

EXPLORING OUTCOME-BASED EDUCATION IN EARLY CHILDHOOD: A DOCUMENTARY ANALYSIS OF SKILL-EVALUATION RECORDS AT FRONTIER CHILDREN'S ACADEMY, PESHAWAR, PAKISTAN

AARIF GOHAR SALARZAI^{1*} AND SYED MUNIR AHMAD²

¹ PHD SCHOLAR, INSTITUTE OF EDUCATION AND RESEARCH, UNIVERSITY OF PESHAWAR, KHYBER PAKHTUNKHWA, PAKISTAN

² ASSOCIATE PROFESSOR, INSTITUTE OF EDUCATION AND RESEARCH, UNIVERSITY OF PESHAWAR, KHYBER PAKHTUNKHWA, PAKISTAN

CORRESPONDING AUTHOR: AARIF GOHAR SALARZAI, EMAIL: ARGOSAL@GMAIL.COM

Abstract

Outcome-based education (OBE) is well established in Pakistani higher education, but little is known about how it operates in preschools. This study used a descriptive, cross-sectional documentary design, nested within a qualitative single-case study, to examine the routine skill-evaluation records of a private preschool in Peshawar. Class teachers had rated 364 children in 16 classes across three year-groups (Play Group, Nursery, and Preparatory) on a four-point ordinal scale (A+ to C, coded 4 to 1). Ratings were averaged into composite scores and summarized with means, grade distributions, and class-level comparisons; Kruskal–Wallis and one-way ANOVA tests, with class as the unit of analysis, were used for an exploratory comparison of year-groups. Mean attainment was similar in Play Group ($M = 3.02$) and Nursery ($M = 3.03$) and lower in Preparatory ($M = 2.76$), although the class-level difference was not statistically significant, $H(2) = 2.21$, $p = .331$. Within Nursery, Writing was the weakest domain ($M = 2.72$) and Gross Motor Skills the strongest ($M = 3.26$). Differences between classes of the same year-group were largest in Preparatory (range = 1.15 points). The patterns are interpreted through Vygotsky's zone of proximal development. We suggest a shared moderation rubric for teacher ratings, structured writing instruction in Nursery and Preparatory, peer observation between higher- and lower-scoring classes, and child-level record keeping that allows the same children to be followed over time.

Keywords: outcome-based education, preschool assessment, early childhood, documentary analysis, school readiness, Pakistan, zone of proximal development.

INTRODUCTION

Outcome-based education (OBE) asks a school to state, in advance, what a learner should be able to do at the end of a course or term, and then to plan teaching and assessment around that statement (Biggs & Tang, 2011; Spady, 1994). The model grew out of school reform and higher education, and most published accounts of its use still come from universities and technical programs. Pakistan follows this pattern. Interest in OBE has been driven largely by the Higher Education Commission and professional accreditation bodies, which regard it as a way of bringing local degrees into line with international expectations (Mahmood, 2021). Whether and how OBE principles appear in preschool classrooms is far less documented.

The preschool years are a particularly important place to ask the question. Early skill formation shapes later learning and adult outcomes (Black et al., 2017; Heckman, 2006), self-regulation and early academic skills at school entry are linked to later adjustment (Blair & Raver, 2015), and programs that deliberately target

specific readiness skills can improve them (Bierman et al., 2008). Evidence from the Effective Pre-school and Primary Education project likewise shows that the quality of a child's preschool experience is associated with later progress (Sylva et al., 2010). Against this, assessing young children is difficult. Judgments are sensitive to who makes them and how, and no single measure captures what a three- or four-year-old can do (Snow & Van Hemel, 2008). A preschool that adopts outcome language therefore needs some way of checking whether its own records show the progression that the language implies.

Most preschools already hold such records. Frontier Children's Academy (FCA), a private school in Hayatabad, Peshawar, requires class teachers to rate every child on a set of learning-outcome domains each term, mainly to report to parents. These ratings were not collected for research, but institutional documents of this kind are an accepted source of evidence in case study research (Bowen, 2009; Yin, 2018). They allow three questions to be asked that are rarely asked of school records. First, does measured attainment rise as children move through the school? Second, are some skill areas weaker than others across the school? Third, do parallel classes at the same level produce comparable results?

This article reports the documentary component of a wider qualitative case study of OBE at FCA (Salarzai, 2026). To the authors' knowledge, no published study has examined routine outcome records from a preschool in Peshawar in this way. The analysis is descriptive and exploratory, and it is intended both to stand as an account of what the records show and to generate focused questions for the interview and focus-group strand of the larger study.

RESEARCH QUESTIONS

The study addressed four questions:

RQ1. How do children at FCA perform, on average, at each of three year-groups (Play Group, Nursery, and Preparatory)?

RQ2. Which learning-outcome domains are relatively strong or weak school-wide, where the records allow this comparison?

RQ3. Does mean attainment rise from one year-group to the next, as a sequenced OBE curriculum would lead one to expect?

RQ4. How much do parallel classes within the same year-group differ from one another?

THEORETICAL FRAMEWORK

The study draws on Vygotsky's (1978) account of learning as a socially mediated process in which children develop through interaction with teachers, parents, and more capable peers. Its central idea for present purposes is the zone of proximal development (ZPD): the distance between what a child can do alone and what the child can do with suitable support. Wood et al. (1976) described the support given within that zone as scaffolding. A teacher who sets a task beyond the ZPD gives the child nothing to hold on to, and a teacher who offers no help leaves the child where they are. Judging how much support to give is, in many respects, the core skill of early-years teaching.

The link with OBE is close. OBE requires teachers to know what a child should be able to do next and to teach toward it, which in practice resembles identifying the ZPD and working inside it (Biggs & Tang, 2011). Where OBE departs from a purely Vygotskian classroom is in insisting that the intended outcome be written down and measured. That measurement, however imperfect with young children, is what makes a records-based analysis possible.

The school's evaluation instrument was not designed with Vygotsky in mind, but its domains can be read through this framework. Spoken Skills, Reading Skills, Writing Skills, Number Concepts, and General Knowledge are shaped directly by classroom instruction and are therefore the domains a teacher can scaffold term by term. Fine and Gross Motor Skills, Social Skills, Discipline and Manners, Confidence, and Behavior develop more diffusely through play, peer contact, and maturation. Parents' Cooperation can be read as an indication of how much support is available at home, a second ZPD running alongside the one in the classroom. This reading yields a testable expectation: instruction-dependent domains should vary

more between classes than diffusely developing ones, because scaffolding quality differs from room to room. The data permit this expectation to be examined only partly, and the Results and Discussion sections say where.

METHOD

Design

The study is the documentary component of a wider qualitative, single-case study of OBE implementation at FCA. The case study follows Yin's (2018) approach of examining one setting closely and combines semi-structured interviews, focus groups, classroom observation, and document analysis (Merriam, 2009). The component reported here is a descriptive, cross-sectional, secondary analysis of existing institutional records. The numerical analysis is therefore descriptive, with a small exploratory inferential step, and it does not seek to estimate effects.

Setting and Records

FCA is a private school in Hayatabad, Peshawar, that enrolls children from a range of socioeconomic backgrounds. At the time of the analysis its Junior section comprised 16 classes in three year-groups: Play Group (the entry stage), Nursery (recorded in the school's files as O-B), and Preparatory (recorded as O-A, the last stage before primary school). Each class has its own teacher. The records covered 364 children (Table 1).

Table 1 Sample by Year-Group

Year-group	Classes	Children	Section labels
Play Group	3	59	PG-1 to PG-3
Nursery (O-B)	6	127	OB-I to OB-VI
Preparatory (O-A)	7	178	OA-I to OA-VII
Total	16	364	—

Instrument and Coding

The records are the school's internal student evaluation sheets, completed by each class teacher on the basis of ordinary classroom observation as part of routine reporting to parents. Each child is rated on every domain using a four-band ordinal scale: A+ (Excellent), A (Very Good), B (Good), and C (Needs Improvement). This is a graded rating scale, not an attitude scale of the kind Likert originally described, and the bands are ordered categories rather than equal-interval measurements. For descriptive purposes the bands were coded A+ = 4, A = 3, B = 2, and C = 1, a convention that is widely used when ordinal ratings are summarized as means (Norman, 2010; Sullivan & Artino, 2013). Because the coding treats the intervals between bands as equal, inferential comparisons were made with a rank-based test as well as with ANOVA (see Data Analysis).

The level of detail in the records differs by stage. For Nursery, full domain-level ratings were available for twelve skills: Spoken Skills, Reading Skills, Writing Skills, Number Concepts, General Knowledge, Fine Motor Skills, Gross Motor Skills, Social Skills, Discipline and Manners, Confidence, Behavior, and Parents' Cooperation. For Play Group, the same twelve domains are reported in three composites (Academic, Motor, and Interpersonal). For Preparatory, the records available reported composite attainment for each section, together with a qualitative summary of strengths and gaps against seven broader OBE competency areas, but not a numerical breakdown by domain. Domain-level findings are therefore reported for Nursery only, and no domain-level figures have been constructed for the other year-groups.

Data Analysis

Analysis proceeded in three steps that correspond to the research questions. First, composite scores for each year-group were calculated as enrolment-weighted means across classes on the 1.00–4.00 scale, and the distribution of grades was tabulated (RQ1, RQ3). Second, where domain-level ratings existed, domain means were compared descriptively (RQ2). Third, each class was compared with the others in its year-group by ranking class means and calculating the range and standard deviation of class means within each year-group (RQ4).

The records available to the authors were tabulated at class level and did not contain child-level scores. Inferential testing was therefore limited to the 16 class means, with the class as the unit of analysis. A Kruskal–Wallis test (Kruskal & Wallis, 1952) and a one-way ANOVA were used to compare the three year-groups; the Kruskal–Wallis test is the primary test because the coded scores are ordinal and the groups are small and unequal. The effect size for the Kruskal–Wallis test is reported as $\epsilon^2 = H/(n - 1)$. Statistical significance was set at $p < .05$, and calculations were performed in Python using the SciPy library. These tests are exploratory. With only 3, 6, and 7 classes per year-group they have little power, and they cannot account for differences among children within a class.

Ethical Considerations

Only de-identified, aggregated figures are reported. No child, parent, or teacher is named or identifiable in the results; classes are identified by section label only. The school gave permission for the secondary use of its records, and the work forms part of the doctoral research project of the first author at the Institute of Education and Research, University of Peshawar. The first author holds an administrative post at FCA. To limit any influence on the findings, the data were analyzed only at the aggregate level and the second author reviewed the analysis.

RESULTS

Attainment Across the Three Year-Groups

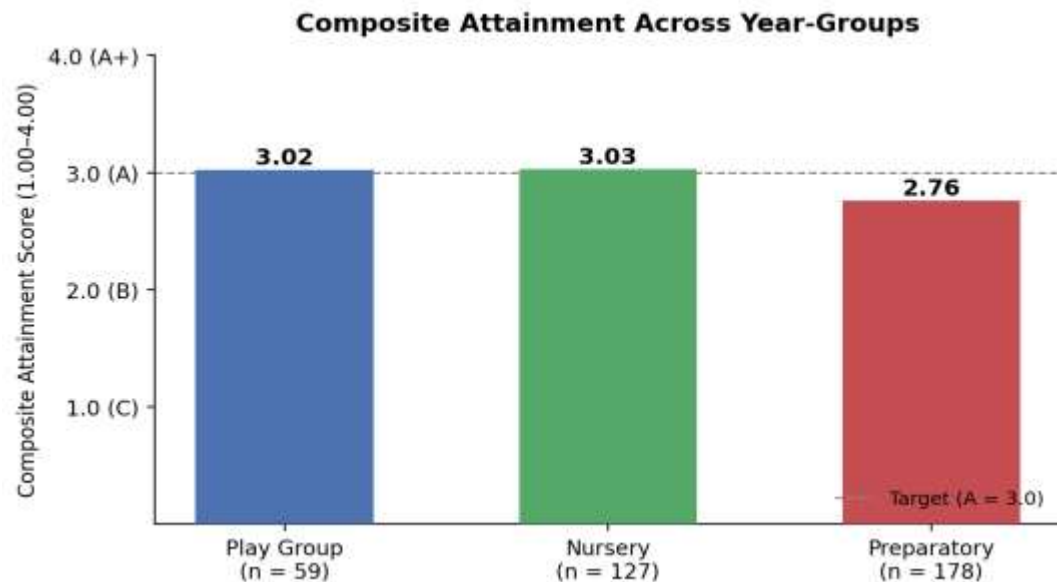
Attainment did not rise from one year-group to the next (Table 2, Figure 1). Play Group ($M = 3.02$) and Nursery ($M = 3.03$) were almost identical. Preparatory, the stage closest to formal schooling, had a lower mean ($M = 2.76$), 0.27 points below Nursery. A sequenced curriculum would more plausibly lead one to expect an increase. This difference was not statistically significant when the 16 class means were compared, $H(2) = 2.21$, $p = .331$, $\epsilon^2 = .15$; $F(2, 13) = 1.23$, $p = .324$. The result should be read as a descriptive pattern, and the small number of classes means that a real difference of this size could not be reliably detected.

Table 2 Composite Attainment and Grade Distribution by Year-Group

Year-group	n	Composite M	% A+	% A	% B	% C
Play Group	59	3.02	41.7	27.4	20.5	10.5
Nursery (O-B)	127	3.03	27.0	57.5	13.4	2.4
Preparatory (O-A)	178	2.76	15.7	31.5	43.3	9.6

Note. Composite scores are on a 1.00–4.00 scale ($A+ = 4$, $A = 3$, $B = 2$, $C = 1$).

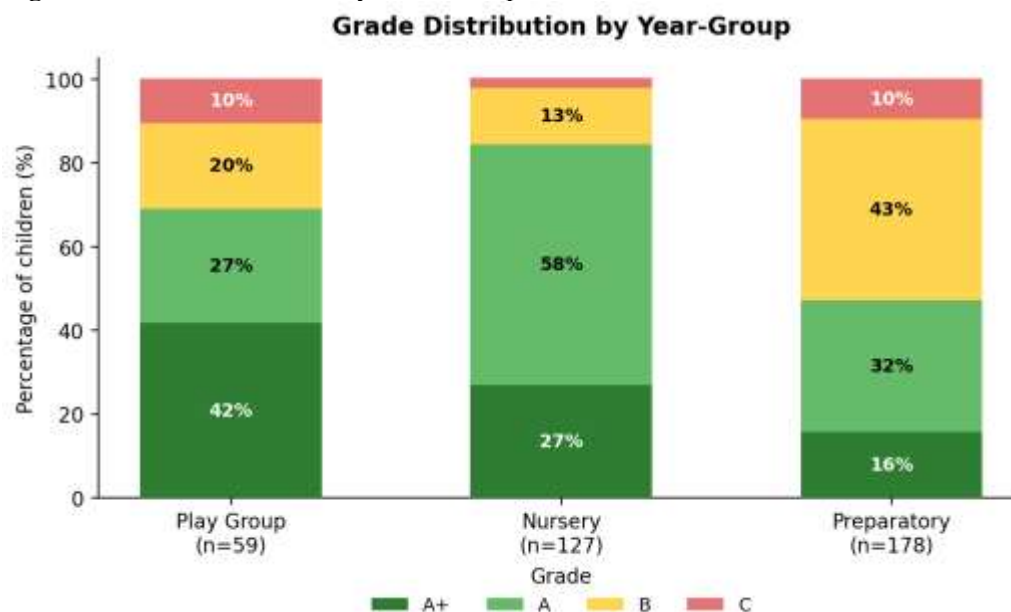
Figure 1 Composite Attainment Across the Three Year-Groups



Note. The dashed line marks the A-grade threshold (3.0).

The shape of the difference is informative. The lower Preparatory mean does not come from more children being rated C: the proportion rated C in Preparatory (9.6%) is similar to that in Play Group (10.5%) and higher than in Nursery (2.4%). The larger change is in the upper bands. The A+ category falls from 41.7% in Play Group and 27.0% in Nursery to 15.7% in Preparatory, and the B band grows to 43.3% of Preparatory children (Figure 2). Few Preparatory children are rated as struggling; many are rated as adequate.

Figure 2 Grade Distribution by Year-Group



Note. Percentages are shown rounded to whole numbers in the figure; Table 2 gives one decimal place.

Domain-Level Performance in Nursery

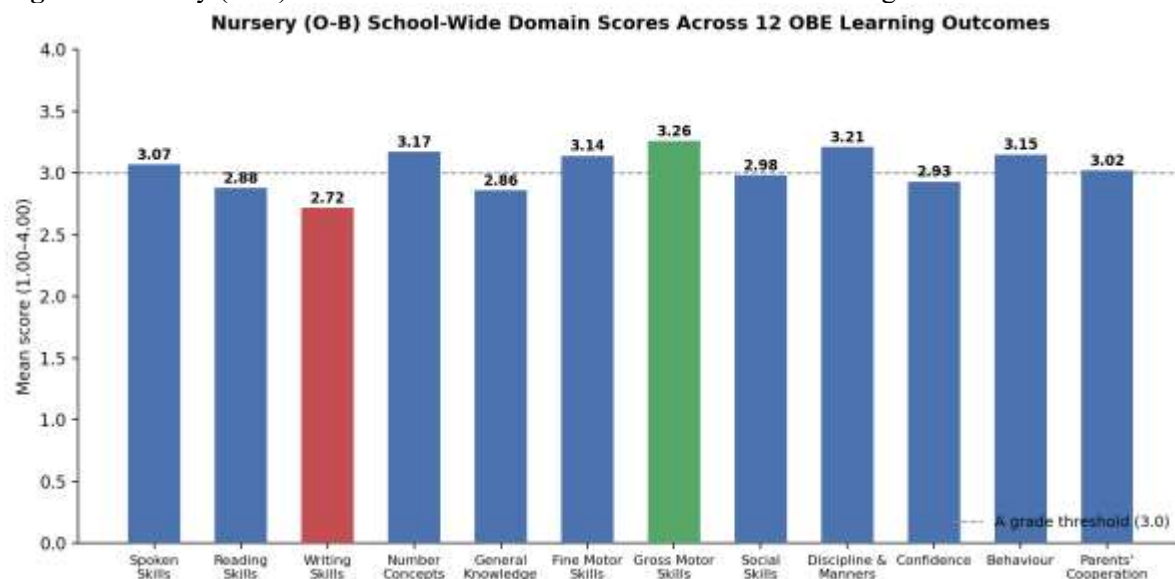
Full domain-level detail was available for Nursery only (Table 3, Figure 3). Writing Skills was the weakest domain ($M = 2.72$), followed by General Knowledge ($M = 2.86$) and Reading Skills ($M = 2.88$). Gross Motor Skills was the strongest ($M = 3.26$), followed by Discipline and Manners ($M = 3.21$) and Behavior ($M = 3.15$). This is broadly in line with the expectation set out in the Theoretical Framework: the domain most dependent on structured instruction, writing in particular, was weakest, whereas domains that develop through maturation, play, and routine were comparatively strong. Number Concepts ($M = 3.17$) is a notable exception among the instruction-dependent domains.

Table 3 Nursery (O-B) School-Wide Mean Rating by Learning-Outcome Domain

Domain	M	Domain	M
Spoken Skills	3.07	Gross Motor Skills	3.26
Reading Skills	2.88	Social Skills	2.98
Writing Skills	2.72	Discipline and Manners	3.21
Number Concepts	3.17	Confidence	2.93
General Knowledge	2.86	Behavior	3.15
Fine Motor Skills	3.14	Parents' Cooperation	3.02

Note. Scale 1.00–4.00. Values are those displayed in Figure 3.

Figure 3 Nursery (O-B) School-Wide Mean Score Across Twelve Learning-Outcome Domains



Note. The dashed line marks the A-grade threshold (3.0). The weakest domain is shown in red and the strongest in green.

At Play Group, where the twelve skills are grouped into three composites, the Interpersonal composite was the strongest in every class, and the Academic composite varied most between classes, from 2.6 in PG-1 to 3.1 in PG-2. For Preparatory, the records reported strengths and gaps qualitatively. The school's own internal analysis flagged communication and literacy as a comparative weakness in the lowest-scoring section (OA-VII). This is consistent with the Nursery pattern but cannot be confirmed with the same precision.

Variation Between Classes Within Year-Groups

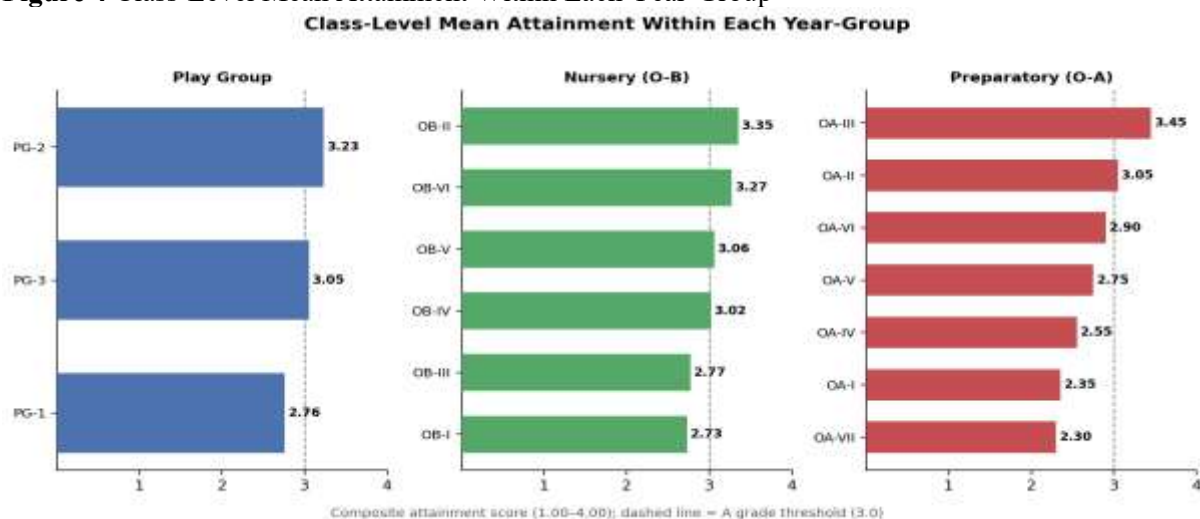
Table 4 and Figure 4 rank each class within its year-group. The gap between the highest- and lowest-scoring classes was widest in Preparatory (1.15 points, from 3.45 in OA-III to 2.30 in OA-VII), narrower in Nursery (0.62 points, from 3.35 to 2.73), and narrowest in Play Group (0.47 points, from 3.23 to 2.76). The standard deviation of class means followed the same order (Table 5). Some Preparatory classes were well above the Nursery average: OA-III (3.45) scored higher than every Nursery section.

Table 4 Composite Mean Attainment by Class, Ranked Within Year-Group

Class	Composite M
Play Group	
PG-2	3.23
PG-3	3.05
PG-1	2.76
Nursery (O-B)	
OB-II	3.35
OB-VI	3.27
OB-V	3.06
OB-IV	3.02
OB-III	2.77
OB-I	2.73
Preparatory (O-A)	
OA-III	3.45
OA-II	3.05
OA-VI	2.90
OA-V	2.75
OA-IV	2.55
OA-I	2.35
OA-VII	2.30

Note. Scale 1.00–4.00. Classes are identified by section label only.

Figure 4 Class-Level Mean Attainment Within Each Year-Group



Note. Classes are ranked from highest to lowest within each year-group. The dashed line marks the A-grade threshold (3.0).

Table 5 Spread of Class Means Within Year-Groups

Year-group	Classes	Mean of class means	SD	Range	Spread
Play Group	3	3.01	0.24	2.76–3.23	0.47
Nursery (O-B)	6	3.03	0.25	2.73–3.35	0.62
Preparatory (O-A)	7	2.76	0.41	2.30–3.45	1.15

Note. Calculated from the class means in Table 4 (unweighted). Year-group comparison of the 16 class means: Kruskal–Wallis $H(2) = 2.21$, $p = .331$, $\epsilon^2 = .15$; one-way ANOVA $F(2, 13) = 1.23$, $p = .324$.

Summary of Results

- Mean attainment was similar in Play Group (3.02) and Nursery (3.03) and lower in Preparatory (2.76). The difference was not statistically significant at class level, and the sample of classes is small.
- The lower Preparatory mean reflects a shift from the A+ band to the B band, rather than an increase in C ratings.
- At Nursery, the only year-group with a numerical domain breakdown, Writing Skills was the weakest domain and Gross Motor Skills, Discipline and Manners, and Behavior were the strongest.
- Differences between classes of the same year-group were largest in Preparatory, where the spread between the highest and lowest classes (1.15 points) exceeded one grade band.

DISCUSSION

Why Preparatory Attainment Is Lower

A sequenced curriculum should produce rising scores as children move through it, because later outcomes are intended to build on earlier ones (Biggs & Tang, 2011). The records show a plateau between Play Group and Nursery and a lower mean in Preparatory, with the change concentrated in the move from the top band to the middle band. Given the lack of statistical support at class level, this is best treated as a signal worth examining, not an established decline. Three explanations are plausible, and the records cannot choose between them.

The first is a change in rating standards rather than in children's skills. A Preparatory teacher preparing children for primary school may expect more before awarding A+, which would make part of the difference a measurement effect. The second concerns pacing. Preparatory classes introduce more formal literacy and numeracy content, and for some children this may sit beyond their current ZPD, producing the mismatch between task and support that Vygotsky's framework predicts (Vygotsky, 1978; Wood et al., 1976). The third is that the difference is real and reflects teaching time, staffing, or consistency at Preparatory level. Because the data are cross-sectional, a further possibility is that the three year-groups simply differ in composition.

Writing as a Priority for Intervention

Writing Skills was the weakest Nursery domain, and the school's own analysis noted communication and literacy as a weakness in its lowest-scoring Preparatory section. Handwriting and written expression typically develop later than spoken language and gross motor control, and they respond well to direct, structured teaching. Because the weakness appears to be instructional and not only developmental, a focused response has a reasonable chance of improving it. The relatively high Fine Motor score ($M = 3.14$) suggests that the difficulty lies less in basic hand control than in the transition from motor control to letter formation and written language, although the records cannot show this directly.

Areas of Strength

Behavior, Discipline and Manners, and Parents' Cooperation were comparatively strong in the Nursery data. These domains are not secondary to the weaker findings; they indicate conditions that can support more demanding instruction. Within Vygotsky's framework, the home is a second site of scaffolded learning, and Parents' Cooperation offers a rough indication of how far that support is present. A school with settled classroom behavior and engaged parents is better placed to introduce a stronger literacy program without losing classroom order.

Differences Between Parallel Classes

The widening spread at Preparatory level suggests that the lower average is not shared evenly across the seven sections. Some Preparatory classes scored close to or above the Nursery average, while others fell well below it. This looks more like variation in how individual classes manage scaffolding, or in how strictly individual teachers apply the rubric, than a uniform school-wide shortfall. Both possibilities are practical problems with practical remedies, because the school already contains classes that appear to be doing well. The records do not show which of the two is responsible, and the second possibility, differences in rater severity, is an important reason for treating the class rankings with caution.

Connection With the Qualitative Strand

The wider case study (Salarzai, 2026) offers some context for these patterns. Teachers there reported that they assess mainly through informal observation and personal judgment, with few structured developmental rubrics, and that they received limited training in formative assessment. They also described large classes and a heavy documentation workload, and the study found some tension between teachers' preference for flexible, play-based work and administrators' interest in measurable outcomes. These accounts bear on how the evaluation grades should be read. If standards are applied from personal judgment and not from shared descriptors, differences between classes and between year-groups may partly reflect rating practice. The qualitative strand did not examine the Preparatory pattern or the writing result directly, and both remain open for targeted follow-up.

Recommendations

The findings are descriptive and local, and the suggestions below are offered as practical options for the school and for similar preschools, not as conclusions the data prove.

1. **Introduce moderation of ratings.** Develop written descriptors and exemplar work for each grade band and domain, and have teachers cross-mark a sample of children each term. This would address the question of rating consistency that the records leave open and would show whether the A+ to B shift at Preparatory level reflects standards or skills.
2. **Strengthen early writing instruction.** Add a short, structured daily sequence of pre-writing, letter formation, and shared writing in Nursery, linked to fine motor work, and monitor Writing as a separate domain in Preparatory.
3. **Use peer observation between classes.** Pair higher- and lower-scoring Preparatory classes for structured observation and joint planning, so that effective scaffolding practice already present in the school can be shared.
4. **Keep comparable records at every stage.** Record all twelve domains numerically at every year-group, and use child-level identifiers so that the same children can be followed from Play Group to Preparatory.
5. **Support teachers in formative assessment.** Provide training in observation-based assessment and in using rubrics, and review whether the current documentation load leaves time for it.

Limitations and Future Research

Several limitations apply. The findings are specific to one private school and are not intended to generalize to other settings (Yin, 2018). The ratings are single-teacher judgments against a shared rubric, and no check

was made on how consistently different teachers apply it. The data are cross-sectional: the three year-groups are different children rated at one point in time, so differences between them are not evidence of progress or decline for any child. The records were available at class level, so inferential tests were limited to 16 class means and have low power; they cannot account for variation among children within classes. The level of detail also differs by year-group, so domain-level comparison across all three was not possible. Future work should use child-level data, with multilevel models that separate child, class, and year-group variation, follow the same children longitudinally, and check rating reliability through double marking.

CONCLUSION

This article examined the routine evaluation records of a Peshawar preschool through an outcome-based lens and a Vygotskian framework. Across 364 children in 16 classes, mean attainment was similar in Play Group and Nursery and lower in Preparatory, with the change concentrated in a move from the highest grade band to the middle one. The class-level difference was not statistically significant, so it is better read as a signal than as a finding. Writing was the weakest Nursery domain, and classroom behavior, discipline, and parental cooperation were strengths. Classes within the same year-group differed most at Preparatory level. The results indicate that a school can learn a good deal from records it already keeps, provided it is clear about what those records can and cannot support. The immediate priorities are moderated ratings, structured writing instruction, sharing practice between classes, and record keeping that allows children to be followed over time.

Declarations

Authors' Contributions

Aarif Gohar Salarzai conceived the study, curated and analyzed the data, and wrote the first draft. Syed Munir Ahmad supervised the study design and analysis and revised the manuscript critically. Both authors approved the final version.

Conflict of Interest

The authors declare no conflict of interest. The first author holds an administrative position at Frontier Children's Academy; the data were analyzed only at the aggregate, de-identified level to limit any influence on the findings. The second author is the first author's doctoral supervisor at the Institute of Education and Research, University of Peshawar.

Ethics Statement

The analysis used de-identified, institutionally aggregated records produced through the school's routine practice. Permission for secondary analysis was obtained from the institution. No child, parent, or teacher is identifiable in the results. The study falls under the ethical approval process of the first author's doctoral research programme at the Institute of Education and Research, University of Peshawar.

Data Availability

De-identified, aggregate data supporting these findings are available from the corresponding author on reasonable request, subject to the institution's data-sharing permissions.

REFERENCES

1. Bierman, K. L., Nix, R. L., Greenberg, M. T., Blair, C., & Domitrovich, C. E. (2008). Executive functions and school readiness intervention: Impact, moderation, and mediation in the Head Start REDI program. *Development and Psychopathology*, 20(3), 821–843.
<https://doi.org/10.1017/S0954579408000394>

2. Biggs, J., & Tang, C. (2011). *Teaching for quality learning at university* (4th ed.). Open University Press.
3. Black, M. M., Walker, S. P., Fernald, L. C. H., Andersen, C. T., DiGirolamo, A. M., Lu, C., McCoy, D. C., Fink, G., Shawar, Y. R., Shiffman, J., Devercelli, A. E., Wodon, Q. T., Vargas-Barón, E., & Grantham-McGregor, S. (2017). Early childhood development coming of age: Science through the life course. *The Lancet*, 389(10064), 77–90. [https://doi.org/10.1016/S0140-6736\(16\)31389-7](https://doi.org/10.1016/S0140-6736(16)31389-7)
4. Blair, C., & Raver, C. C. (2015). School readiness and self-regulation: A developmental psychobiological approach. *Annual Review of Psychology*, 66, 711–731. <https://doi.org/10.1146/annurev-psych-010814-015221>
5. Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal*, 9(2), 27–40. <https://doi.org/10.3316/QRJ0902027>
6. Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science*, 312(5782), 1900–1902. <https://doi.org/10.1126/science.1128898>
7. Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, 47(260), 583–621. <https://doi.org/10.1080/01621459.1952.10483441>
8. Mahmood, K. (2021). Practicing outcome based education in Pakistan universities: A step towards globalization. *Pakistan Journal of Humanities and Social Sciences Research*, 4(2), 189–201.
9. Merriam, S. B. (2009). *Qualitative research: A guide to design and implementation*. Jossey-Bass.
10. Norman, G. (2010). Likert scales, levels of measurement and the “laws” of statistics. *Advances in Health Sciences Education*, 15(5), 625–632. <https://doi.org/10.1007/s10459-010-9222-y>
11. Salarzai, A. G. (2026). *Preschool outcome-based education: An exploratory case study of Frontier Children’s Academy, Peshawar, Pakistan* [Unpublished doctoral dissertation]. Institute of Education and Research, University of Peshawar.
12. Snow, C. E., & Van Hemel, S. B. (Eds.). (2008). *Early childhood assessment: Why, what, and how*. National Academies Press.
13. Spady, W. G. (1994). *Outcome-based education: Critical issues and answers*. American Association of School Administrators.
14. Sullivan, G. M., & Artino, A. R., Jr. (2013). Analyzing and interpreting data from Likert-type scales. *Journal of Graduate Medical Education*, 5(4), 541–542. <https://doi.org/10.4300/JGME-5-4-18>
15. Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I., & Taggart, B. (Eds.). (2010). *Early childhood matters: Evidence from the Effective Pre-school and Primary Education project*. Routledge.
16. Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
17. Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, 17(2), 89–100. <https://doi.org/10.1111/j.1469-7610.1976.tb00381.x>
18. Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). SAGE.