the program to keep as they were or to change, observations were organized into two categories of information: components that were working well and those that needed improvement. We present here one or two from each with examples. Most decisions on changes to the program waited until all the data were in. However, during monthly meetings, technical officers and teachers who

participated decided to eliminate the pocket charts used for literacy. Teachers were not confident to use them. Technical officers were concerned about the integration of too many new materials and activities, but teachers and students were pleased with the variety, so that did not change.

3.3.1. Storytelling

The number of storybooks in each pilot school was increased from 10 to 50, recommendations were given to rotate these frequently and teachers were trained on how to read books to children. Observations showed that although children enjoyed the books the teachers were not skilled at encouraging participation from the children. Teachers tended to ask simple memory-based or closed-ended questions, such as "What happened next?" or "Did you like the story?" rather than questions related to their comprehension of the story that would promote language development, such as "Why is the man in the story happy?" and "What does this word mean?" It was also observed that the teachers were not rotating the new storybooks often enough, taking approximately 1 week per book, rather than 1 or 2 days, resulting in the children becoming bored. It was decided to provide the teacher with more challenging questions to engage children.

3.3.2. Math

Teachers were making use of the new hands-on materials but although the materials were varied, generally the concepts were not. When learning to count, children practiced counting cubes, leaves, sticks, and pictures of balls, bananas and aeroplanes. They counted in groups, in pairs and on their own. This was repetitive and children lost interest. Teachers had been given math bags for each child's use but their creative use was not added to lesson plans except as another item to count. It was observed that teachers responded well to being given a structured program; for example, they successfully adopted a new format of teaching where children worked in pairs and small groups because this was written into the lesson plans. Therefore, it was decided that Technical Officers would need to provide a series of varied reasoning-based exercises for math and language to be used by teachers.

3.3.3. Free play

Five activity corners were set up for the children, for example, there was a corner for blocks, dramatic play, sand and water, reading, and games and puzzles. Free play was observed to be a popular activity and the majority of children were actively engaged in play for the full 45 min.

Children enjoyed playing with the blocks but the puzzles were less popular for several reasons. The quality of the puzzles was poor and they quickly became damaged. Also, the teachers were unfamiliar with the concept of puzzles and thought it could be learned through a demonstration. Other teachers did not present the puzzles in a graded way according to difficulty—they gave 4 and 12 piece puzzles at the same time, resulting in children becoming overwhelmed and frustrated. This highlighted the need for teachers to be given exercises in their training sessions to verbally coach children on the use of new materials. The games were popular, particularly the memory game, which was usually played with a group of children. Where this was most successful, the teacher played the game with the children once, introducing them to the rules and sequence of actions; and then the children played by themselves with new children joining in. Other games, such as pattern blocks, were much less interesting to the children and were largely ignored, probably because there were no interesting models. It was decided to create a booklet of model patterns, graded for difficulty.

3.3.4. Teacher–child communication

Teachers consistently made an effort to engage with the children, particularly during free play, when they were often observed to go to each activity corner in turn and interact with an individual or a small group. This was a positive change from the previous year when teachers were either passive or highly instructional during playtime. Nonetheless, teachers were still observed to be rather intrusive in their attempts to engage the children, often taking control of the child's game and instructing the child to watch the teacher build a tower or complete a puzzle, rather than supporting the children to do the task themselves. The curriculum included a new emphasis on child expression connected to two new activities that encouraged more talk with children, namely a daily news session and a journal session in which children were encouraged to talk about a picture they had drawn. Children were ready to talk, but some teachers needed to encourage them to use more than a single sentence by asking follow-up open-ended questions.

4. Discussion

The purpose of the study was to examine whether changes to a preschool program in rural Bangladesh, guided by a prior evaluation using the ECERS, were feasible, low-cost and beneficial. The first hypothesis was that the quality of the pilot preschools would show greater gains than regular preschools from the start to the end of the school year. The second hypothesis was that child cognitive and social outcomes would likewise improve more in pilot than regular preschool children. To a certain extent both hypotheses were confirmed, particularly with respect to quality and to non-verbal reasoning and school readiness. These findings will be discussed along with implications from the action research on the difficulties of implementing a higher quality program in a low-resource country using para-professional teachers.

4.1. Preschool quality

Regarding quality of the pilot program, there were significant gains in both subscales of the ECERS-R, namely Activities and Program Structure. The regular program showed some increase but it was not significant with a paired *t*-test. Over a period of 7 months a pilot preschool that was initially scoring approximately 3.5 on the two important subscales was able to improve sufficiently to reach a score of 5.5 (4.7 on Activities and 6.5 on Program). The improvements were done at a cost of approximately US$35 extra for the whole year, or US$1.50 per child per year. This is a minimal expense for a 2-point rise in ECERS quality. Furthermore, the positive relation between cognitive outcomes (i.e. school readiness) and program quality suggests that even over a short implementation period program improvements were associated with some observable gains in child outcomes. Some would suggest that the ECERS sets too high a standard for resource-poor countries such as Bangladesh. Concerning actual finances, rural communities and government ministries may be unable to support preschools costing US$300–$500 on a yearly basis. However, others would argue that it is worth spending a bit more in order to have a beneficial outcome than to spend less for a preschool of mediocre quality with minimal benefits for children.

Changes to the curriculum, materials and teacher instruction had clear links to quality items on which the pilot preschools gained impressively, for example informal language, math, group time and free play. The action research observations provided meaningful examples of the gains and limits. Informal language, for example, was built into the new curriculum by providing time for individual children to talk about their daily experiences in a morning news session or talking about a picture they had drawn to accompany a journal entry. Responsive talk was difficult for most, but not all teachers, the majority of whom did not build on the child's contributions. This was particularly the case in the context of planned exercises such as literacy and maths and also during free play, where teachers had previously been taught to instruct children. So, although informal talk increased in pilot schools, action research showed there was still a tendency for teachers to adopt a directive and instructional style.

Math also increased in quality on the ECERS largely because new math bags with materials were given to each child and teachers were trained to have children use materials to learn math concepts. However, observations showed that teachers were still rather repetitive in their use of different materials to instill the same skill such as counting, and there was still too much learning by copying rather than reasoning.

Teachers in the pilot schools were successful in adopting a new format for learning. This was reflected in higher quality ratings on the Program Structure subscale. Changes on the ECERS (group time) showed that pilot children were spending much less of their time in one large group and more time working individually, in pairs or in small groups for subjects such as math, reading and journal writing. Pilot schools also successfully adopted a daily period of free play into their curriculum; previously children had been assigned to play corners by the teacher. By the posttest, pilot preschools obtained a maximum score of 7 on the free play item of the ECERS, meaning that play materials were varied and regularly rotated, and that teachers supported play activities.

A number of the Activity dimensions rose or remained at the same level in both groups, namely dramatic play, sand and water, blocks, music and art. Both groups were given some of the extra materials before the pretest, as they arrived from the distributors. For example, both were given double the number of blocks from the previous year, this time varying in color, size and shape. So both groups started higher than the previous year but did

not increase over time. Although it was recommended that dramatic play items be improved to cover more than the household theme along with dress-up material, these were not observed. Similarly, dramatic play did not increase on Rubin's Play Observation measure. Without materials to arouse their imagination, children did not sustain social dramatic play.

4.2. Child outcomes

Pilot preschool children also made greater gains compared to regular preschoolers on some but not all cognitive and social outcomes. Regarding cognitive measures, the pilot preschool children benefited more than the regular preschool children on Matrices and Block Design. Both reflect non-verbal reasoning skills, which the children had performed poorly on the year prior. Consequently, it appears that the introduction of new math materials and math games, as part of the math class and free play, may have benefited them (Greenes et al., 2004). Greater improvements on School Readiness by pilot children showed that they were better prepared for the demands of Grade 1 on key concepts of language and math. However, both groups performed well on this measure, as they had the previous year (Aboud, 2006). This is attributed to the strong instructional lessons given to children on literacy and math.

There was no improvement on Vocabulary among pilot preschoolers despite the introduction of 50 instead of 10 storybooks and time for daily reading. Action research suggested reasons for this; teachers were continuing to use storybooks as an exercise of memory rather than as an opportunity for expressive language and vocabulary development (Aram and Biron, 2004; Hargrave and Senechal, 2000). Also, teachers were still reading the same book every day for the whole week instead of rotating them. A decision was taken to introduce dialogic reading (Opel et al., 2006a) and provide the teachers with open-ended questions to engage students in thoughtful dialog.

In terms of social play, there were no significant differences between pilot and regular children. However, examination of the means shows that pilot children made notable but not reliable gains relative to regular children on interactive play. Interactive play is the most socially sophisticated type of play and it is particularly important for social and cognitive development (Rubin et al., 1983). In terms of cognitive play, pilot children played significantly more group games than control children. This is a positive finding as it indicates that the children in pilot schools were using the new math and language games available during free play. As discussed earlier, it is disappointing that the children in the pilot preschools did not show more dramatic play, perhaps due to the limited materials for different themes such as transportation and occupations. This and other noted limitations have since been rectified.

Despite small improvements on cognitive outcomes, which may not match the more impressive increases in quality, we felt that benefits were being passed on to the children. The new instructional style clearly was difficult to implement in such a short period. Teachers required more support from technical officers to develop reasoning and conceptualization skills amongst their students. It is not reasonable to expect para-professional teachers, who themselves have limited education and in-service training, to make lesson plans that include student reasoning. Teacher training introduced a new child-focussed style of learning, compatible with the local organization's emphasis on individualized and child-friendly learning, and teachers were actively discouraged from using only copying and repetition. However, they were not given concrete instructions on how to use the materials in a varied way during lessons. As a result of the action research it was decided that Technical Officers would provide the teachers more lesson plans with a range of exercises in math and language, for them to follow (Opel et al., 2006b).

The most obvious study limitation relates to the problem of contamination. Both schools received some new materials (e.g. books and blocks) and although the schools were geographically distant, they were supervised by the same Technical Officers. Following the original evaluation (Aboud, 2006), the Technical Officers received training in children development and preschool education from specialists. For the first time, many visited and conducted skilled observations of the schools they supervised, both regular and pilot. It is likely that regular schools benefited from the new enthusiasm and training of their supervisors. This is most clearly evident in the data from the ECERS-R, where regular schools also showed notable improvement on some items over the study year and in comparison with the prior year's evaluation. Another limitation was the difficulty in keeping a full

record of the action research's qualitative data; it seems that the cyclical processes was not fully used in that after observing and reflecting there was no new plan or new action. However, our interpretation of the quantitative findings became clearer for all researchers, quantitative and action researchers, when juxtaposed with the qualitative observations. The two data sets were fully integrated for purposes of interpreting the findings and recommending improvement.

5. Implications

A current debate exists on whether developing countries such as Bangladesh should introduce a high-quality preschool program guided by an international tool such as the ECERS-R. Arguments include the different learning styles of children from different cultures, and the detrimental negative evaluation that results from an unattainably high standard (Myers, 2004). Concerning the first point, it is clear that there are different teaching styles, in that the teachers here preferred to teach by demonstrating and requiring the children to copy and repeat. However, local field staff noted that children had stopped attending and thinking by the fifth repetition. The cultural emphasis on memorization of math phrases and stories, rather than on reasoning and vocabulary, jeopardized children's interest and comprehension. Furthermore, higher ECERS qualities were empirically associated with higher cognitive outcomes among poor rural Bangladeshi children (Aboud, 2006). Consequently, although adults may prefer to teach by demonstration and repetition, children may not learn math and language best by these methods. Children from different cultures may benefit similarly from the qualities proposed by the ECERS (Burchinal and Cryer, 2003).

This study addresses the second point head-on by demonstrating that when used as a tool to guide improvement, in addition to its summative evaluation, the ECERS is useful in Bangladesh. In its first year of implementation, the new improved program showed significant gains in quality. Two key subscale qualities were addressed, namely Activities and Program Structure, by introducing new materials, daily activities, and individualized, non-instructional talk. For an extra US$ 1.50 per child per year, quality and child outcomes were improved. The benefits to all concerned were so noticeable that it was impossible to prevent regular preschools from adopting many of the piloted changes. In the year since, some of the shortcomings noted here have been remedied and all preschools of Plan and its local partners around the country have adopted the new program. Serious concerns may be raised as to the quality of the scaled-up program; this requires continued assessment. So the same measure of quality was, and will be, useful for making decisions about the current status of the program and ways to improve it (Chatterji, 2003).

Acknowledgments

We would like to acknowledge the financial support of Plan International Bangladesh and the scientific support of the International Centre for Health and Population Research (ICDDR,B), Dhaka, where the research was subjected to scientific and ethical review, Grant no. 2003-029. We would also like to thank Plan's Early Childhood coordinator, Mahmuda Akhter, the Technical Officers, teachers, research assistants, and the mothers and children who participated.

References

Aboud, F.E., 2006. Evaluation of an early childhood preschool program in rural Bangladesh. Early Childhood Research Quarterly 21, 46–60.

Aram, D., Biron, S., 2004. Joint storybook reading and joint writing interventions among low SES preschoolers: differential contributions to early literacy. Early Childhood Research Quarterly 19, 588–610.

Asimeng-Boahene, L., 2004. The prospects and problems of African social studies teachers as action researchers: a spotlight on Ghana. Educational Action Research 12, 273–285.

Bartlett, S., Arnold, C., Sapkota, P., 2003. What's the difference? The Impact of Early Childhood Development Programs. Save the Children (US), Kathmandu, Nepal:

Burchinal, M.R., Cryer, D., 2003. Diversity, child care quality, and developmental outcomes. Early Childhood Research Quarterly 18, 401–426.

Chatterji, M., 2003. Designing and Using Tools for Educational Assessment. Allyn & Bacon, New York.

EFA Global Monitoring Report, 2005. The Quality of ECCE and Literacy Programs. UNESCO (Chapter 2).

Engle, P.L., Black, M.M., Behrman, J.R., Cabral de Mello, M., Gertler, P.J., Kapiriri, L., Martorell, R., Eming Young, M., and the International Child Development Steering Committee, 2007. Strategies to avoid the loss of developmental potential among over 200 million children in the developing world. Lancet 369, 229–242.

Greenes, C., Ginsburg, H.P., Balfanz, R., 2004. Big math for little kids. Early Childhood Research Quarterly 19, 159–166.

Hargrave, A.C., Senechal, M., 2000. A book reading intervention with preschool children who have limited vocabularies: the benefits of regular reading and dialogic reading. Early Childhood Research Quarterly 15, 25–90.

Harms, T., Clifford, R.M., Cryer, D., 1998. Early Childhood Environment Rating Scale, revised edition. Teachers' College, Colombia University, New York.

Isley, B.J., 2001. Tamil Nadu Early Childhood Environment Rating Scale. M.S. Swaminathan Research Foundation, Chennai, India.

Kagitcibasi, C., Sunar, D., Bekman, S., 2001. Long-term effects of early intervention: Turkish low-income mothers and children. Applied Developmental Psychology 22, 333–361.

Kember, D., 2002. Long-term outcomes of educational action research projects. Educational Action Research 10, 83–104.

Love, J.M., Harrison, L., Sagi-Schwartz, A., et al., 2003. Child care quality matters: how conclusions may vary with context. Child Development 74, 1021–1033.

Meyer, J., 2005. Qualitative research in health care. Using qualitative methods in health related action research. British Medical Journal 320, 178–181. HYPERLINK 〈http://bmj.com/cgi/content/full/320/7228/178#BIBL〉.

Myers, R.G., 2004. In search of quality in programs of early childhood care and education. Background paper for Education for All, Global Monitoring Report 2005. UNESCO, Paris, 〈www.unesco.org/education/gmr_download/references_2005.pdf〉. Retrieved October 30, 2005. Also Mimeo, Hacia Una Cultura Democratica, Mexico, D.F.

NICHD Early Child Care Research Network, 2003. Does quality of child care affect child outcomes at age $4\frac{1}{2}$? Developmental Psychology 39, 451–469.

Opel, A., Ameer, S.S., Aboud, F., 2006a. A Literacy Intervention for Preschool Children in Bangladesh: The Benefits of Dialogic Reading. BRAC University Institute of Educational Development, Dhaka, Bangladesh.

Opel, A., Camellia, S., Aboud, F., 2006b. Playing With Mathematics: A Pilot Intervention to Develop Basic Mathematical Skills Among Preschoolers in Bangladesh. BRAC University Institute of Educational Development, Dhaka, Bangladesh.

Peisner-Feinberg, E.S., 2004. Childcare and its impact on young children's development. In: Tremblay, R.E., Barr, R.G., Peters, R. (Eds.). Encyclopedia on Early Childhood Development, Centre of Excellence for Early Childhood Development, Montreal, Que., pp. 1–7. Available at HYPERLINK 〈www.excellence-earlychildhood.ca/documents/Peisner-FeinergANGxp.pdf〉.

Prochner, L., 2002. Preschool and playway in India. Childhood: A Global Journal of Child Research 9, 435–453.

Rubin, K., 2003. Play Observation Scale: A Manual for Coding Free Play Behaviours of Young Children. University of Maryland.

Rubin, K., Fein, G.G., Vandenberg, B., 1983. Play. In: Mussen, P.H., (Ed.), Handbook of Child Psychology, Vol. IV: Socialization, Personality, and Social Development (Hetherington, E.M., Vol. Ed.). Wiley, New York.

Sahaesewiyon, K., 2004. Working locally as a true professional: case studies in the development of local curriculum through action research in the context of Thai schools. Educational Action Research 12, 493–531.

Sylva, K., Siraj-Blatchford, I., Taggart, B., Sammons, P., Melhuish, E., Elliot, K., Totsika, V., 2006. Capturing quality in early childhood through environmental rating scales. Early Childhood Research Quarterly 21, 76–92.

Taiwo, A.A., Tyolo, J.B., 2002. The effect of pre-school education on academic performance in primary school: a case study of grade one pupils in Botswana. International Journal of Educational Development 22, 169–180.

Tomal, D.R., 2003. Action Research for Educators. A Scarecrow Education Book, USA.

UNICEF, 2001. Baseline survey for caregivers' KAP on early childhood development in Bangladesh. Dhaka.

Wechsler Preschool and Primary Scale of Intelligence—Third Edition (WPPSI-III), 2002. The Psychological Corporation.

Wijesundera, S., 2002. School improvement: an action-based case study conducted in a disadvantaged school in Sri Lanka. Educational Action Research 10, 1–88.

World Health Organization, 1983. Measuring change in nutritional status. Guidelines for assessing the nutritional impact of supplementary feeding programs for vulnerable groups, Geneva.