SmartStart 2023
Programme Evaluation

smartstart
GIVING EVERY CHILD THE POWER TO SUCCEED.

G:ENESIS
UNLOCKING VALUE

DOCUMENT REFERENCE
SmartStart 2023 Programme Evaluation

DATE
June 2025

CONTACT INFORMATION

Genesis Analytics (Pty) Ltd South Africa
Office 3, 50 Sixth Road Hyde Park, **Johannesburg**

Genesis Analytics Ltd Kenya
9th Floor Europa Towers, Lantana Road, Westlands, **Nairobi**

Genesis Analytics Côte d'Ivoire
Royal Work Club, Immeuble Arc-En-Ciel, Avenue Chardy, **Abidjan**

Genesis Analytics United Kingdom
Office 4.06, 4th Floor, 88 Kingsway, Holborn, London, WC2B 6AA, **London**

www.genesis-analytics.com

AUTHORS
Genesis Analytics

CONTACT PERSON
Dr Sue Aitken
suea@genesis-analytics.com

Table of Contents

Figures

Tables

Abbreviations

Abbreviation	Meaning
CEF	Cognitive and Executive Functioning
ECD	Early Childhood Development
ELL	Early Language and Literacy
ELOM	Early Learning Outcomes Measure
ELP	Early Learning Programme
ENM	Early Numeracy and Mathematics
FMCVMI	Fine Motor Coordination and Visual Motor Integration
GMD	Gross Motor Development
HLE	Home Learning Environment
OLS	Ordinary Least Squares
NPO	Non-Profit Organisation
PQA	Programme Quality Assurance
SDG	Sustainable Development Goals
SEM	Structural Equation Modelling
TbF	Thrive by Five

Executive Summary

Introduction, Purpose and Scope

The global push to expand access to Early Childhood Development (ECD) services has gained considerable momentum in recent years, driven by mounting evidence that high-quality early learning experiences form a foundation for lifelong success. Governments, multilateral agencies, and civil society organisations worldwide are increasingly recognising that investing in the early years can help break cycles of poverty and inequality. The United Nations Sustainable Development Goals (SDGs) explicitly highlight early childhood care and education as key to achieving equity in learning outcomes and broader social progress.

Despite this global consensus, many countries continue to face profound challenges in scaling ECD programmes in ways that ensure both broad coverage and consistent quality. South Africa is no exception. Although significant strides have been made in policy and practice, a substantial number of children still lack access to affordable, high-quality early learning opportunities. This shortfall has serious implications not only for the children themselves, whose potential can be constrained by inadequate developmental support, but also for South Africa's broader human capital development goals and economic aspirations.

It was in response to this pressing need that SmartStart was founded in 2015. Positioned as South Africa's first national early learning delivery platform, SmartStart has been at the forefront of efforts to reach underserved communities with effective, low-cost Early Learning Programmes (ELPs). By coupling a social franchise model with an extensive partner network, SmartStart works to mobilise existing resources in under-resourced communities and transform them into sustainable ECD hubs. The aim is twofold: to bring affordable, high-quality ECD services to children who need them most, and to create pathways for local employment and social enterprise.

In the years since its inception, SmartStart has rapidly expanded its reach and built a robust ecosystem of early learning practitioners, trainers and coaches. Yet the central challenge facing any ECD intervention at scale remains whether quality can be safeguarded and meaningful child outcomes sustained as a programme grows.

This evaluation thus aims to fill critical knowledge gaps in understanding how and whether it is possible to provide quality ECD at scale, and what specific mechanisms drive any observed successes. In doing so, it engages directly with the global conversation on scaling ECD: how can programmes maintain fidelity and impact as they expand, and what lessons can be gleaned for other developing countries grappling with similar challenges?

To address these questions, this report combines empirical data with established indicators of child development and programme quality. Specifically, it draws on Early Learning Outcomes Measure (ELOM) data from a representative sample of 551 children in 325 SmartStart-supported ELPs, complemented by Programme Quality Assurance (PQA) data, practitioner interviews and a Home Learning Environment (HLE) assessment. Additionally, findings are compared with both historical data from a previous SmartStart evaluation (2018/19) and the Thrive by Five (TbF) Index 2021 which offers a nationally representative benchmark for preschool children's development in South Africa.

Through this rigorous inquiry, the report offers vital evidence on whether SmartStart's delivery model can deliver on its central promise at scale: reaching large numbers of children with high-quality early learning support. Its conclusions - highlighting the pathways by which the programme achieves impact and the areas in which further improvements are needed - speak to the broader global imperative of closing the quality gap in ECD services. As such, this evaluation not only informs SmartStart's ongoing efforts to strengthen service delivery but also contributes to the international discourse on how best to equitably expand early learning opportunities in the most vulnerable communities.

Methods

A cohort field study design was used to track the progress of a representative sample of 551 children in 325 SmartStart ELPs over an eight-month period. The ELOM was used to measure learning and development progress, allowing for a comparison of child outcomes in the 2023 sample with both the 2018 SmartStart sample and the TbF 2021 Index. Practitioner interviews provided further insight into systemic quality factors, and a HLE questionnaire was administered to parents and caregivers.

To comprehend the impact of SmartStart's programme on children's ELOM scores between baseline and endline, a combined descriptive and multivariate regression analysis approach was used. The ELOM scores of the 2023 SmartStart sample at endline

were further compared to the TbF 2021 Index. This comparison aims to evaluate how children at SmartStart's ELPs perform relative to national norms. Furthermore, the 2023 SmartStart evaluation sample was compared with the 2018/2019 sample to assess whether SmartStart had maintained improvements in ELOM performance at scale.

Multivariate regression methods were further employed to assess whether quality, as measured by SmartStart's total and domain-disaggregated PQA tool, was associated with higher ELOM scores. A detailed analysis of the fully disaggregated PQA tool was conducted to identify which specific components of the PQA observations were associated with the most significant gains in ELOM scores.

Lastly, multivariate regressions were used to explore the extent to which key components of the SmartStart support model (e.g., the number and quality of coach visits, club meetings, etc.) were associated with PQA quality measures and ELOM gains.

Findings

Key Attributes of Sampled ELPs

The children in the study were an average age of 54.4 months old at baseline and 62.3 months old at endline, with a gender distribution of 53% girls and 47% boys. Over half of the children, 52%, had participated in SmartStart for three years, while 31% had participated for two years. These children represented nine language groups, with a majority coming from isiZulu and isiXhosa speaking homes.

The ELPs represented all nine provinces, with 63% located in urban areas. A significant majority of the ELPs, almost 90%, were Early Childhood Development (ECD) centres primarily operating from home and community venues, while the remainder were childminders or play groups. Nearly half of the practitioners did not complete their Matric year, with only 33% obtaining a Matric or National Senior Certificate as their highest qualification.

Approximately one-third of ELPs were conducted in informal structures, and 54% charged a monthly fee of less than R200 (approximately $11), while 14% did not impose any fees. The ELPs most commonly (37%) maintained a ratio of 6-10 children per adult, with adults including both practitioners and assistants. There were also substantial shares of facilities with ratios of 11-15 children per adult (24%) and 1-5 children per adult (23%). A substantial proportion (77%) of ELPs, had a Green PQA rating.

Learning Outcomes

There was a notable increase in the proportion of SmartStart children 'On track' between baseline and endline, rising by 20 percentage points from 45% to 65% in just eight months. The proportion of children categorised as 'Falling far behind' decreased from 26% to 14%, a significant drop of 12 percentage points. In terms of ELOM scores, SmartStart children achieved an average ELOM gain of 11.6 points between baseline and endline, equivalent to 1.47 points per month.

Performance improvements were observed across all ELOM domains, with the largest increase in Early Numeracy and Maths (ENM), followed by Fine Motor Coordination and Visual Motor Integration (FMCVMI), and Early Language and Literacy (ELL). Smaller performance gains were observed in Cognitive and Executive Functioning (CEF) and Gross Motor Skills (GMD). Somewhat consistent with these findings, the proportion of children 'On track' at endline was higher in ELL (65%) and ENM (62%), while GMD (51%), CEF (47%), and FMCVMI (44%) reported lower proportions.

The greatest baseline-to-endline gains were observed among ELPs in low-income quintiles[1]. While ELPs in Quintile 1 and Quintile 2-3 experienced average Total ELOM score increases of 11 and 13 points respectively, ELOM scores of ELPs in Quintile 4-5 rose by only nine points. Consequently, the performance gap between children attending ELPs in the highest and lowest income quintiles significantly decreased, from 25 points to six points.

The SmartStart sample outperformed the TbF 2021 Index across all domains, even after adjusting for differences between the samples. The average total ELOM score for SmartStart children at endline was 11 points higher than the TbF Index, with a significantly higher proportion of SmartStart children being 'On track' at endline (65% vs 46%). The disparity was most pronounced in ENM (+24.6 percentage points), followed by FMCVMI (+11.8 percentage points) and ELL (+9.8 percentage points). The regression analysis, which controlled for sample differences, revealed that the SmartStart sample outperformed the

---

[1] In South Africa, public schools are categorised into five quintiles (quintile 1 to quintile 5) based on the relative socio-economic status of the communities they serve. This classification is determined by national poverty indicators such as income levels, unemployment rates, and general living conditions within the school's catchment area. Quintile 1 schools serve the poorest communities, while quintile 5 schools are located in the wealthiest areas. The quintile system is primarily used to guide the allocation of government funding, with lower quintile schools receiving greater financial support per learner. Most schools in quintiles 1-3 are designated as no fee schools. For ELPs, the quintile of the nearest primary school is used as a proxy.

TbF sample across all domains, although the effect was not statistically significant for the CEF domain.

In the 2018 evaluation, the proportion of children 'On track' based on the total ELOM score showed a greater baseline-to-endline improvement (26.9 percentage points) compared to 2023 (20 percentage points). However, the proportion of children 'On track' at endline was higher in the 2023 sample (65%) compared to the 2018 sample (59%).

To understand the reasons behind these differences in outcomes between the 2018 and 2023 evaluations, a regression analysis was conducted to control for sample differences in PQA scores and child/ELP characteristics. For example, the 2018 sample was smaller, all the ELPs were rated 'Green' on the PQA status and most were playgroups and day mothers. In contrast, the 2023 evaluation used representative samples across all PQA categories (i.e. Red, Orange and Green).

Even after accounting for these factors, the differences remained statistically significant. This could be due to omitted variable bias (some variables present in the 2018 evaluation were not available in the 2023 evaluation) or genuine differences in performance. These findings should thus be interpreted with this limitation in mind.

Programme Quality and ELOM Gains

The descriptive analysis revealed that the majority (77%) of ELPs had a Green PQA rating, 10% had an Orange and 3% a Red PQA rating. The remaining 10% did not have a PQA rating. On average, the scores of items within each domain were high, with many items nearing the maximum score of 2.

Several components of the SmartStart programme's design and practice, as measured by the 34 individual PQA items, were significantly associated with ELOM gains. ELPs that consistently engaged in at least eight out of ten observed SmartStart activities showed an average ELOM gain of 4.57 more points compared to those in ELPs that observed fewer SmartStart activities, suggesting a substantial positive learning impact of faithful implementation of the SmartStart routine. Children in ELPs allowing ample time for recall and reflection also experienced an average ELOM gain of 3.22 points, while those encouraged to learn at their own pace gained 4.22 points. In contrast, children in environments where adults imposed solutions on conflicts or didn't listen to children's ideas experienced significantly lower gains in overall ELOM, as well as in the ELL domain.

While several individual PQA items were associated with ELOM gains, the total PQA score for an ELP, the Green/Orange/Red ratings, and the scores on the six PQA subdomains did not show a significant association with gains in ELOM scores. This is similar to other studies which have found that the strength of the relationship between quality assurance tools, designed to monitor good practice and target support, and child outcomes tend to be modest.[2] [3] This finding suggests that the current PQA scoring system and categorisations may not accurately reflect programme quality in terms of ELOM gains.

The Role of SmartStart Support

The descriptive analysis produced positive findings for the different support mechanisms orchestrated by SmartStart.

- SmartStart coaches: 86% of practitioners found coach visits to be beneficial, rating them as useful (20%), very useful (28%), or extremely useful (38%). Almost 80% of practitioners reported receiving constructive feedback from the coach after a visit. Frequency of coach visits varied, with an average of 4.92 visits per year. Just under a third of ELPs were visited once or twice a year (29%), followed by 3-4 times per year (23%) and 5-6 times per year (16%).

- SmartStart clubs: A substantial 93% of practitioners confirmed their participation in a Club, and of these 95% attended club meetings, indicating widespread engagement with this collaborative component of the SmartStart programme. Additionally, 91% of club members expressed the value of club participation as either useful (25%), very useful (33%), or extremely useful (33%).

- Business skills training: 69% of practitioners reported participating in SmartStart's comprehensive two-day business skills workshop. The feedback on the training was overwhelmingly positive, with only 4% of attendees finding the training to be not at all useful or just slightly useful.

- Practitioner satisfaction: Practitioners showed strong support for the SmartStart programme. When asked to assess their likelihood of recommending SmartStart to another ELP, only 15% provided a likelihood rating of less than 7.

---

[2] Ulferts, H., Wolf, K. M., & Anders, Y. (2019). Impact of process quality in early childhood education and care on academic outcomes: Longitudinal meta-analysis. Child Development, 90(5), 1474-1489.
[3] Burchinal, P., Kainz, K., Cai, K., Tout, K., Zaslow, M., Martinez-Beck, I., & Rathgeb, C. (2010). Early care and education quality and child outcomes. Child Trends.

The frequency of coach visits was found to have measurable associations with PQA scores and ELOM gains. ELPs that received one or fewer visits annually had PQA scores 4.48 points lower than those with two or more visits per year. The frequency of coach visits was also associated with ELOM gains. However, this was only statistically significant when comparing ELPs receiving 0-1 visits annually to those receiving 11-24 visits, with the latter group achieving an average of 6.14 additional ELOM points.

The quality of coach visits mattered as well, with satisfaction regarding coaching effectiveness significantly linked to ELOM improvements. ELPs where coaching visits were perceived as "not at all useful" experienced an average increase of 8.37, while those rating visits as "extremely useful" saw an increase of 12.45 points. Coaches were more likely to receive higher "usefulness" ratings if they provided practical techniques for stimulating environments and assisted in child assessments and planning.

Although club meetings were well-attended and highly valued by practitioners, neither the frequency of club meetings nor their perceived usefulness to practitioners showed measurable associations with improvements in ELOM scores or PQA scores. This finding was not unexpected, given that the primary focus of clubs is mutual support and peer engagement rather than directly influencing programme quality. However, it is important to note that the current evaluation was only able to test for quantitative, statistical associations of club participation with elements of the ELOM and the PQA, and there may be benefits of club participation which could not be captured quantitatively by the measures used in this study.

Recommendations

Based on the evaluation's results, in order for SmartStart to deepen its impact and continue to maintain programme quality at scale, it is recommended that SmartStart:

1. Capitalise on ELOM and ENM Successes

SmartStart has demonstrated substantial ELOM progress, with a 20% increase in children meeting the standard and outperforming national benchmarks. Results were especially notable in ENM, with ENM accounting for a 64% share of programme gain, and scores 24.6 percentage points above the national average. These successes in ENM appear to be particularly correlated with ELP environments that also adopt specific practices, such as:

- Allocating significant time for free play (45-55 minutes).

- Extending large group activities beyond 10 minutes.
- Employing effective language development methods.
- Adjusting the learning pace to suit individual children.
- Using storybooks to familiarise children with reading materials.

2. Refine the PQA Tool Usage

This evaluation suggests that many of the observational elements of the PQA are significantly associated with ELOM gains. However, the current way in which the PQA tool's observational items are being aggregated and used by SmartStart for accreditation does not align well with actual ELOM gains. Instead of using total PQA scores as a benchmark, SmartStart should focus on the individual PQA items that directly contribute to ELOM gains. This includes:

- Prioritising observations related to appropriate play and learning pacing, reflection time for children, and the number of SmartStart activities observed.
- Enhancing SmartStart monitoring, feedback, and quality assurance based on these key items.
- Refining and improving the observer rating and feedback process around the key PQA observations that appear to be associated with ELOM gains.

3. Emphasise Quality and Quantity of Coach Visits

The analysis confirms a significant association between monthly coach visits and ELOM gains. SmartStart should focus on exploring the optimal number of coach visits to achieve sufficient ELOM gains at scale, as well as ensuring that coaches focus on the elements where their support appears to be the most valued, namely in providing practical techniques for creating stimulating environments and supporting child assessments and planning.

4. Further Explore the Role of Club Attendance

Club attendance was highly valued and widespread, and club attendance may perform many important functions around peer support and retention that were not explicitly measured in this evaluation. However, there were no direct significant associations between club attendance variables, PQA scores and ELOM gains. It is therefore recommended that SmartStart should continue to qualitatively unpack the benefit of club attendance and document how club attendance is supporting SmartStart in ways not directly related to the ELOM and/or the PQA.

Conclusion

The evaluation results demonstrate that SmartStart's national early learning delivery platform, focusing on home- and community-based settings in low-income communities, can significantly improve early childhood outcomes, even at scale. The evaluation, while not covering all dimensions of the SmartStart architecture, provides critical insights into the association between programme practices, ELP quality, and child outcomes. The SmartStart sample outperformed the TbF national benchmark across all ELOM domains, even after adjusting for sample differences.

The findings on the mechanisms that lead to observed gains in child outcomes underscore the importance of regular, high-quality coach visits and adherence to the SmartStart routine in building quality early learning opportunities. This provides critical validation that features of SmartStart's programme design are effectively improving outcomes for children. Furthermore, these impactful programme features were found to be simple practices, underscoring that the value of a programme is not necessarily determined by expensive facilities and resources, but rather by what the programme offers and how it is supported.

In summary, the 2023 programme evaluation of SmartStart provides compelling evidence that:

1. A carefully designed delivery system for home- and community-based ELPs can improve early childhood outcomes, even when implemented at scale.
2. In under-resourced contexts, ELPs with the crucial ingredients of quality in place can close the opportunity gap for children in low-income communities compared to their peers from higher income backgrounds.
3. Given that this delivery method can be implemented quickly and affordably, it should be a priority for funding solutions and should inform a more contextually-appropriate regulatory framework.

1 | Introduction

The global conversation on expanding access to ECD services is growing ever more urgent, given the strong evidence linking high-quality early learning to a range of improved life outcomes. Yet many countries, including South Africa, still grapple with the dual challenge of rapidly scaling programmes without compromising quality. In this context, SmartStart stands out as an innovative example of how home- and community-based ECD solutions might be delivered at scale. By mobilising local resources and partners through a social franchise model, SmartStart aims to address persistent gaps in ECD provision and help break cycles of inequality. This evaluation report situates SmartStart's efforts within the broader global debate on how to expand ECD services equitably, assessing whether - amid significant growth - SmartStart has preserved the high-quality practices necessary for meaningful child outcomes.

Background

In South Africa, early learning opportunities are a critical yet often under-resourced component of child development. South Africa has nearly seven million children under six years of age, with 70 percent of these young children living in households below the official "upper-bound" poverty line[4]. However, around 1.1 million children aged 3-5 years are not enrolled in any kind of ELP.[5] There evidently remains a need for considerable expansion of ELPs to achieve the National Development Plan goal of universal access to two years of early childhood development by 2030.[6] [7]

---

[4] Hall K, Almeleh C, Giese S, Mphaphuli E, Slemming W, Mathys R, Droomer L, Proudlock P, Kotze J, and Sadan M.(2024) South African Early Childhood Review 2024. Cape Town: Children's Institute University of Cape Town and Ilifa Labantwana. Available at: https://ilifalabantwana.co.za/wp-content/uploads/2024/07/SA-early-childhood-review-2024-FINAL.pdf

[5] Slemming, W., Biersteker, L. & Lake, L. (2024) South African Child Gauge 2024. Cape Town: Children's Institute, University of Cape Town. Available at:https://ci.uct.ac.za/sites/default/files/media/documents/ci_uct_ac_za/533/full-publication-child-gauge-2024.pdf

[6] Republic of South Africa.(2013). National Development Plan 2030: Our Future-make it work. Pretoria: National Planning Commission. Available at: https://www.gov.za/sites/default/files/gcis_document/201409/ndp-2030-our-future-make-it-workr.pdf

[7] Please note that the statistics reported in this paragraph are from nationally representative surveys conducted in 2022.

Children who have access to structured learning opportunities are more likely to succeed later in life. Children with access to at least two years of high-quality early learning are more likely to acquire the cognitive, social and emotional skills to manage the transition into formal schooling.[8] Access to quality early learning opportunities has also been found to be associated with longer-term gains, enhancing children's pathways through school and into adulthood.[9] Although all children benefit from high-quality ELPs, those from disadvantaged backgrounds experience the most significant developmental gains.[10] This emphasises the critical need for quality ELPs in the highly unequal context of South Africa.

Ensuring high-quality provision is essential for children to benefit from ELPs.[11] Progress on child outcomes requires learning programmes to be implemented with both structural and process quality, with process quality - what children actually experience in a programme - shown to have a greater influence on child development than structural variables.[12][13][14][15][16] Furthermore, increasing evidence suggests that specific elements of ELPs are more strongly linked to school readiness. These include extended time in small group

---

[8] Yoshikawa, H, Weiland, C, Brooks-Gunn, J, Burchinal, MR, Espinosa, LM, et al. (2013). Investing in our future: The evidence base on preschool education. Society for Research in Child Development. Available at: https://www.fcd-us.org/wp-content/uploads/2016/04/Evidence-Base-on-Preschool-Education-FINAL.pdf

[9] Nores, M, & Barnett, WS. (2010). Benefits of early childhood interventions across the world: (Under) Investing in the very young. Economics of Education Review, 29(2):271-282.

[10] Yoshikawa, H, Weiland, C, Brooks-Gunn, J, Burchinal, MR, Espinosa, LM, et al. (2013). Investing in our future: The evidence base on preschool education. Society for Research in Child Development. Available at: https://www.fcd-us.org/wp-content/uploads/2016/04/Evidence-Base-on-Preschool-Education-FINAL.pdf

[11] United Nations Children's Fund. (2019). A World Ready to Learn: Prioritizing quality early childhood education - Advocacy brief. New York: UNICEF. Available at: https://www.unicef.org/media/51751/file/A-world-ready-to-learn-advocacy-brief-2019-eng.pdf

[12] Sabol TJ, Soliday Hong Sl, Pianta RC, Burchinal MR. (2013). Can rating pre-K programs predict children's learning? Science. 341(6148):845-846.

[13] Rao N, Sun J, Wong JMS, Weekes B, Ip P, Shaeffer S, Lee D. (2014). Early Childhood Development and Cognitive Development in Developing Countries: A rigorous literature review. Department for International Development.

[14] Torii K, Fox S, Cloney D. (2017). Quality is Key in Early Childhood Education in Australia. (Paper No. 01/2017). Melbourne: Mitchell Institute.

[15] Diamond KE, Justice LM, Siegler RS, Snyder PA. (2013). Synthesis of IES Research on Early Intervention and Early Childhood Education. (NCSER 2013-3001). Washington, DC: National Center for Special Education Research, Institute of Education Sciences, U.S. Department of Education.

[16] Maldonado-Carreño C, Yoshikawa H, Escallón E, Ponguta LA, Nieto AM, Kagan SL, Aragon CA. (2022). Measuring the quality of early childhood education: Associations with children's development from a national study with the IMCEIC tool in Colombia. Child Development. 93(1):254-268.

activities[17] and positive interactions between children and between adults and children[18]. Learning through play is also widely recognised as one of the most effective methods for delivering high-quality early learning programmes[19] [20]. These findings highlight the need to understand the links between specific programme elements and child outcomes, particularly in low-income, under-resourced contexts such as South Africa.

Founded in 2015, SmartStart is a comprehensive early learning delivery platform designed to bridge the gap in access to early learning in South Africa. By focusing on home- and community-based settings in low-income communities, SmartStart aims to improve early childhood outcomes through an accessible, cost-effective model that leverages existing community resources. Since SmartStart's last evaluation in 2018, the SmartStart network has more than doubled, raising the need to assess whether programme quality and impact have been maintained at scale. Consequently, it is necessary to investigate whether SmartStart has been able to maintain its programme quality and impact at scale, and, if so, what mechanisms are driving the observed impact.

This report sets out to assess the degree to which the SmartStart programme was associated with improved outcomes for participating children in 2023 and whether potential mechanisms of change can be identified. In terms of the mechanisms of change, associations will be explored between child outcomes and a) SmartStart's PQA tool, and b) components of SmartStart's platform design and support.

The evaluation draws on data from baseline and endline ELOM assessments conducted among a sample of SmartStart franchisees. These data are complemented by PQA assessments, interviews with ELP practitioners, interviews with household caregivers, data from the 2018/2019 SmartStart evaluation, and data from the TbF 2021 Index.

---

[17] McDaniel M, Townley-Flores C, Sulik M, Obradovic J. (2022). Widely used measures of classroom quality are largely unrelated to preschool skill development. Early Childhood Research Quarterly. 59:243-253.
[18] Burchinal MR, Garber K, Foster T, Bartsch-Hines M, Franco X, Peisner-Feinberg E. (2021). Relating early care and education quality to preschool outcomes: The same or different models for different outcomes? Early Childhood Research Quarterly. 55:35-51.
[19] Zosh JN, Hirsh-Pasek K, Hopkins EJ, Jensen H, Liu C, Neale D, Whitebread D. (2018). Accessing the inaccessible: Redefining play as a spectrum. Frontiers in Psychology. 9:1124.
[20] Zosh JN, Hopkins EJ, Jensen H, Liu C, Neale D, Hirsh-Pasek K, Whitebread D. (2017). Learning through Play: A review of the evidence. Billund, Denmark: LEGO Fonden.

SmartStart

Early learning opportunities for children, before they commence formal schooling, are strong determinants of later life outcomes and success[21]. However, in South Africa, a substantial number of children lack access to these pivotal ELPs, particularly in low-income communities. To address this issue, SmartStart was founded in 2015 with a mission to broaden access to affordable, high-quality early learning programmes.

SmartStart - South Africa's first national early learning delivery platform - leverages a three-tiered social franchise model. The organisation works with regional branches and Non-Profit Organisation (NPO) partners, known as Franchisors, to recruit, train, license and support a network of early learning practitioners to deliver the same, high-quality learning programme targeted at three- to five-year olds. The early learning practitioners are supervised by coaches associated with each Franchisor. SmartStart, the orchestrator of this platform, ensures that support mechanisms are effectively maintained across the various levels of the model.

Coaches, clubs, and the PQA tool constitute three significant mechanisms employed by SmartStart to ensure quality implementation across the three-tiered network:

1. Implementing partners and branches employ coaches who oversee and support a group of ELPs. These coaches are trained in SmartStart's programme methodology, broader aspects of quality and the administration of SmartStart's PQA tool.
2. SmartStart's PQA tool is a formative observational assessment tool designed to check for the presence of quality components, incentivise good practice, and identify areas requiring targeted support and capacity building.
3. Clubs are local communities of practice where practitioners meet monthly to exchange ideas and offer personal and professional support. While clubs share good practices and sometimes offer programme support, their primary focus is fostering a sense of collective endeavour, peer support, and recognition of effort and progression.

This delivery system is designed to be effective at scale. The objective is to implement affordable home- and community-based ELPs that leverage the strengths and assets of low-income communities. To do this, the delivery system draws on a national network of

---

[21] Nores, M, & Barnett, WS. (2010). Benefits of early childhood interventions across the world: (Under) Investing in the very young. Economics of Education Review, 29(2):271-282.

implementing partners to improve children's readiness for learning, school performance, and overall life success. Simultaneously, it aims to create employment and social enterprise opportunities and stimulate economic activity in the most under-resourced communities in South Africa.

Purpose of the Evaluation

SmartStart's 2018 child outcomes evaluation provided strong indications that the SmartStart delivery platform was effective in improving early learning and development outcomes for children. However, by 2023, the SmartStart network had more than doubled to 8,629 ELPs and 66,051 children, and by 2025 it had grown further to nearly 13,000 ELPs and 120,000 children. It is thus necessary to investigate whether SmartStart has been able to maintain its programme quality and impact at scale, and, if so, what mechanisms are driving this impact.

Thus, in order to understand whether SmartStart's home- and community-based approach can maintain impact across a much larger network of programmes, this evaluation report seeks to answer five key questions:

1. Are quality SmartStart programmes on average ensuring that participating children achieve better outcomes than those achieved on average by other children in South Africa?

2. How do gains in ELOM scores over time compare with gains seen during SmartStart's 2018/2019 outcome evaluation?

3. Is quality, as measured by SmartStart's PQA tool and process, associated with higher ELOM scores?

4. Which of the various components of the PQA are associated with the biggest gains in ELOM scores?

5. To what extent is SmartStart support associated with the PQA quality measures? Which aspects of support are most beneficial?

The research questions thus focus on the degree to which SmartStart's social franchise model is improving outcomes for children and whether potential mechanisms of change can be identified. In terms of the latter area of inquiry, associations between child outcomes

and a) SmartStart's PQA tool, and b) components of SmartStart's platform design and support will be investigated.

2 | Methodology

This section outlines the methodological approach used to address the research questions, particularly in terms of the sampling process, data, and methods of analysis.

Sample

Baseline

Early Learning Programmes

A sample of 404 ELPs was drawn from a comprehensive pool of 8,993 SmartStart franchisees. The procedure ensured both randomisation and logistical efficiency through a two-stage pairing process, where ELPs were geographically paired to facilitate assessors' ability to visit two sites per day:

1. 202 ELPs were randomly selected, with probability proportional to size (where size refers to the number of SmartStart ELPs) at the province level, from the sampling frame.
2. From the neighbourhoods surrounding each of these 202 selected ELPs, another 202 ELPs were purposefully selected.

In theory, this sampling approach should produce a representative sample. However, obtaining a representative sample posed certain challenges:

1. One hundred and seventy-two of all franchisees lacked a discernible address, leading to their exclusion.
2. Areas with only a single franchisee were omitted from the sample due to cost constraints.
3. There were unexpected difficulties in locating ELPs during fieldwork, particularly smaller, informal ELPs in townships.
4. Missing or incomplete addresses and contact details, such as phone numbers, in the list of ELPs.

Initially, assessors attempted to contact ELPs through phone calls or SmartStart coaches. If no phone number was available, assessors physically went to the provided location. Where they struggled to reach ELPs, assessors turned to coaches for help.

If they could not find an ELP following this approach, the assessors replaced it with the next randomly selected ELP in the area.

Children

Two children aged between 50 and 61 months were randomly selected at baseline from each of the 404 franchisees, amounting to 808 children. The decision to select two children per franchisee site was influenced by the relatively high attrition rate in the previous outcome evaluation and the need to ensure that at least one child per franchisee remained in the study for both baseline and endline assessments. During fieldwork, efforts were made to achieve gender balance whenever possible, ensuring that the randomly selected children ideally represented a balanced gender distribution (one boy, one girl).

Since the evaluation aimed to explore the relationship between an ELP's PQA score and the gains made by children in that ELP, a large sample of ELPs was necessary to capture variability in programme quality and its impact on child development outcomes. For this reason, only two children per ELP were sampled, to maximise between ELP variation.

Endline

Several factors affected the sample between baseline and endline assessments, with certain ELPs having to be excluded from the study for various reasons outlined in Table 1 below. Consequently, a total sample of 325 ELPs was used for the analysis. Across these ELPs, 551 children had valid ELOM assessments at baseline and endline.

Table 1: ELPs excluded from the study

Reason for ELP Exclusion	Count	Note
Inactive before the commencement of endline	19	Identified by SmartStart
No longer affiliated with SmartStart or inactive during the fieldwork phase	10	Identified by fieldworkers
Two sampled children did not attend the ELP from baseline to endline	27	Dropped out of the ELP
Refusal to participate by SmartStarters	5	Principals did not want to participate
Inaccessible	16	Two children at ELP unreachable at deadline
Total	**77**	

Data

Data Collection Tools

Three data collection tools were employed at baseline and endline, namely the child assessment, parent/caregiver interview and practitioner interview.

ELOM Assessment

The ELOM is an assessment tool used to evaluate and measure the developmental progress and learning outcomes of children aged between 4-6 years. It encompasses five domains of early childhood development including Gross Motor Development (GMD), Fine Motor Coordination and Visual Motor Integration (FMCVMI), Early Numeracy and Mathematics (ENM), Cognitive and Executive Functioning (CEF), and Early Language and Literacy (ELL). ELOM assessments help ELPs, programme designers, researchers, and policymakers understand children's developmental trajectories, identify areas where children may need additional support, and track the effectiveness of ELPs and interventions.

Parent/Caregiver Interview

Telephonic interviews were carried out with the caregivers of the sampled children using the HLE questionnaire. This questionnaire is designed to measure home aspects associated with early language, numeracy and cognitive functioning for children aged 2-7 years old. Specifically, the HLE questionnaire measures the availability of early learning resources (e.g. books, toys, and other utensils used for play), early learning activities (e.g. reading, telling stories, counting, and singing), and caregiver time spent with the child.

Practitioner Interview

Practitioners were interviewed to evaluate the structural and systemic quality of the ELP. These interviews aimed to gather insights into the SmartStarters' qualification levels, the physical setting, teacher-child ratios, group sizes, availability of learning materials, considerations for health and nutrition, attendance records, and the operational schedule (i.e. the number of days an ELP is open per week).

Complementary Data

In addition to the data collected, PQA and 2018/2019 SmartStart evaluation data were used to assess quality measures and their changes over time.

Programme Quality Assurance Data

The PQA data provides a measurement of programme quality for each of the sampled ELPs. The PQA tool measures six subdomains with each subdomain being the aggregate score of the items falling within that subdomain. The six subdomains are:

1. A stimulating and adequately resourced learning environment
2. Consistent use of the SmartStart routine
3. A stable and nurturing environment where children feel safe and loved
4. Positive and plentiful adult-child interactions which encourage rich use of language
5. Opportunities for child-directed, open-ended play, supported and extended by adults
6. Interactive storytelling that introduces children to new language and learning

There are 34 individual observational items falling across each subdomain, which are rated as either Good (a score of 2), Basic (a score of 1) or Inadequate (a score of 0).

Following the scoring of the 34 observation items, the 34 items on the PQA are summed, and cut-offs are observed, which correspond to a Green PQA rating (meets SmartStart quality status benchmarks), Orange PQA rating (is on the borderline of meeting the quality benchmark), or Red PQA rating (falls far below quality benchmarks).

2018 SmartStart Data

Data from a previous SmartStart evaluation was used as a benchmark for this analysis. The 2018 SmartStart data is a smaller sample (N=199) that, unlike the current evaluation, was purposefully sampled and therefore not representative of SmartStart's ELPs. Moreover, the 2018 evaluation sample excluded ELPs that did not have a Green PQA rating and consisted only of playgroups and day mothers, while the 2023 sample included ECD centres with all PQA ratings.[22] Other key differences included the number of years children had been part of SmartStart[23] and whether fees were charged by the facility[24]. However, to a certain extent, we were able to adjust or control for these differences through regression analyses.

---

[22] The 2023 sample was 89% ECD centres, 7% play groups, and 4% 'full week' centres.
[23] In 2023 52% of children had been in SmartStart for three years, 31% for two years and 17% for one year, while for the 2018 sample it was 30% of children in the programme for three years.
[24] In 2023 14% of children attended no fee facilities while in 2018 52% of children attended no fee facilities.

While the 2018 evaluation contains fewer background variables related to programme characteristics, the ELOM data, HLE data, and practitioner interview data were otherwise collected under the same procedure as the 2023 evaluation. Baseline data were collected in February-March 2018 and the endline data were collected in October 2018.

Thrive by Five Index 2021 Data

The 2021 Thrive by Five Index is a nationally representative survey of 5,222 South African preschool children. It contains ELOM assessment data on each child and background information on the ELP the child attends, making the dataset ideal for generating benchmark statistics against which to compare SmartStart.

Analysis Approach

The five core research questions have been answered using the data sources listed in the previous section. Below we list the five research questions and the approach taken to answer them:

1. **Are quality SmartStart programmes on average ensuring that participating children achieve better outcomes than those achieved on average by other children in South Africa?**

To address this evaluation question, we presented three different approaches. The first descriptive approach was to show the aggregate improvements in average ELOM scores and the change in the proportion of children 'On track' across domains. This approach allowed us to assess the proportion of children 'On track', given the anticipated ELOM norms for children in a similar age bracket. To contextualise the SmartStart sample to other children in South Africa, we compared the endline ELOM scores and proportion of children 'On track' to the nationally representative Thrive by Five sample. We initially presented a descriptive comparison between the two samples and then employed Ordinary Least Squares (OLS) regression analysis to compare ELOM scores while controlling for sample differences. Control variables included age, gender, years in the programme, fee amount, practitioner education, class size, whether the class has books, and child's height-for-age.

2. **How do gains in ELOM scores over time compare with gains seen during SmartStart's 2018/2019 outcome evaluation?**

Similarly to the first research question, we initially approached the second research question with descriptive statistics, comparing the proportion of children 'Achieving the

ELOM standard' between the 2018 and 2023 samples. Next, we used a statistical approach where we controlled for differences between the two samples by using multivariate Ordinary Least Squares (OLS) regression analysis[25] where the outcome variable was the change in ELOM score between baseline and endline. Control variables included gender, age, child-adult ratio, practitioner education, whether the facility charged fees, height-for-age, child years in the programme, and the facility's PQA score.

3. **Is quality, as measured by SmartStart's PQA tool and process, associated with higher ELOM scores?**

4. **Which of the various components of the PQA are associated with the biggest gains in ELOM scores?**

We used OLS regression to explore these questions. All coefficients reported for multivariate models controlled for a range of ELP and child-specific factors, including baseline ELOM score, height-for-age at endline, child age at endline, gender, attendance, practitioner-to-child ratio, urban vs. rural metropolitan status, and the number of years the child has been attending the ELP. Additionally, we incorporated controls for the HLE of the child, encompassing factors such as the level of education of the child's primary caregiver, a sum of incidences in the household of various activities during the past week (e.g. reading books, telling stories, singing songs, outdoor activities, playing, naming things, counting, drawing), and the duration of time the primary caregiver spends playing with the child on both weekends and weekdays.

5. **To what extent is SmartStart support associated with PQA quality measures? Which aspects of support are most beneficial?**

Structural equation modelling (SEM) for path analysis[26] demonstrated that none of the PQA domains had a significant effect on ELOM gains. As a result, a SEM was ineffective in

---

[25] Multivariate Ordinary Least Squares (OLS) regression is a statistical method used to model the relationship between one continuous dependent variable and two or more independent (explanatory) variables. The technique estimates the linear associations between the dependent variable and each independent variable, while holding the effects of the other variables constant.
OLS regression assumes a linear relationship, minimises the sum of the squared differences between the observed and predicted values (residuals), and provides coefficient estimates that can be interpreted as the average change in the dependent variable associated with a one-unit change in an independent variable, all else being equal.
[26] Structural Equation Modelling (SEM) for path analysis is a statistical technique used to examine hypothesised causal relationships among a set of observed variables. It extends multiple regression by allowing for the estimation of multiple dependent relationships

exploring the contribution of aspects of SmartStart support to ELOM gains via the pathway of PQA improvements. Additionally, because the 34 individual PQA items were not quantified on a metric scale, they were inappropriate for further SEM analysis. As a result, multivariate OLS and logistic regressions were primarily used to explore the relationship between SmartStart support, PQA and ELOM gains.

---

simultaneously, including both direct and indirect effects. In path analysis, SEM is used to test the strength and direction of relationships within a predefined model.
SEM for path analysis enables researchers to assess whether the data support the proposed model and to explore how mediating variables transmit the effects of one variable onto another. It provides a rigorous framework for testing theoretical models and understanding the underlying mechanisms driving educational outcomes.

3 | Descriptive Statistics

Children

A crucial aspect of the SmartStart evaluation involves learning outcomes achieved by children who attend the facilities. The ELOM tool was used to measure whether children were developmentally on track for their age group, whether ECD programmes were effective in preparing children for entry into the Foundation Phase of their schooling career, and to identify areas for improvement.

In this comprehensive evaluation of SmartStart programmes in South Africa, a cohort of 551 children underwent ELOM assessments at both baseline and endline, depicting various developmental trajectories. The evaluation spanned approximately eight months. At baseline, these learners were aged between 42-65 months, with an average age of 54.4 months. At endline, they were aged between 49-73 months, averaging 62.3 months — indicating an average age increase of 7.9 months observed between the baseline and endline ELOM assessments. The sample offered a relatively balanced gender split, with 53% of the participants being female and 47% being male. Furthermore, in 2023, 52% of children had participated in SmartStart for three years, 31% for two years, and 17% for one year.

Although SmartStart is not primarily focused on health-related issues such as stunting, the ELOM does include height-for-age measurements. The results indicate a significant decrease in stunting levels from baseline to endline, demonstrating an improvement in children's growth and development over the programme's duration. Initial height-for-age measurements indicated that 10% of the children experienced stunted growth, while the endline statistics painted a more optimistic picture, showing a decline in the stunting rate to 7%.

Beyond age, gender, and physical development, the study captured the diverse languages of South Africa. Each ELOM assessment was conducted in the child's home language and was largely reflective and representative of the multilingual fabric of South Africa. Figure 1, below, shows the sample comprising a wide range of home languages and depicts isiZulu (31%) and isiXhosa (26%) as the most prevalent home languages, which match their status as the most widely spoken national languages. Sesotho contributed to just over 12% of the linguistic landscape, while other languages comprised between 1%-9% of the total sample.

Figure 1: Child's home language
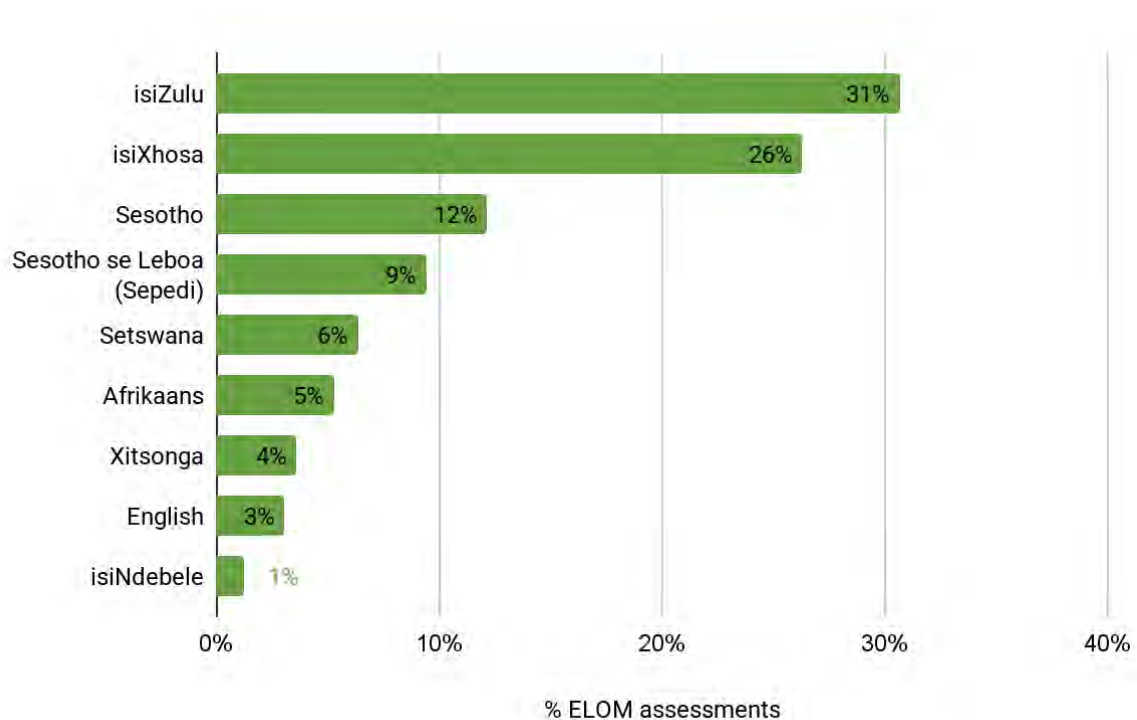
Early Learning Programmes

The sample of 325 ELPs was representative of SmartStart facilities across the country (Figure 2). Gauteng hosted the majority (34%) of the sampled ELPs, which was over double the prevalence of ELPs in the Eastern Cape (15%). KwaZulu-Natal and the Western Cape had similar shares, each contributing 14% and 13%, respectively, while the remaining provinces each comprised between 2%-9% of the sample. We categorised each ELP as either urban or rural, utilising the 2011 Census Enumeration Areas. The distribution leans towards urban settings, with 63% of ELPs falling into this category, while 37% of ELPs are situated in rural areas.

Figure 2: Distribution of sampled ELPs per province
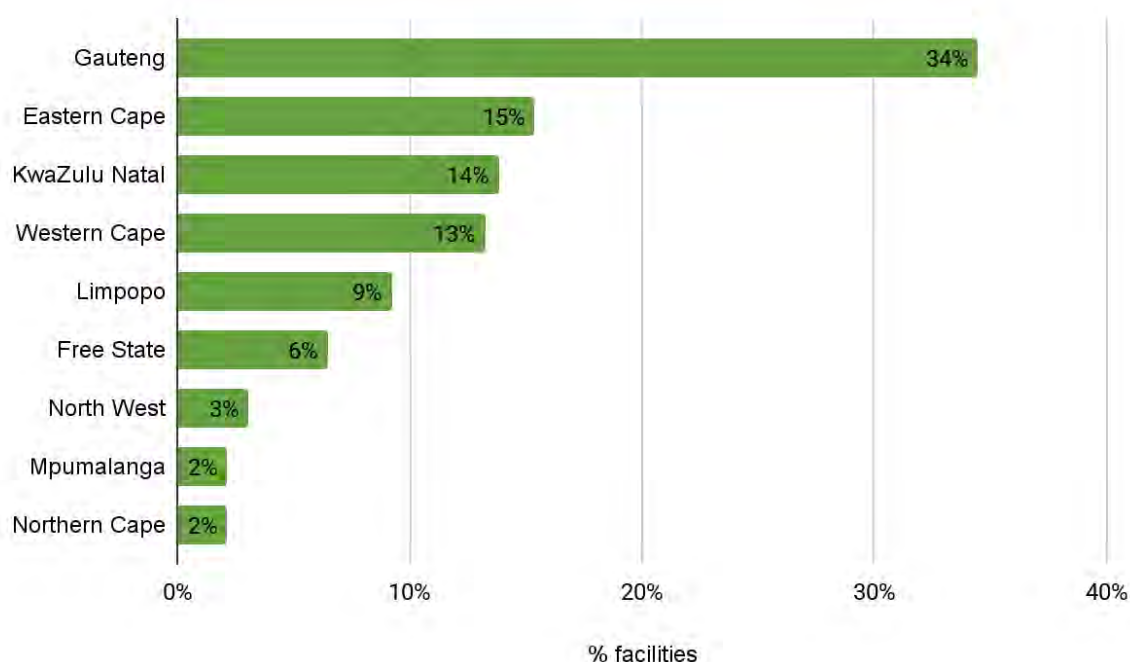
ELPs were further categorised into one of four types: SmartStart ECDs (full-time programmes with more than six children), ECD Centres (classes taught by SmartStart teachers in an ECD centre), Play Groups (part-time programmes with up to 12 children), and Full Week Day Mothers (full-time programme with six or fewer children) (Figure 3). SmartStart ECDs and ECD Centres accounted for the majority of ELP sites at 46% and 42%, respectively. Play Groups and Full Week Day Mothers contributed 7% and 4%, respectively, to the overall composition.

Figure 3: ELP type
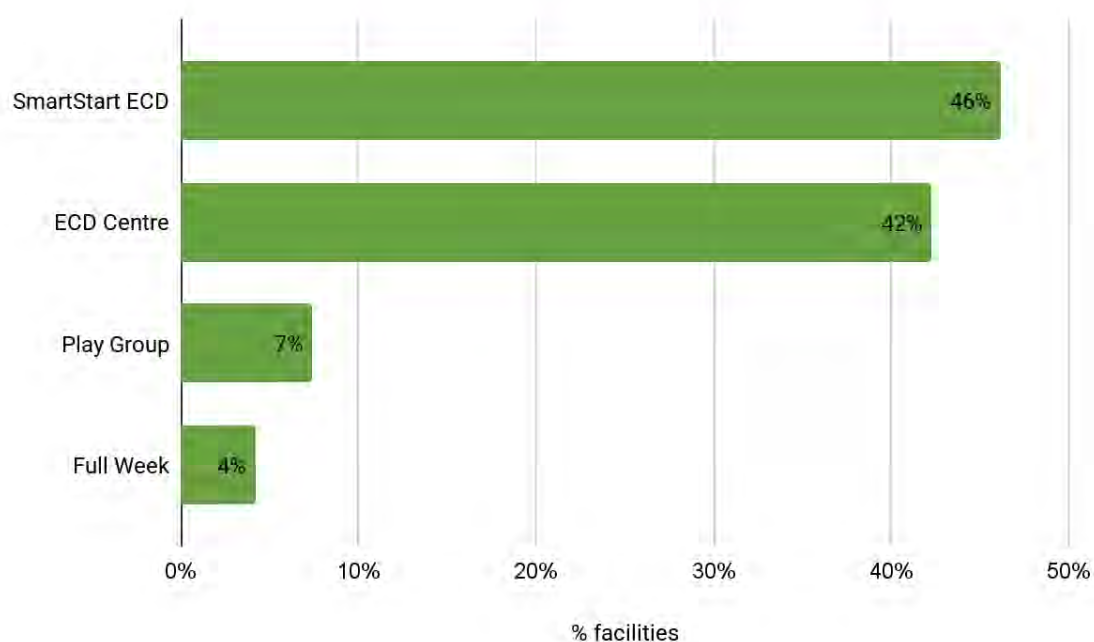
The fee structures implemented by the sampled ELPs reflect the economic context of the children they serve, the majority of whom are from low-income households. According to Figure 4, a substantial 54% of these facilities levied fees below R200 per month, and a noteworthy 14% imposed no charges at all.

Figure 4: Fees
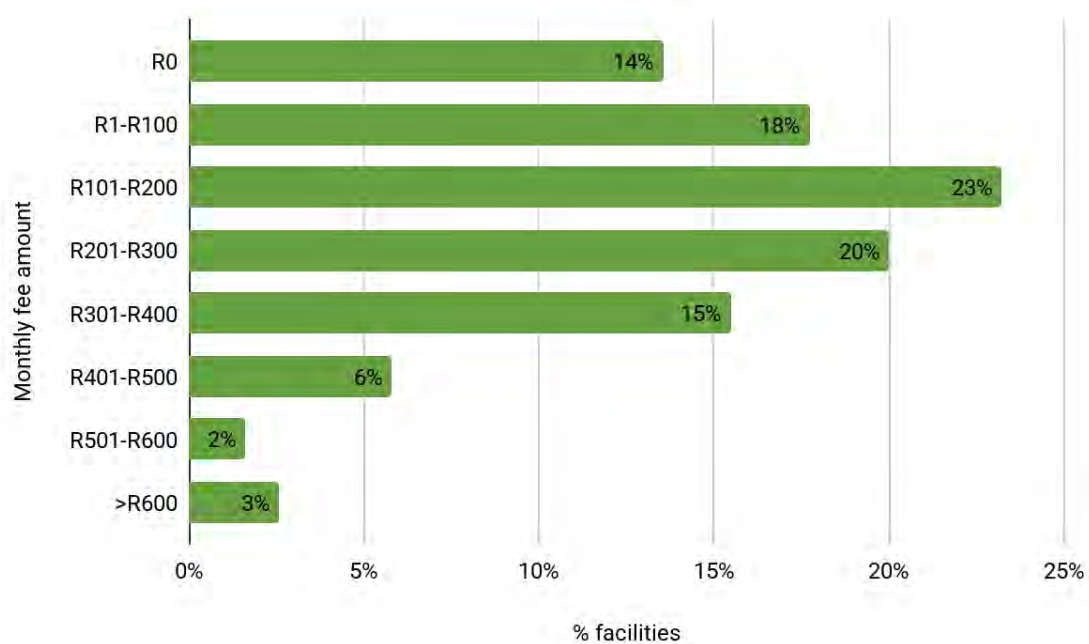
The economic background of these low-fee charging programmes was also evident in the physical infrastructure and building materials of the ELP facilities, with approximately 30% of ELPs located in informal structures such as informal housing, containers, or traditional dwellings (Figure 5).

Figure 5: Building type[27]
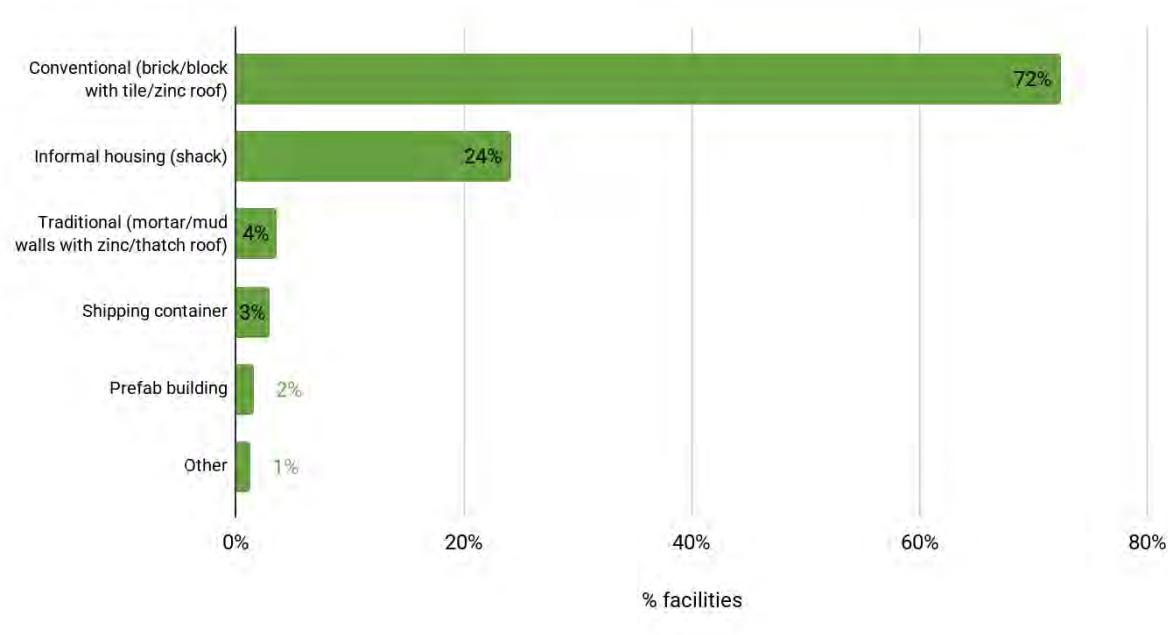
This economic context was further mirrored in the limited availability of crucial classroom resources, as shown in Table 2. While drawing, painting, and writing materials (95%), as well as children's books (93%) were widely available, access to other important classroom resources varied. Notably, close to 35% of ELPs lacked materials for counting, 33% lacked chairs, desks, and tables, and 18% reported an absence of toys, dolls, or dress-up clothes. This aligns with the fact that the majority of ELPs were located in rural and low-income areas, and were thus more likely to face funding constraints.

Table 2: Classroom resources

Classroom Material	% Facilities
Drawing, painting, writing materials	95%
Children's books	93%
Picture cards, posters	86%
Puzzles, games	86%
Blocks (e.g., LEGO)	82%

---

[27] The question allowed for more than one building type to be selected.

Classroom Material	% Facilities
Toys, dolls, dress-up clothes	82%
Chairs, desks, tables	67%
Materials for counting	66%
Clay, play dough	59%
Carpet, sleeping mats	57%
Balls, hula-hoops, sandbags	53%
Skipping ropes, scooters	45%
Buckets, spades, sand moulds	41%
Theme tables	38%
Musical instruments	34%

Child-adult ratios are an important metric for understanding children's access to programme facilitators, directly influencing their quality of learning. In our analysis, presented in Figure 6, we found that facilities most commonly (37%) maintained a ratio of 6-10 children per adult, with adults including both practitioners and assistants. A substantial share of facilities also had ratios of 11-15 children per adult (24%) and 1-5 children per adult (23%). However, it is essential to note that including assistants as "classroom adults", who are often unqualified, may present a misleadingly favourable picture of staffing levels. Any imbalances, such as high child-adult ratios, can compromise overall educational quality, emphasising the importance of appropriate staffing levels in ELPs.

Figure 6: Adult-children ratios
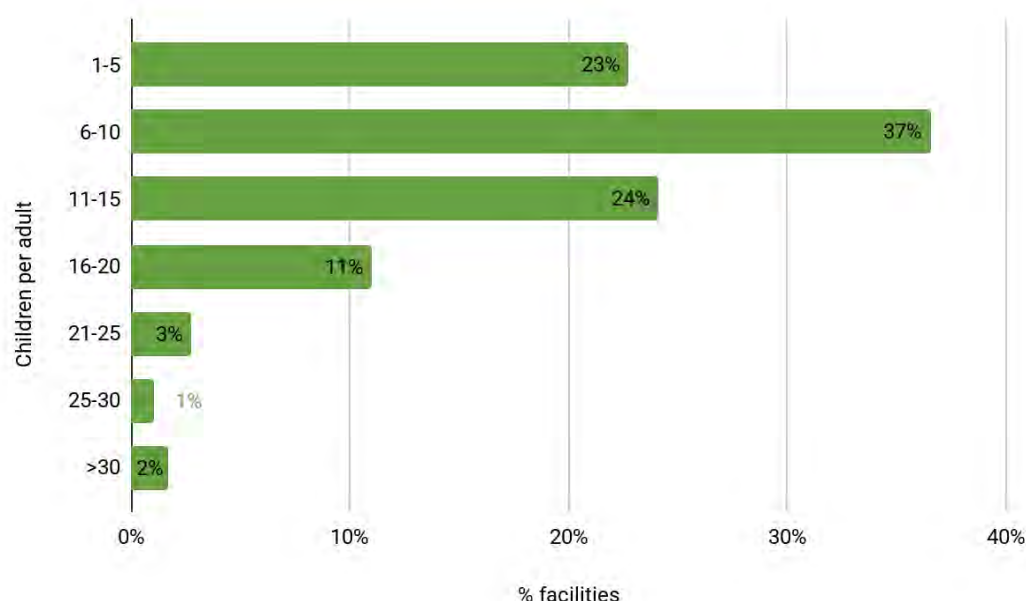
PQA

SmartStart uses the PQA tool to measure programme quality at SmartStart ELPs. The tool comprises 34 rating items administered by classroom observation that can be aggregated to assess the quality of six domains:

1. A stimulating and adequately resourced learning environment
2. Consistent use of the SmartStart routine
3. A stable and nurturing environment where children feel safe and loved
4. Positive and plentiful adult-child interactions that encourage rich use of language
5. Opportunities for child-directed, open-ended play, supported and extended by adults
6. Interactive storytelling that introduces children to new language and learning

Each of the 34 observational items was given a score of 0 (Inadequate), 1 (Basic), or 2 (Good). Scores across all observational items can be summed to give a total PQA score, with specific cut-offs for overall quality rating. The majority (77%) of ELPs had a Green PQA rating, with 10% an Orange PQA rating and 3% with a Red PQA rating. The remainder (10%) did not have a PQA rating.

Table 3 shows the average score for each item that comprises the six domains. Overall, the average scores for each domain were high with many items having an average score very

close to 2 (the maximum possible score). The domain "A stimulating and adequately resourced learning environment" scored the best with an average score of 1.8, while the lowest scoring domain was "Positive and plentiful adult-child interactions that encourage rich use of language" which had an item average of 1.5. All other domains had an average item score of 1.6.

Table 3: Average scores for each PQA item

Domain	Item	Average Score
A stimulating and adequately resourced learning environment	Supervision	1.8
	Learning space	1.9
	Using the toy kit	1.8
	Labelling	1.7
	Toys and storybooks	1.8
	Displays	1.5
Consistent use of the SmartStart routine	SmartStart Activities	1.9
	SmartStart Routine	1.9
	Small group time	1.7
	Making plans	1.5
	Free play	1.5
	Recall	1.4
	Storytime	1.6
	Large group time	1.6
A stable and nurturing environment where children feel safe and loved	Warm & respectful interactions	1.9
	Individual attention	1.7
	Acknowledgement & encouragement	1.6
	Looking after upset children	1.5
	Maintaining order	1.7
	Resolving conflict	1.3
Positive and plentiful adult-child interactions which encourage rich use of language	Talking to children & encouraging communication	1.8
	Listening & responding	1.6
	Using talk to extend learning	1.5
	Building language	1.4

Domain	Item	Average Score
	Encouraging initiative	1.4
Opportunities for child-directed, open-ended play, supported and extended by adults	Letting children make choices	1.5
	Facilitating children's play	1.7
	Participating in children's play	1.7
	Extending learning through play	1.5
	Ensuring play & learning is at the right level	1.6
Interactive storytelling that introduces children to new language and learning	Encouraging conversation during story time	1.7
	Explaining new words & ideas	1.5
	Asking questions	1.6
	Familiarising children with books	1.5

ELP Practitioners

Figure 7 displays the highest level of education obtained by the ELP practitioners, and shows that almost half of the practitioners did not complete their Matric year, with only 33% obtaining a Matric or National Senior Certificate as their highest qualification. This highlights the importance of programmes such as SmartStart in equipping practitioners with the knowledge and tools needed to provide better educational support to their learners.

Figure 7: Practitioner education
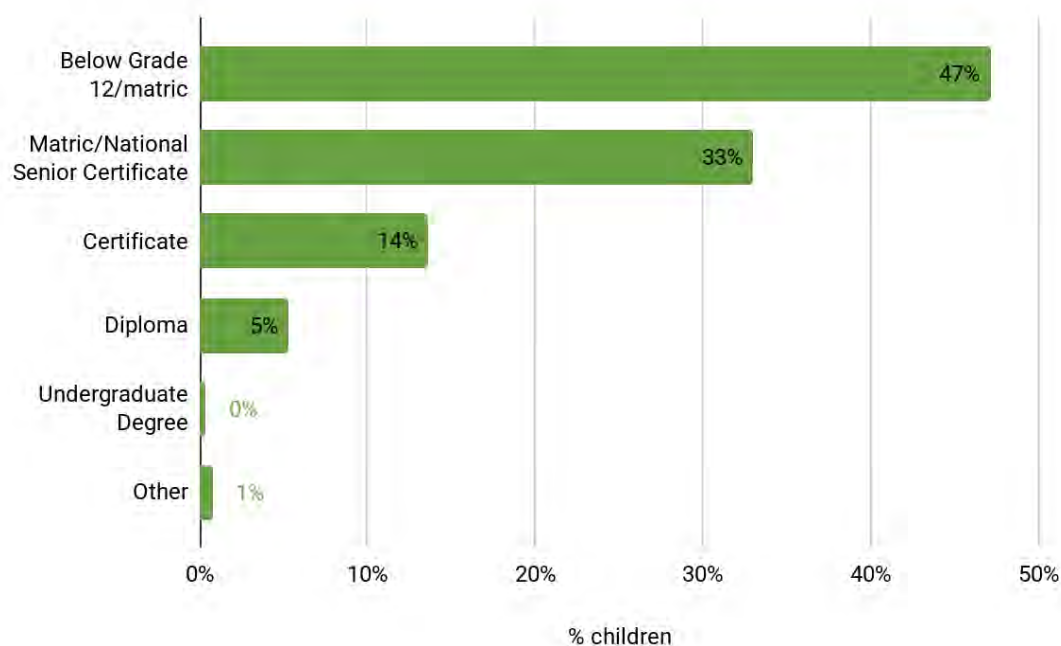
SmartStart

The surveyed ELPs all affirmed their affiliation with SmartStart, and their practitioners provided insights into their experiences with the programme. The duration of practitioners' involvement varied widely, spanning from two months to eight years. Notably, a significant portion of the sampled ELPs had relatively brief tenures with SmartStart, with almost half having been involved with SmartStart for less than two years. As shown in Figure 8, this range of participation reflects both long-standing and recent commitments to improving the national educational and ECD landscape.

Figure 8: Months in SmartStart
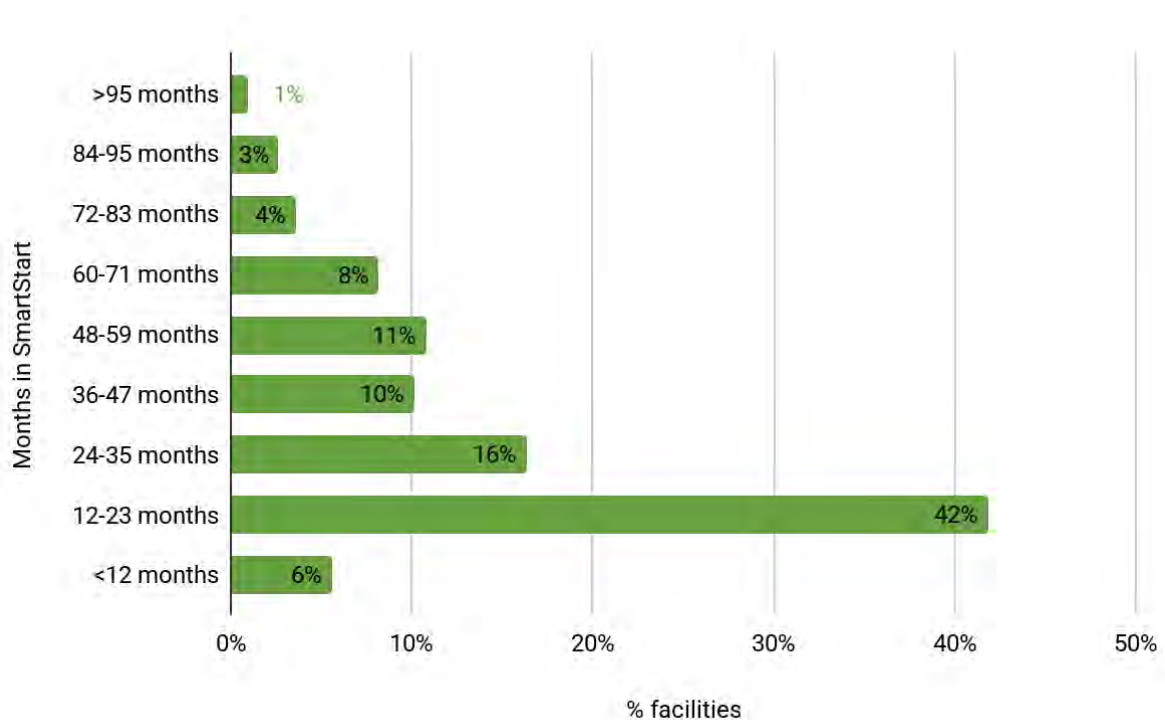
SmartStart's influence in ELPs was highlighted by the impressive uptake of SmartStart clubs and dedicated coaches. A substantial 93% of practitioners confirmed their participation in a club, showcasing widespread engagement with this collaborative component of the SmartStart programme. Similarly, 92% of practitioners reported having a coach assigned to their ELP, a crucial support mechanism provided by SmartStart. Among SmartStart club members, a substantial 95% attended club meetings, with 91% of members expressing the value of club participation as either useful (25%), very useful (33%), or extremely useful (33%).

In terms of coach visits, 86% of practitioners found them to be valuable, rating them as useful (20%), very useful (28%), or extremely useful (38%). Almost 80% of practitioners highlighted that their SmartStart coaches provided constructive feedback after visits. These feedback sessions often included discussions on tangible ways to enhance the overall functioning of their centres.

While coach visits were valuable, the frequency of coach visits varied, ranging from no visits in a year (6%) to more than 13 visits in a year (3%) (Figure 9). The majority of coaches visited the ELPs once or twice a year (29%), followed by 3-4 times per year (23%) and 5-6 times per year (16%) (Figure 9). A noteworthy 77% of practitioners expressed a desire for more

frequent coach visits, signalling the demand for heightened support and guidance within their ELPs and their commitment to improving their students' educational outcomes.

Figure 9: Frequency of coach visits
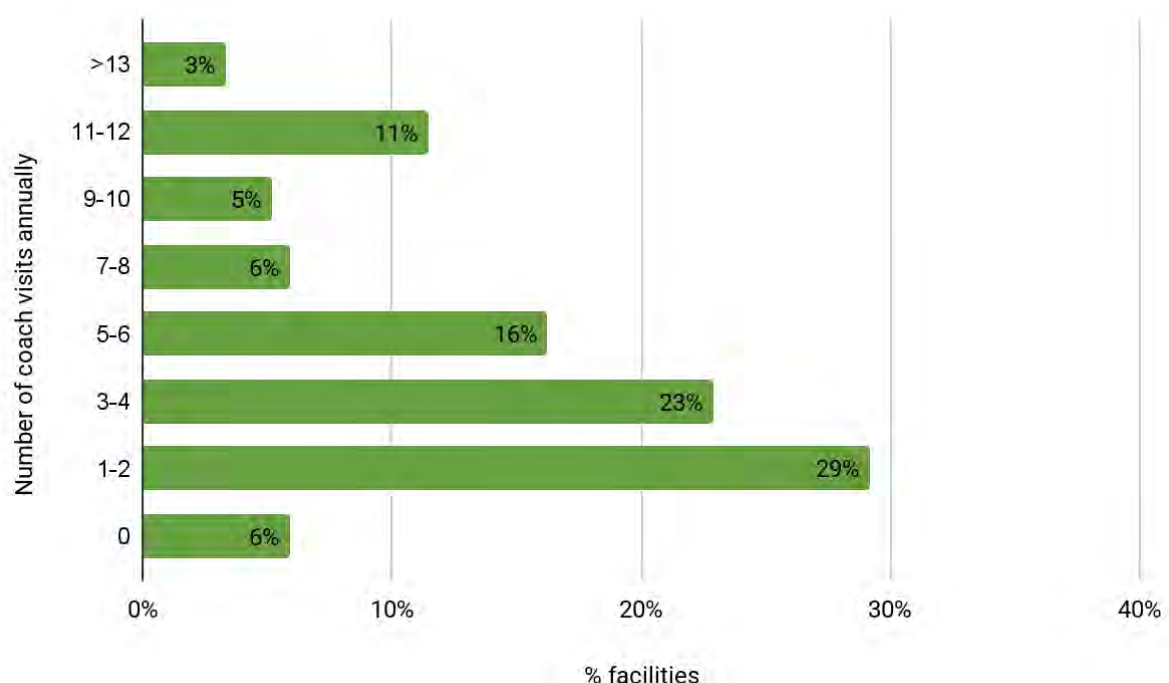
The choice of communication medium can play an important role in fostering better interaction between practitioners and coaches. As shown in Figure 10, mobile communication dominated as the preferred medium. In this regard, 86% of the practitioners disclosed that they were in contact with the coaches through WhatsApp and 78% through phone calls. Additionally, 26% communicated with coaches through meetings, and only 11% used SMS. The least popular means of communication - email or written communication - was used by less than 10% of practitioners.

Mobile communication formats enable instant communication, which is particularly important for practitioners located in rural areas whose facilities may lack various resources and access to WiFi networks. There is also a reduced sense of formality, which helps to foster a comfortable environment to enable ongoing discussion, feedback and the exchange of ideas. The popularity of WhatsApp, in particular, suggests that these benefits of mobile communication were being utilised. That being said, it is important to ensure that all practitioners are digitally active and able to engage with coaches through mobile platforms in order to realise the potential benefits of instant mobile communication.

Figure 10: Coach contact with practitioners
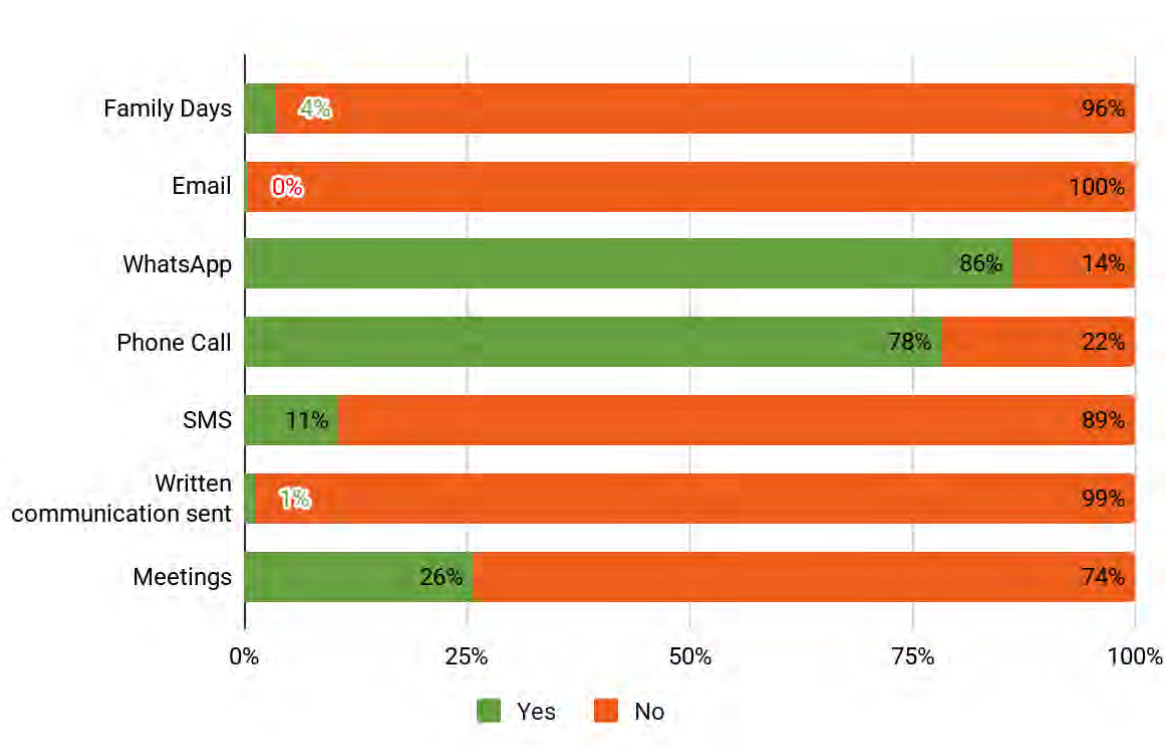
To further bolster the capacity of ELP practitioners, SmartStart provides all practitioners with the opportunity to attend a comprehensive two-day business skills workshop. The response was encouraging, with 69% of SmartStarters having participated in the workshop, and the feedback was overwhelmingly positive, with nearly unanimously positive feedback. Only a minimal 4% of attendees found the training to be not at all useful or just slightly useful.

Finally, to gauge the programme's impact, practitioners were asked to evaluate their likelihood of recommending SmartStart to another ELP. The findings, depicted in Figure 11 below, indicated strong support for SmartStart, with only 15% providing a likelihood rating of less than 7. This endorsement suggests a general impression regarding the programme's overall efficacy, and the positive reception it garnered among the participating practitioners.

Figure 11: Likelihood (out of 10) of referring another ELP to SmartStart
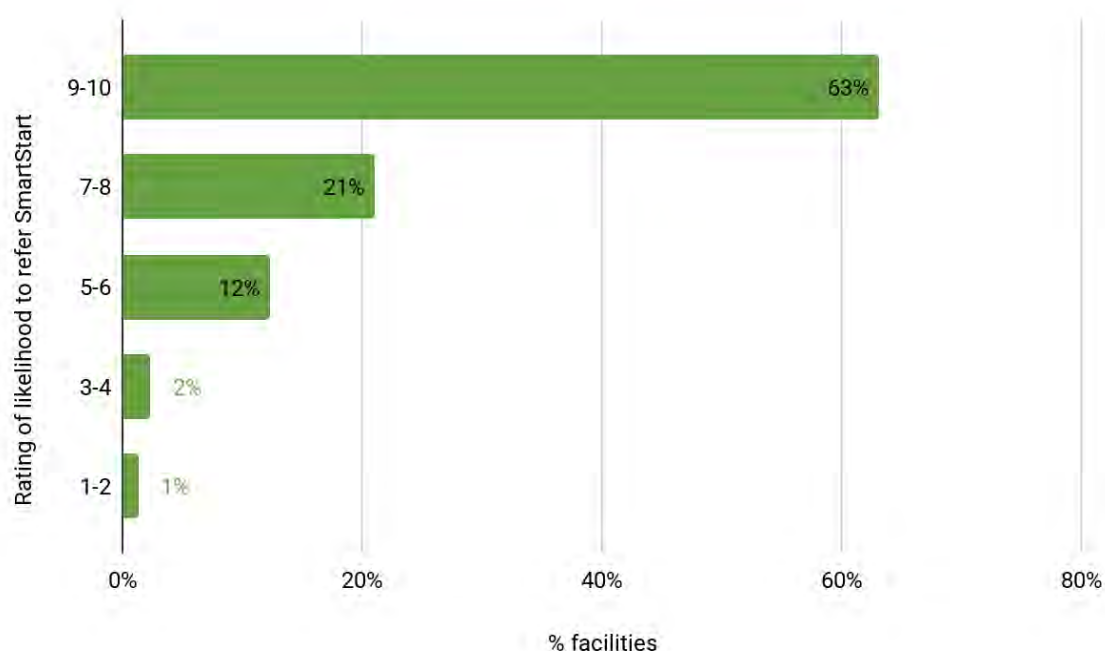
The descriptive statistics offer valuable insights into the operational landscape of SmartStart ELPs. The data show how predominantly low-income households are supported, demonstrating the initiative's success in reaching one of the most economically vulnerable segments. Additionally, the PQA assessment paints an overall reassuring picture of programme quality, with consistently high average scores observed across the six domains, albeit with nuanced variations in specific items within the domains.

The SmartStart initiative's support mechanisms, as seen in the clubs, coaches, and business skills workshops, were widely embraced by the participating practitioners and ELPs. This widespread adoption signifies a positive impact on the operational aspects and overall quality of early childhood education provided by these programmes. The overwhelmingly positive feedback from practitioners, both in terms of training and their likelihood to recommend SmartStart to other ELPs, serves as a testament to the programme's overall success and indicates the potential for continuous improvement and enhancement in early childhood education practices through SmartStart.

4 | Analysis

Evaluation Question 1

Are quality SmartStart programmes on average ensuring that participating children achieve better outcomes than those achieved on average by other children in South Africa?

We employ two approaches to answer the first evaluation question. Firstly, we aim to determine the change in SmartStart children 'On track' according to the ELOM standards from baseline to endline. An increase in the proportion of children meeting the age-controlled benchmark is suggestive of a positive association of SmartStart on the ELOM domains. Secondly, we utilise regression analysis to estimate the predicted endline ELOM scores for children in SmartStart compared to those in non-SmartStart ELPs, while controlling for various confounding variables, with the 2021 Thrive by Five Index sample serving as the comparison group.

1. Increase in SmartStart Children 'On track'

SmartStart programmes saw, on average, a net positive outcome achieved by children, with average ELOM gains of 1.47 points per month[28] (Table 4). All 551 children in the sample completed baseline and endline ELOM assessments, with the average increase in age between baseline and endline being almost 8 months. Looking at the ELOM outcomes, we see an average total improvement in ELOM scores of 11.6 between baseline and endline. Across ELOM domains, the increase in score was greatest for ENM (+2.6) and CEF (+2.5) and lowest for GMD (+1.9).

---

[28] It is important to note that this 1.47 points per month includes both a maturation effect - children's normal improvement in ELOM scores due to ageing - and a programme effect - the improvement in scores due to their participation in the ELP, independent of ageing. Dawes and Henry (2023) estimate a maturation effect of 1.04 ELOM points per month, with a confidence interval of 0.85 - 1.22. While Van der Berg (2021) estimates a lower bound maturation effect of 1.02 and upper bound of 1.23 ELOM points per month of age. Using Dawes and Henry's maturation effect estimate, SmartStart's programme effect, independent of the natural ageing effect, is approximately 0.43 ELOM points per month. This is equivalent to raising children's scores by 3.44 points, on average, on the total ELOM score over an eight month period.

Table 4: Average ELOM scores at baseline vs endline

Timepoint	Total ELOM (out of 100)	GMD (out of 20)	FMCVMI (out of 20)	ENM (out of 20)	CEF (out of 20)	ELL (out of 20)	Average age (months)
Baseline	44.7	8.3	10.9	8.5	6.3	10.7	54.4
Endline	56.3	10.1	13.3	11.1	8.7	13.1	62.3
Change	+11.6	+1.9	+2.3	+2.6	+2.5	+2.3	+7.9

Each of a child's ELOM scores (Total and each of the five domains) can be categorised into one of three performance bands: 'On track', 'Falling behind', and 'Falling far behind'. These benchmarks were established following consultation with education experts and a review of performance across an Age Validation sample of South African children, with performance inspected for the full sample as well as within the three school quintile groups (1, 2/3, and 4/5)[29]. Children with scores that fall within the 'On track' band are situated within the top 40% of the age validation sample distribution. Those classified as 'Falling far behind' significantly lag behind the standard and require substantial assistance to meet expectations, while children deemed to be 'Falling behind' may attain the standard with additional support.

Table 5 below shows the performance bands for the ELOM Total score and each of the domains, according to two age groups 50-59 months and 60-69 months. To account for maturation, the thresholds for all performance bands are higher once a child is aged 60-69 months. As the majority of the children in our sample moved between age categories between baseline and endline, when using the three performance bands to analyse change, maturation is automatically accounted for.

Table 5: ELOM scores reflecting standards and performance bands for children aged 50-59 and 60-69 months

Scores	50 - 59 months			60 - 69 months		
	Falling far behind	Falling behind	On track	Falling far behind	Falling behind	On track
Total ELOM	0-36.01	36.02-46.31	46.32-100	0-43.23	43.24-54.37	54.38-100
GMD	0-5.40	5.41-8.59	8.60-20	0-7.21	7.22-10.53	10.54-20
FMCVMI	0-9.70	9.71-12.31	12.32-20	0-11.46	11.47-14.12	14.13-20

---

[29] For more information, see Dawes, A., Biersteker, L., Girdwood, E., Snelling, M. & Tredoux, C. (2016). Early Learning Outcomes Measure: Technical Manual. Innovation Edge.

Scores	50 - 59 months			60 - 69 months		
	Falling far behind	Falling behind	On track	Falling far behind	Falling behind	On track
ENM	0-6.34	6.35-9.32	9.33-20	0-6.90	6.91-10.23	10.24-20
CEF	0-4.07	4.08-7.16	7.17-20	0-5.84	5.85-9.26	9.27-20
ELL	0-6.53	6.54-10.25	10.26-20	0-7.97	7.98-11.64	11.65-20

Source: Dawes, A. et al. 2016. Table 13.

Figure 12 below shows the change in the proportion of SmartStart children with total ELOM scores that moved from being 'Falling far behind' or 'Falling behind' to 'On track' between baseline and endline. The proportion of children 'On track' increased substantially, with under half at baseline and almost two-thirds at endline, an improvement of 20 percentage points. We also see that the proportion of children 'Falling far behind' fell substantially from 26% to 14%, a decline of 12 percentage points between baseline and endline.

Figure 12: Proportion of children 'Falling far behind'/'Falling behind'/'On track' based on their total ELOM scores at baseline and endline
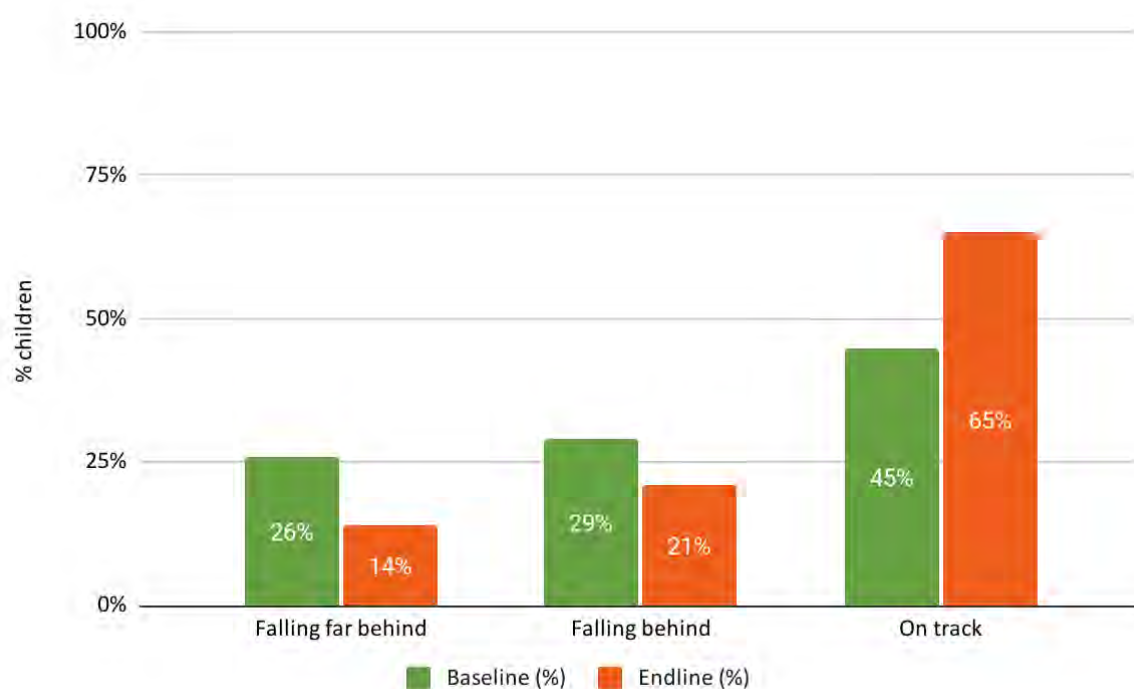

Figure 13 shows that the gains in the proportion of children 'On track' varies widely between domains. With a 25 percentage point increase, ENM saw the biggest gain in the proportion of children 'On track' between baseline and endline, going from the second worst performing domain at baseline to the best performing domain at endline. FMCVMI and ELL

saw similar gains of about 10 percentage points while GMD and CEF saw relatively modest gains of about 4 percentage points. At endline, the proportion of children 'On track' also varied widely between domains. ELL (65%) and ENM (62%) had the largest proportion of children 'On track', while this was lower for GMD (51%), CEF (47%), and FMCMVI (44%), suggesting SmartStart should place additional focus on these domains.

Figure 13: Proportion of children 'On track' at baseline and endline for each ELOM domain
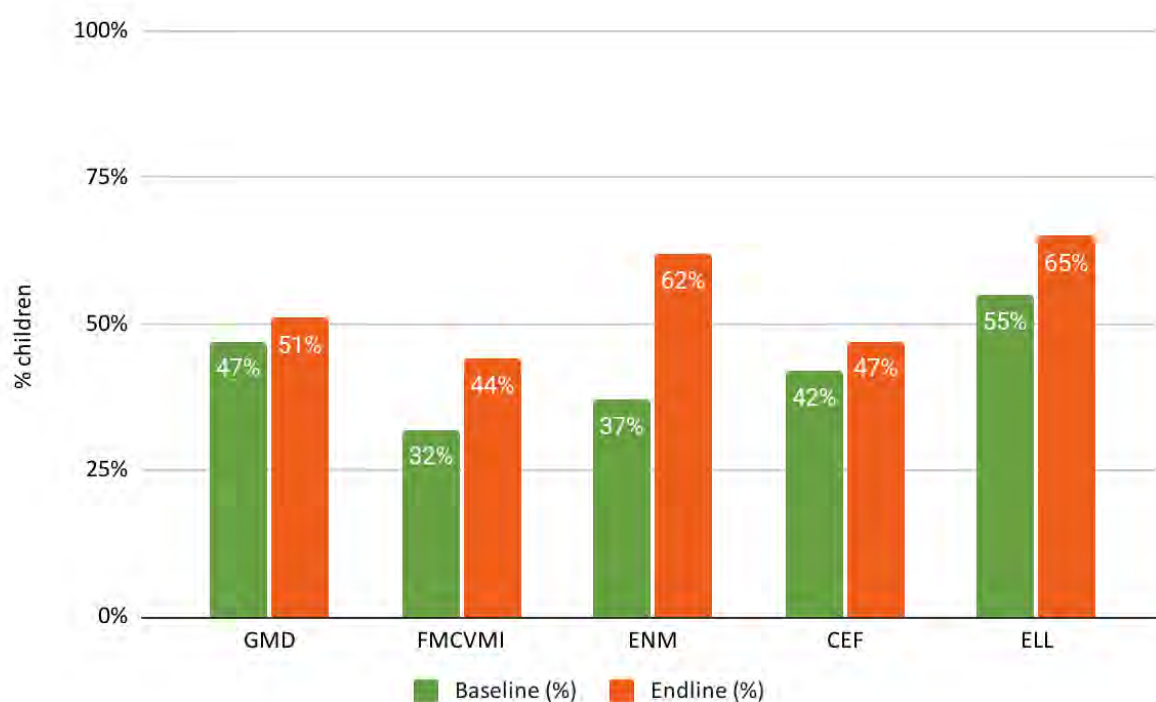
2. Comparison Between SmartStart and Thrive by Five 2021

To investigate whether children enrolled in SmartStart ELPs achieve better outcomes on average than children in other programmes, we compare the endline SmartStart ELOM scores to the 2021 Thrive by Five Index ELOM scores. Ideally, we would compare SmartStart ELOM gains to another sample's ELOM gains, however, there is no representative ELOM dataset available that has an analogous 'baseline-endline' design that would allow for this. Nevertheless, the Thrive by Five sample is nationally representative which means that it is a good benchmark of the 'average ECD-aged child' against which to compare SmartStart's sample.

Figures 14 and 15 below compare the proportion of children 'On track' between SmartStart and Thrive by Five. Mirroring the results in Table 6, Figure 14 shows a statistically

significantly greater proportion of children 'On track' in the SmartStart sample (65%) than in the Thrive by Five sample (46%).[30] Figure 15 then compares SmartStart to the Thrive by Five sample by domain, where for all domains there is a greater proportion of SmartStart children 'On track' than the Thrive by Five sample and the differences in proportion of 'On track' children varies widely by domain. The difference between SmartStart and Thrive by Five is greatest for ENM (+25 percentage points), followed by FMCVMI (+12 percentage points) and ELL (+10 percentage points), with statistically insignificant differences for CEF (+5 percentage points) and GMD (+3 percentage points).[31]

Figure 14: Proportion of children 'On track' for total ELOM scores

---
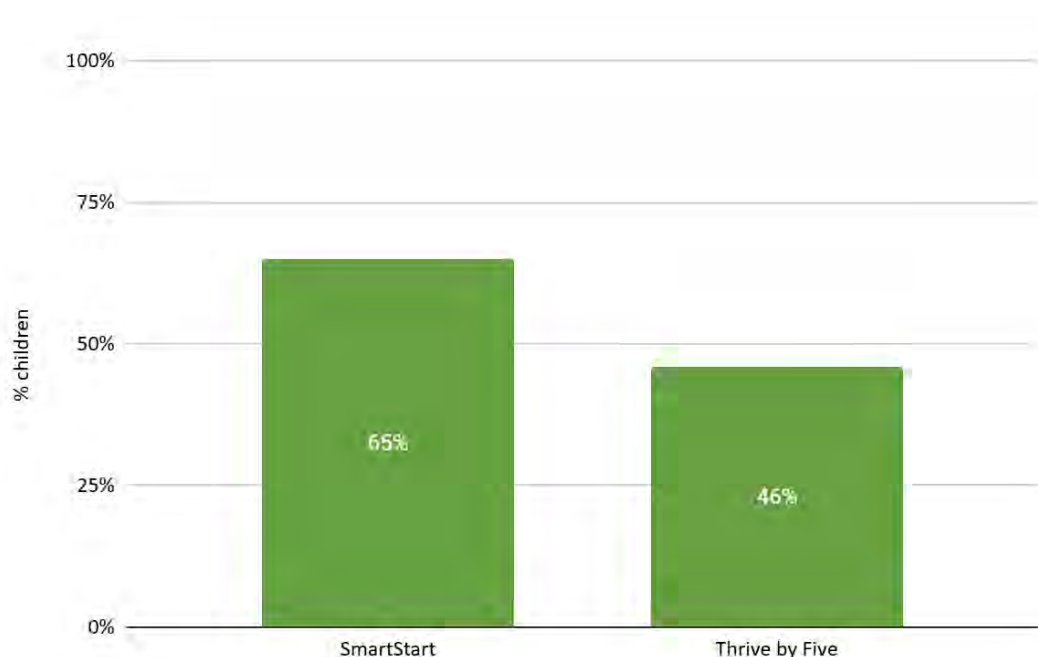
[30] Total ELOM: X2 (5,772, N = 5,773) = 66.22, p = 0.000)

[31] GMD: X2 (5,772, N = 5,773) = 1.69, p = 0.210; FMCVMI: X2 (5,772, N = 5,773) = 26.25, p = 0.000; ENM: X2 (5,772, N = 5,773) = 109.85, p = 0.000; CEF: X2 (5,772, N = 5,773) = 2.260, p = 0.13; ELL: X2 (5,772, N = 5,773) = 16.89, p = 0.000.

Figure 15: Proportion of children 'On track' for each ELOM domain
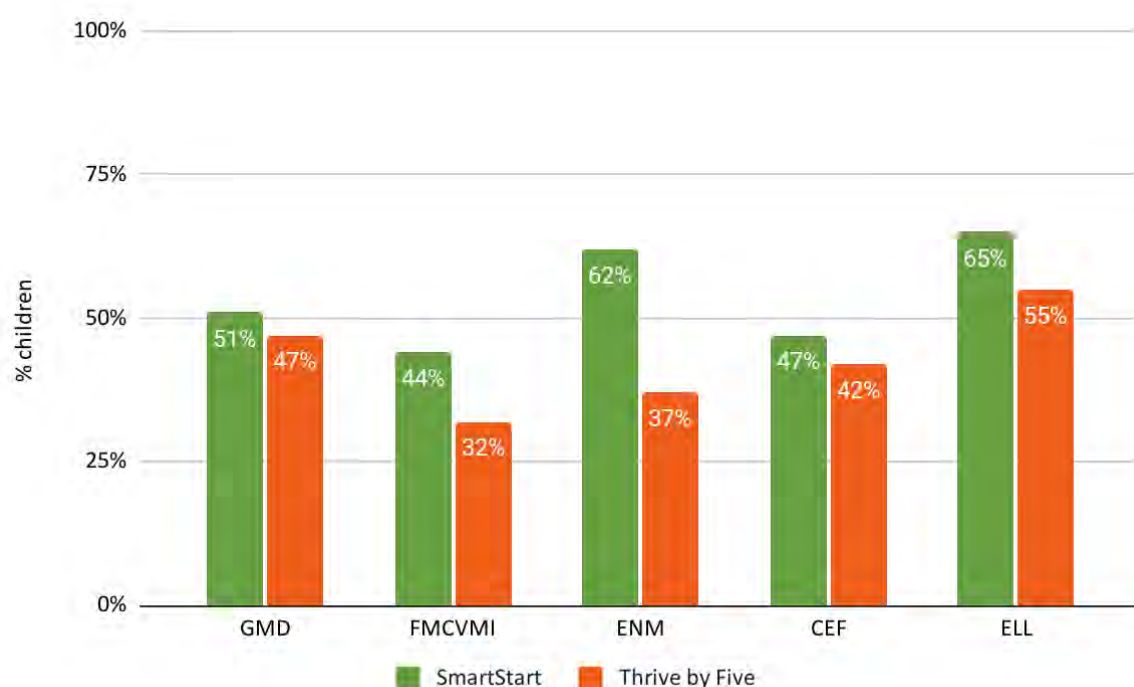
Table 6 below compares the average endline SmartStart ELOM scores to those of the 2021 Thrive by Five Index. The table shows children in SmartStart, on average, scored better across all domains compared to children in the Thrive by Five samples. The difference was greatest for ENM (+2.8 points) and ELL (+2.5 points) and lowest for CEF (+1.7).[32]

Table 6: ELOM score comparison: endline 2023 SmartStart to 2021 Thrive by Five

Study	Total ELOM	GMD	FMCVMI	ENM	CEF	ELL
SmartStart	56.3	10.1	13.3	11.1	8.7	13.1
Thrive by Five	45.1	8.1	11.0	8.3	7.0	10.6
Difference	+11.3	+2	+2.3	+2.8	+1.7	+2.5

---

[32] Total ELOM: $t(5,766) = 17.77.36$, $p = 0.000$; GMD: $t(5,766) = -10.89$, $p = 0.000$; FMCVMI: $t(5766) = -14.88$, $p = 0.000$; ENM: $t(5766) = 14.16$, $p = 0.000$; CEF: $t(5766) = 9.69$, $p = 0.000$; ELL: $t(5766) = 13.36$, $p = 0.000$

Comparison Between SmartStart and Thrive by Five Controlling for Sample Differences

A comparison of the 2023 SmartStart sample and the Thrive by Five 2021 sample presents two key limitations. First, the Thrive by Five data lacks baseline ELOM scores making it impossible to estimate programme gain and compare the change in ELOM scores between the two programmes. A higher average programme gain within the SmartStart sample would theoretically serve as the strongest evidence for programme effectiveness, as it addresses the potential issue of differing baseline ELOM scores between the two samples. It is possible that the SmartStart sample's average initial ELOM scores are lower than what the Thrive by Five program's initial ELOM scores might have been (had they been captured), as SmartStart children are generally from more deprived home environments. Consequently, the SmartStart sample's average endline ELOM scores could have been even higher, had the sample consisted of children identical in composition to the Thrive by Five groups.

The second limitation is that the two samples are dissimilar along several important characteristics, most notably average age in months. Similar arguments could be made for gender differences, class resourcing, practitioner qualifications and socio-economic status. These are confounding variables because they affect both who receives the SmartStart programme, and how well the children do on the ELOM assessment. To at least partially address this, we use regression analysis to statistically "control" for these confounding factors, which means it holds them constant so we can better estimate the effect of the programme, as if everyone had the same background.

Table 7 presented below displays the OLS regression results, showing the correlation between ELOM score and sample while controlling for various confounding variables. These covariates were chosen based on theoretical considerations and previous models utilised in the 2018 SmartStart Outcomes Evaluation. Regrettably, certain potentially impactful covariates, such as 'rural/urban setting', child socio-economic status and practitioner experience, could not be included in the analysis due to the unavailability of data. Consequently, the results presented should be approached with caution, as the absence of these variables may introduce potential omitted variable bias.

The results show that, with the exception of CEF, across domains the SmartStart sample statistically significantly outperformed the Thrive by Five 2021 sample. The

best-performing domain relative to Thrive by Five was ELL, where on average and holding all else equal, SmartStart children scored 1.34 points better than Thrive by Five children. In line with the earlier analysis, the SmartStart sample also scored better in ENM (+0.89 points) while for the other domains, the relationship between the sample and ELOM score was not statistically significant at the 5% threshold. For the total ELOM score, SmartStart children scored, on average and holding all else equal, 3.53 points more than Thrive by Five children.

Table 7: Regression analysis, relationship between ELOM scores and sample while controlling for sample differences

	Total ELOM	GMD	FMCVMI	ENM	CEF	ELL
SmartStart	3.53**	0.64*	0.46*	0.89***	0.20	1.34***
	(1.15)	(0.33)	(0.26)	(0.34)	(0.31)	(0.37)
Endline age	1.02***	0.18***	0.25***	0.21***	0.21***	0.16***
	(0.1)	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)
Years in programme	1.17**	0.10	0.2**	0.37***	0.13	0.37***
	(0.37)	(0.1)	(0.09)	(0.11)	(0.1)	(0.12)
Female	2.73***	-0.15	1.02***	0.46***	0.45**	0.95***
	(0.48)	(0.14)	(0.12)	(0.15)	(0.15)	(0.15)
Fee amount	0.008***	0.00	0.002***	0.002***	0.003***	0.002***
	(0)	(0)	(0)	(0)	(0)	(0)
Has matric	0.12	0.02	0.04	0.00	-0.17	0.23
	(0.75)	(0.21)	(0.17)	(0.21)	(0.2)	(0.25)
Class size	0.04	0.00	-0.01	-0.01	-0.02*	-0.01
	(0.04)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Class has books	0.02	-0.12	0.10	-0.08	0.13	-0.01
	(0.86)	(0.23)	(0.19)	(0.25)	(0.24)	(0.29)
Height-for-age	1.35***	0.34***	0.25***	0.28***	0.29***	0.19***
	(0.33)	(0.1)	(0.07)	(0.08)	(0.07)	(0.07)
Constant	-18.88***	-2.28	-4.61***	-5.24***	-5.49***	-1.25

	Total ELOM	GMD	FMCVMI	ENM	CEF	ELL
	(4.9)	(1.47)	(1.16)	(1.42)	(1.45)	(1.6)
Observations	2641	2641	2641	2641	2641	2641
R-squared	0.21	0.07	0.18	0.13	0.15	0.12

Notes: *p<0.05, **p<0.01, ***p<0.001. Models report robust standard errors to account for clustering at the ELP level.

Evaluation Question 2

How do gains in ELOM scores over time compare with gains seen during the 2018/2019 outcome evaluation?

To answer the second evaluation question we take two approaches. We first compare the proportion of children 'On track' in the 2023 SmartStart sample to the 2018 SmartStart sample. We then examine the raw differences in ELOM point gains between the two samples and use regression analysis to control for sample differences. For total ELOM score we find that, while the proportion of children 'On track' is greater in the 2023 endline sample, the gain in ELOM scores between baseline and endline is greater for the 2018 sample. Furthermore, for all domains except ELL, we find the proportion of children 'On track' is greater in the 2018 endline sample, and that apart from ENM, the gain in ELOM scores is greater in the 2018 sample.

1. Proportion of Children 'On track': 2018 versus 2023

The proportion of children 'On track' increased to 65% in the endline assessment, up from 59% recorded in the 2018 evaluation. Table 8 demonstrates that when comparing baselines, the 2018 sample had a lower proportion of children 'On track' compared to the 2023 sample across all domains except ENM. At endline, however, the difference between the two samples reduces and reverses, such that for the ENM, CEF, and GMD domains, there is a greater proportion in the 2018 sample 'On track' than the 2023 sample. Curiously, there is

still a greater proportion of children 'On track' for total ELOM score in the 2023 sample than in the 2018 sample, despite lower proportions for GMD, FMCVMI, ENM, and CEF.[33]

The increase in the proportion of children 'On track' was significantly more pronounced in the 2018 sample, except for the ENM domain. Specifically, in the total ELOM score, the 2018 sample demonstrated a 26.9 percentage point improvement in the proportion of children 'On track', greater than the 20 percentage point improvement seen in the 2023 sample. The contrast is most evident in the GMD domain, where the 2018 sample showed a substantial 34.8 percentage point improvement, contrasting with the modest 2.6 percentage point improvement in the 2023 sample. Notably, ENM is the sole domain where the 2023 sample outpaced the 2018 sample, with a 22.8 percentage point improvement compared to the 2018 sample's 20.9 percentage point improvement.

Table 8: Change in children 'On track', 2018 versus 2023

Study and Time Point		Total ELOM	GMD	FMCVMI	ENM	CEF	ELL
SmartStart 2018	Baseline	32%	34%	20%	45%	28%	40%
	Endline	59%	68%	46%	66%	52%	59%
	Change (p.p.)	+27	+34	+26	+21	+24	+19
SmartStart 2023	Baseline	45%	48%	27%	39%	39%	53%
	Endline	65%	51%	44%	62%	47%	65%
	Change (p.p.)	+20	+3	+17	+23	+8	+12

2. Descriptive Comparison of ELOM Gain

Looking at average ELOM baseline scores, Table 9 shows that the 2023 sample had average ELOM scores greater than the 2018 sample for all domains except ENM (which is only marginally greater). At endline, this trend then switches with the 2018 sample having average scores marginally greater than the 2023 sample for all domains except ELL. Thus, the 2023 sample had substantially lower ELOM point gains than the 2018 sample - for Total ELOM, the ELOM point gain for the 2023 sample was just over half that observed in the 2018 sample (11.6 points versus 21 points). Across domains, the reduction in ELOM gains in 2023,

---

[33] Each of the ELOM domains have a different range which determines the 'Falling far behind', 'Falling Behind' or 'On track'. Eg. GM's 'Falling far behind' categorisation ranges between 0-5.4, FMCVMI's 'Falling far behind' ranges between 0-9.7.

compared to in 2018, was greatest for GMD (gain reduced by 3.2 points in 2023), trailed by FMCVMI (2 points), CEF (1.6 points), ENM (1.5 points) and ELL (1 point). All ELOM point gain differences were statistically significant. [34]

Table 9: SmartStart ELOM gain, 2018 versus 2023

Time Point		Total	GMD	FMCVMI	ENM	CEF	ELL
Baseline score (average)	2018	39.4	6.6	9.3	8.8	5.4	9.3
	2023	44.7	8.3	10.9	8.5	6.3	10.7
Endline score (average)	2018	60.4	11.8	13.5	13.0	9.5	12.6
	2023	56.3	10.1	13.3	11.1	8.7	13.1
ELOM point gain	2018	+21	+5.1	+4.3	+4.1	+4.1	+3.3
	2023	+11.6	+1.9	+2.3	+2.6	+2.5	+2.3
ELOM point gain difference		-9.4	-3.2	-2	-1.5	-1.6	-1

3. Regression Analysis: Controlling for Sample Differences

While attending a SmartStart ELP is associated with positive ELOM gains in both the 2018 and 2023 samples, the gain is slightly lower in the 2023 sample. Accurately gauging the significance of this apparent decrease would require that the 2023 and 2018 samples were comparable. Regression analyses were used to allow us to control for these confounding factors in the absence of a comparable sample.

Table 10 below displays regression results examining the relationship between sample year and ELOM gain, while controlling for child gender, child age, classroom child-adult ratio, whether the teacher completed matric, years of child attendance in the programme, height-for-age, fee versus no fee, and classroom Total PQA score (as a measure of programme quality). Years of practitioner experience, a potentially important predictor of learning gains in children, was missing from the 2023 data.

As with the regressions employed when answering the first evaluation question, covariates were chosen based on theoretical considerations and availability across the two samples. In addition, it was important to control for PQA score, as the original 2018 evaluation only

---

[34] Total ELOM: $t(743) = 8.36$, $p = 0.000$; GMD: $t(743) = 7.92$, $p = 0.000$; FMCVMI: $t(743) = 6.25$, $p = 0.000$; ENM: $t(743) = 4.54$, $p = 0.000$; CEF: $t(743) = 4.74$, $p = 0.000$; ELL: $t(743) = 2.47$, $p = 0.000$

sampled ELPs with a Green PQA rating, whereas the current evaluation incorporated a more representative sample of ELPs.

The regression findings indicate that after controlling for sample differences, children in the 2023 SmartStart sample exhibited lower ELOM point gains compared to those in the 2018 sample across all domains. Specifically, the 2023 sample was shown, on average, to have gains of 10.23 Total ELOM points lower compared to the 2018 sample (p<0.01). Moreover, within the individual domains, the 2023 sample had ELOM gains between 1.19 points (in ELL) to 3.44 points (in GMD) lower than the 2018 sample, with statistical significance observed across all domains (p<0.01 for all). Notably, certain control variables such as the child-adult ratio (for Total ELOM and ELL), teacher matric qualification (for ENM),[35] fee versus no fee (for Total ELOM and FMCVMI), and endline height-for-age (for Total ELOM and CEF) demonstrated statistically significant coefficients at the 5% significance level.

Table 10: Regression results on the relationship between ELOM gain and sample year while controlling for sample differences

Factors	Total ELOM	GMD	FMCVMI	ENM	CEF	ELL
2023 SmartStart	-10.23***	-3.44***	-2.11***	-1.70	-1.79***	-1.19*
	(1.77)	(0.58)	(0.43)	(0.48)	(0.53)	(0.64)
Female	1.06	-0.11	0.13	0.19	0.19	0.67*
	(1.09)	0.43	(0.32)	(0.35)	(0.34)	(0.39)
Endline age	0.08	-0.01	-0.01	0.02	0.01	0.07
	(0.2)	0.07	(0.05)	(0.06)	(0.06)	(0.08)
Child-adult ratio	-0.25**	-0.09*	-0.05*	-0.01	-0.03	-0.08**
	(0.11)	(0.04)	(0.03)	(0.03)	(0.03)	(0.04)
Has matric	-2.30*	-0.41	-0.24	-0.76	-0.60**	-0.29
	(1.18)	(0.45)	(0.31)	(0.36)	(0.37)	(0.44)
Fee ELP[36]	3.12*	0.36	0.93**	0.03	0.66	1.14*
	(1.64)	(0.57)	(0.43)	(0.48)	(0.48)	(0.59)
Total PQA	-0.01	0.01	-0.02	-0.02	0.00	0.01

---

[35] Notably, coefficients across all domains are counterintuitively negative for this covariate - the negative covariate for ENM suggests ENM programme gain is lower when the practitioner has a matric.
[36] R0; R1-R100; R101-R200; R201-R300; R301+

Factors	Total ELOM	GMD	FMCVMI	ENM	CEF	ELL
	(0.06)	0.02	(0.02)	(0.02)	(0.02)	(0.02)
Endline height-for-age	1.35**	0.29	0.27	0.04	0.44***	0.31
	(0.63)	(0.21)	(0.17)	(0.19)	(0.17)	(0.23)
Endline years in programme	0.12	0.13	-0.07	-0.06	0.06	0.06
	(0.72)	(0.27)	(0.19)	(0.22)	(0.23)	(0.29)
Constant	17.71	6.41	5.73	4.25*	3.39	-2.08
	(12.7)	(4.23)	(3.23)	(3.87)	(3.6)	(4.99)

Notes: *p<0.05, **p<0.01, ***p<0.001. Models report robust standard errors to account for clustering at the ELP level.

Evaluation Questions 3 & 4

Is quality as measured by SmartStart's PQA tool and process, associated with higher ELOM scores?

Which of the various components of the PQA are associated with the biggest gains in ELOM scores?

Total PQA Score

SmartStart provided aggregated PQA total and subdomain scores from 475 ELPs. In addition, SmartStart provided fully disaggregated PQA data, with individual rankings on all 34 PQA items, for 337 ELPs.

The total PQA score for an ELP did not meaningfully predict changes in children's ELOM scores. In other words, there was no clear link between an ELP's PQA score and how much children's ELOM scores improved from baseline to endline. This finding was unexpected, as it seems reasonable to anticipate that there would be a positive correlation between total PQA score and ELOM gains. The scatter plot in Figure 16 implies that the overall PQA score is a weak indicator of programme quality, particularly if programme quality is presumed to be reflected in the observed gains in the ELOM of children attending that ELP.

Figure 16: Association between ELOM gains and total PQA score
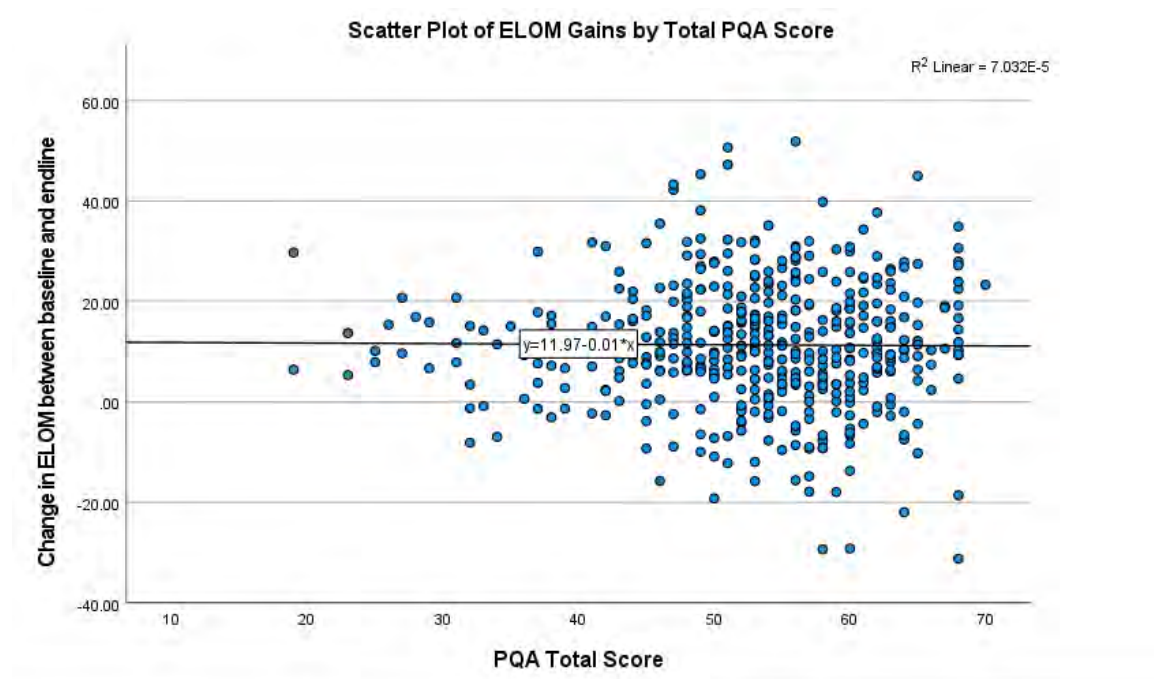
SmartStart uses PQA ratings of Red, Orange and Green depending on cut-off scores on the PQA score. The achievement and maintenance of a Green PQA rating by the ELP is seen as an operationally important benchmark of quality. However, there was no notable variance in ELOM gains observed between baseline and endline for children enrolled in ELPs with Red, Orange or Green PQA ratings based on the PQA. Figure 17 illustrates the relationship between baseline and endline ELOM scores for children in uncategorised ELPs, as well as those with Red, Orange, and Green PQA ratings. In addition to no visible differences in the relationships between baseline and endline ELOM scores across ELPs with different PQA ratings (as shown in Figure 17), a statistical analysis confirmed that the differences in the slopes of the four regression lines were not significant.

Figure 17: Association between baseline and endline ELOM scores by PQA status
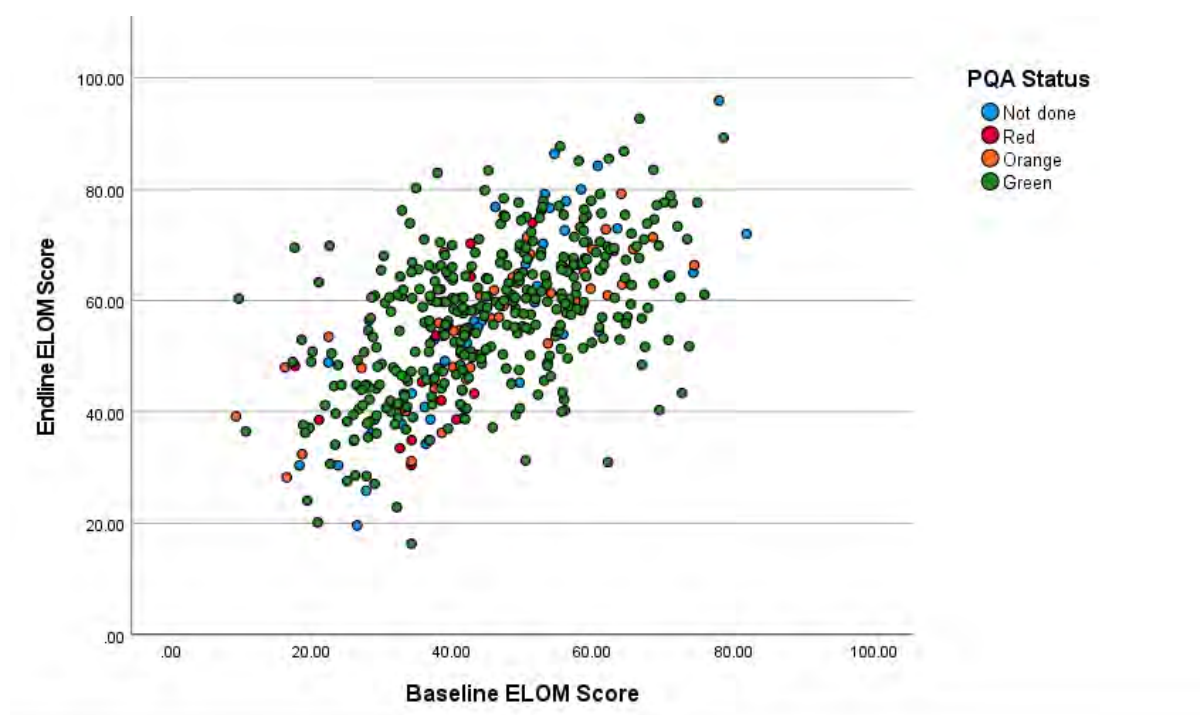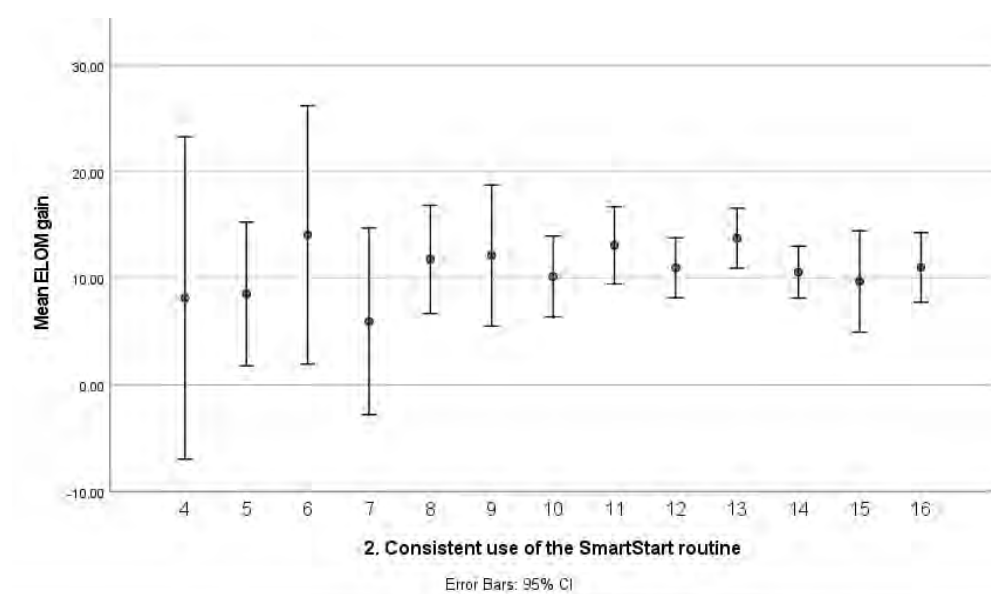

Six PQA Sub-Domains

SmartStart typically breaks the full PQA into six subcategories, and how an ELP scores in each is important for quality assurance. The analysis looked at whether scores on these subscales were linked to gains in children's ELOM scores. The results showed no significant link between an ELP's score on any of the PQA areas and the ELOM improvements between baseline to endline.

Figures 18-23 show that there is no clear link between ELOM gains and scores on any of the six PQA subdomains. If a strong positive relationship existed, we would expect to see ELOM gains increase as PQA scores increase. However, the figures show no such pattern. Additional multivariate analyses (not presented in this report) confirmed that even after accounting for learner-specific factors, none of the PQA subdomains were linked to ELOM gains.

Figure 18: Mean ELOM gains by ELP score on PQA subdomain 1 "A stimulating and adequately resourced learning environment"
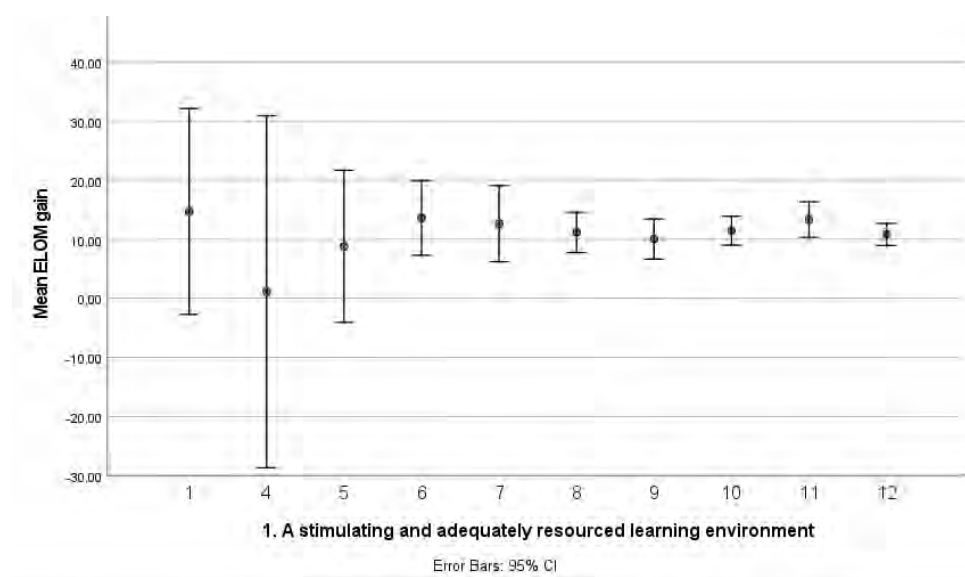
Figure 19: Mean ELOM gains by ELP score on PQA subdomain 2 "Consistent use of the SmartStart routine"

Figure 20: Mean ELOM gains by ELP score on PQA subdomain 3 "A stable and nurturing environment where children feel safe and loved"
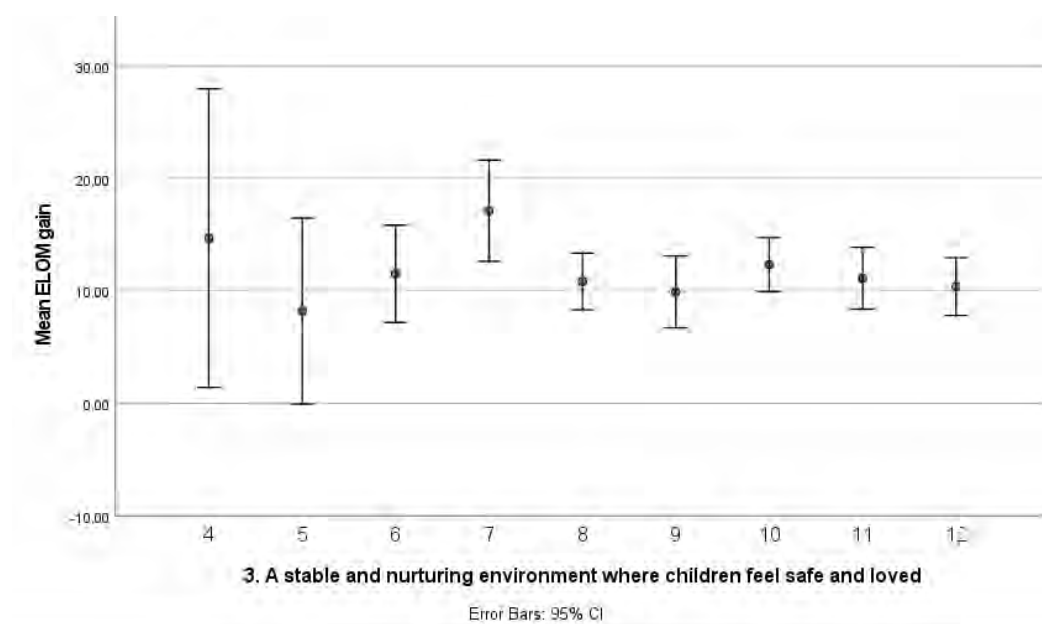
Figure 21: Mean ELOM gains by ELP score on PQA subdomain 4 "Positive and plentiful adult-child interactions which encourage rich use of language"
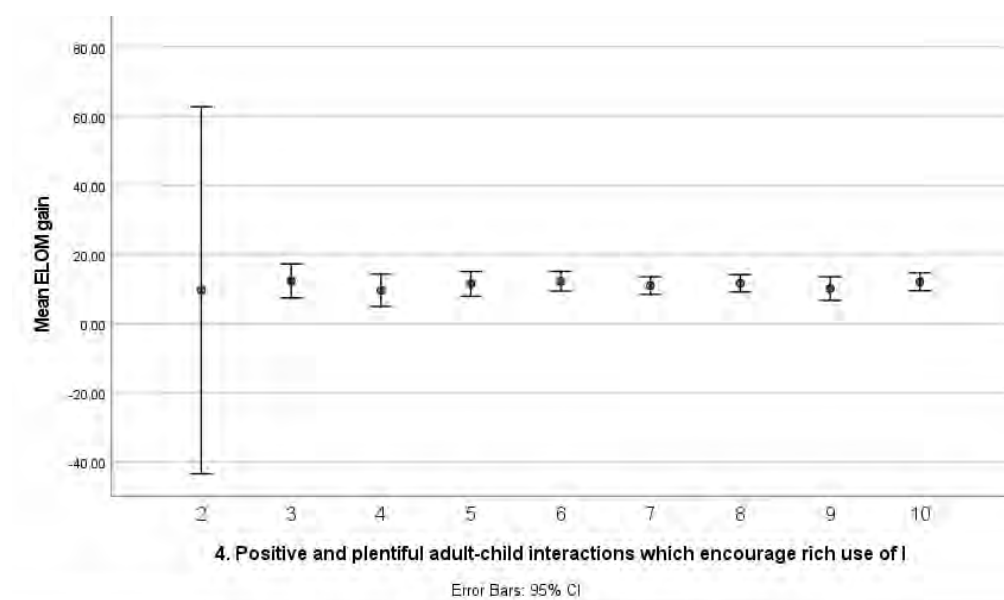
Figure 22: Mean ELOM gains by ELP score on PQA subdomain 5 "Opportunities for child-directed, open-ended play, supported and extended by adults"
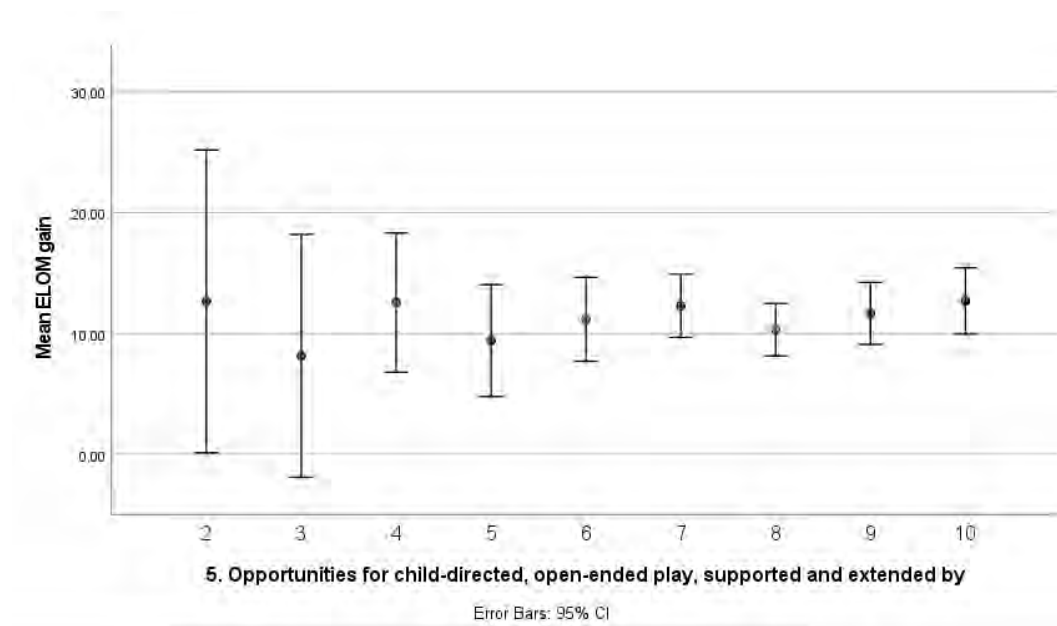
Figure 23: Mean ELOM gains by ELP score on PQA subdomain 6 "Interactive storytelling that introduces children to new language and learning"
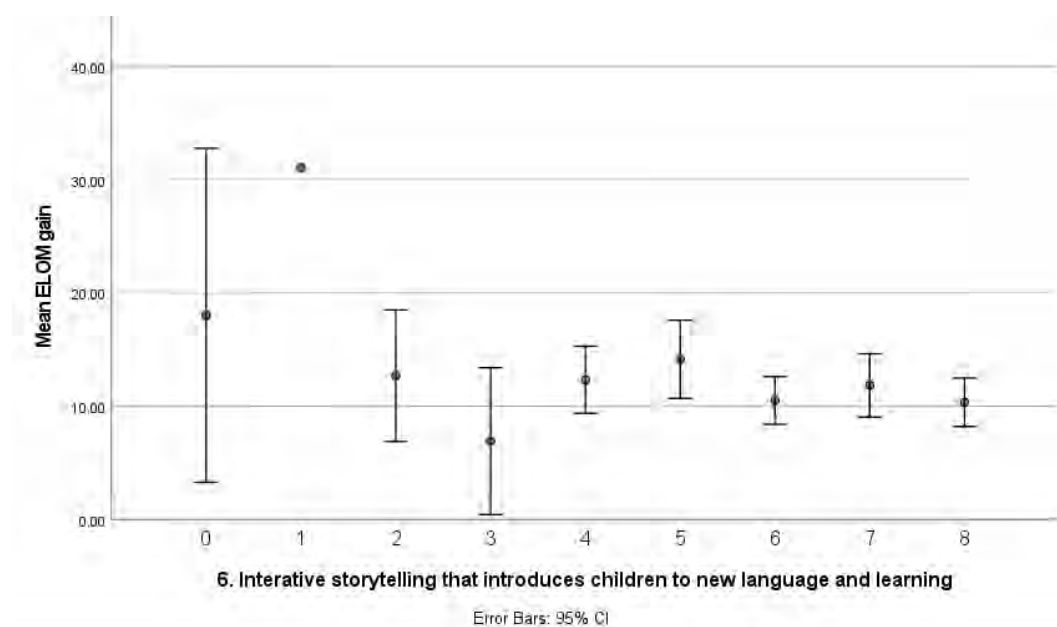
Individual PQA Items

The findings so far show that neither the overall PQA score nor the subdomain scores are significantly linked to children's ELOM gains. However, these results should be interpreted with caution. The absence of an association with these aggregations of the PQA does not necessarily mean that any of the 34 individual PQA items are not associated with ELOM gains. Rather, it is possible that the way the PQA items are currently being summed and aggregated is, potentially, obscuring the effects of specific quality drivers on ELOM gains.

In the subsequent section, we take a closer look by using multivariate analysis on detailed PQA data from 339 ELPs. This approach lets us test whether scores (i.e., 0, 1 or 2) on any of the 34 individual PQA items are linked to ELOM gains, even after accounting for child-specific factors. The goal is to better understand how specific parts of the PQA relate to changes in ELOM scores.

The rankings of the 34 individual items on the PQA are assessed on a scale of 0, 1 or 2 by an observer, where 2 always corresponds to complete achievement of the ranked item, 1 corresponds to partial achievement, and 0 corresponds to inadequate or absent achievement. Unfortunately, exploring differences between ELPs scoring 0 on a specific domain, and those scoring a 1 or 2 on the domain, proved challenging due to the low incidence of 0 scores for most of the 34 items, making reliable statistical analysis difficult. As a solution, the 0 and 1 categories were collapsed to a single category, and ELPs were thus typically re-coded as either high or low scorers on a PQA domain, where a high score corresponds to a 2 ranking, and a low score corresponds to a 0 or 1. This re-coding was required to ensure an adequate sample size across categories for meaningful analysis. However, the analysis for PQA item 3.1, "Warm and respectful interactions," was not feasible due to only 10% of ELPs failing to achieve the maximum rating of 2 in this category. Consequently, item 3.1 was excluded from the analysis.

Table 11 presents findings for the 33 PQA items that exhibited a statistically significant association with ELOM gains.

Table 11: Modelled association of significant PQA items with ELOM gains

PQA Item	Output	Total ELOM	GMD	FMCVMI	ENM	CEF	ELL
1.2 The learning space was divided into three or more interest areas.	b	NS	NS	1.23*	NS	NS	2.08*
	SE			0.54			0.79
	R²			0.11			0.26

PQA Item	Output	Total ELOM	GMD	FMCVMI	ENM	CEF	ELL
	N			233			233
2.1 During the session, between 8-10 SmartStart activities were observed	b	4.57*	2.05*	NS	NS	NS	NS
	SE	1.86	0.9				
	R²	0.34	0.15				
	N	233	233				
2.5 There was substantial time set aside for free play (at least 45-55 minutes)	b	NS	NS	NS	1.4*	NS	NS
	SE				0.56		
	R²				0.2		
	N				233		
2.6 There was enough time set aside for children to recall and reflect on their activities	b	3.22*	NS	0.95*	NS	NS	NS
	SE	1.57		0.43			
	R²	0.35		0.12			
	N	233		233			
2.8 The large group activity lasted for more than 10 minutes	b	NS	NS	NS	1.56**	NS	NS
	SE				0.54		
	R²				0.21		
	N				233		
3.3. Adults acknowledge children's efforts and ideas and offer encouragement and praise	b	NS	NS	NS	NS	NS	1.26*
	SE						0.62
	R²						0.25
	N						231
3.6 Adults actively involve children in solving conflicts, by acknowledging their feelings, listening carefully and trying their solutions.	b	NS	NS	NS	NS	NS	1.9**
	SE						0.64
	R²						0.27
	N						231
4.4 Adults frequently used appropriate methods to build children's language, such as explaining new words or correcting language	b	NS	NS	NS	1.13*	NS	NS
	SE				0.56		
	R²				0.19		
	N				233		
5.5. Adults allowed children to play and	b	4.22**	NS	NS	1.5**	NS	1.61*
	SE	1.55			0.57		0.63

PQA Item	Output	Total ELOM	GMD	FMCVMI	ENM	CEF	ELL
learn at a level and pace which was appropriate to them	R²	0.43			0.2		0.26
	N	233			233		233
6.4 If sharing a storybook, adults did so in ways that helped children become familiar with the books and print	b	NS	NS	NS	1.23*	NS	NS
	SE				0.57		
	R²				0.2		
	N				233		

Notes: *p<0.05, **p<0.01, ***p<0.001. Outcome variable = Gains in ELOM scores. NS = Not Significant. Models control for baseline ELOM, height-for-age, child age, gender, attendance[37], practitioner-to-child ratio, urban vs. rural, number of years attending, education of primary caregiver, summed count of incidents in the household of the following in the last week (reading books, telling stories, singing songs, outdoor activities, playing, naming things, counting, drawing), duration of time the primary caregiver spends playing with the child on both weekends and weekdays. Models report robust standard errors to account for clustering at the ELP level.

Table 11 shows that, even after accounting for factors like the child's home learning environment (HLE), individual characteristics, and key ELPs features (such as practitioner-child ratio and provider education levels), certain practices are still linked to greater ELOM gains. These practices included closely observing SmartStart activities, fostering an environment that promotes recall and reflection within the classroom, and allowing children to engage in play and learning at their own level and pace.

Children in ELPs that made time for recall and reflection (item 2.6) gained, on average, 3.22 more points on the ELOM than those ELPs that did not. Further analyses showed that most of the gains came in the FMCVMI domain. In particular, children in ELPs that strongly supported recall and reflection score an average of 0.95 points higher in FMCVMI than those in ELPs where this practice was lacking.

Children in ELPs where providers tailored play and learning to each child's level and pace (item 5.5) gained 4.22 more points on the ELOM compared to those in ELPs where adults imposed activities that were too easy, too difficult, or not appropriately matched to the child's needs. This ability to adapt to individual capabilities was especially important for gains in ENM and ELL. Children in ELPs with adaptable providers scored, on average, 1.50 points higher in ENM and 1.61 points higher in ELL, highlighting the importance of this practice in fostering positive outcomes.

---

[37] For this model attendance rate data comes from the overall ELP attendance rate reported in the Practitioner Interview. SmartStart collected another attendance dataset that recorded each child's attendance over 2023. See Appendix C for analysis on this dataset.

SmartStart activities recorded by observers (item 2.1) encompass greeting time, message boarding, small group activities, planning time, playtime, clean-up time, recall time, large group time, storytime, and outside time. Children in ELPs that regularly did at least eight of these ten activities gained, on average, 4.57 more points on the ELOM than those in ELPs with fewer activities. This points to a strong link between consistent use of the SmartStart routine and higher overall ELOM gains. More detailed analysis shows that much of this gain came from improvements in the GMD domain, where children in high-performing ELPs scored an average of 2.05 points higher than their peers.

Some individual PQA items were linked to gains in specific ELOM domains, even if they didn't affect overall ELOM scores. For example, when learning spaces were divided into three or more interest areas (item 1.2) children gained on average 1.23 more points on FMCVMI and 2.08 more points on ELL - both highly statistically significant (p<0.01). Setting aside enough time for free play (45-55 minutes, item 2.5) was linked to a 1.40 point gain in ENM. When adults frequently used appropriate methods to build children's language, such as explaining new words or correcting language (item 4.4), ENM gains were 1.13 points higher. Large group activities (item 2.8) were also important: children in ELPs with at least 10 minutes of group time gained 1.56 more points in ENM. Finally, book sharing (item 6.5) was important. Where practitioners made no or limited effort to help children become familiar with the books or print, children gained on average 1.23 less points in ENM.

Two PQA items related to a stable and nurturing environment were particularly associated with gains in ELL. Specifically, children gained on average 1.26 less points on ELL in ELPs where adults rarely or seldom noticed children's efforts or ideas, or tended to be critical (item 3.3). Conflict resolution was also important (item 3.6); In ELPs where adults imposed solutions on conflicts, and did not spend enough time listening to children's accounts, feelings and ideas, children gained on average 1.9 points less in ELL. This finding was highly significant (p<0.01).

Evaluation Question 5

To what extent does SmartStart support affect the PQA quality measures? Which aspects of support are most beneficial?

Coach Support

Practitioners reported an average of 4.92 coach visits to ELPs annually. However, there was significant variance in this, with 26 ELPs (5%) not receiving any visits in the previous year and 67 ELPs (14%) having only one visit. In addition to coach visits, WhatsApp was the most common mode of communication with the coach (86%, n=413) followed by phone calls (78%, n=375).

Insights from the Practitioner Survey point to significant gaps in coaches' operational use of the PQA to enhance quality. Perhaps surprisingly, nearly a quarter (25%) of practitioners were unaware of SmartStart's PQA assessments at their centres. Additionally, more than one-fifth (21%) reported receiving no post-visit feedback, and a similar proportion (22%) highlighted a lack of constructive feedback on improvement opportunities from the coaches.

In examining the association between the frequency of coach support and the resultant PQA scores among ELPs, our analysis identified a significant relationship between the frequency of coach visits in the prior year and the PQA scores of ELPs [$F_{(6,6)} = 2.20$, $p=0.042$]. ELPs benefitting from just one or fewer visits annually had PQA scores 4.48 points lower (95% CI = 1.29-7.67) than those with two visits per year ($p=0.006$). Moreover, ELPs benefitting from 11-24 visits annually demonstrated a 4.15-point increase in PQA scores (95% CI = 1.12-7.12) compared to programmes with zero or one visit ($p=0.007$). These findings indicate that coach visits play a role in enhancing PQA scores, underlining the importance of regular visits, preferably on a monthly basis.

Figure 24: Mean PQA score attained by an ELP by frequency of coach visits to the ELP in the preceding year
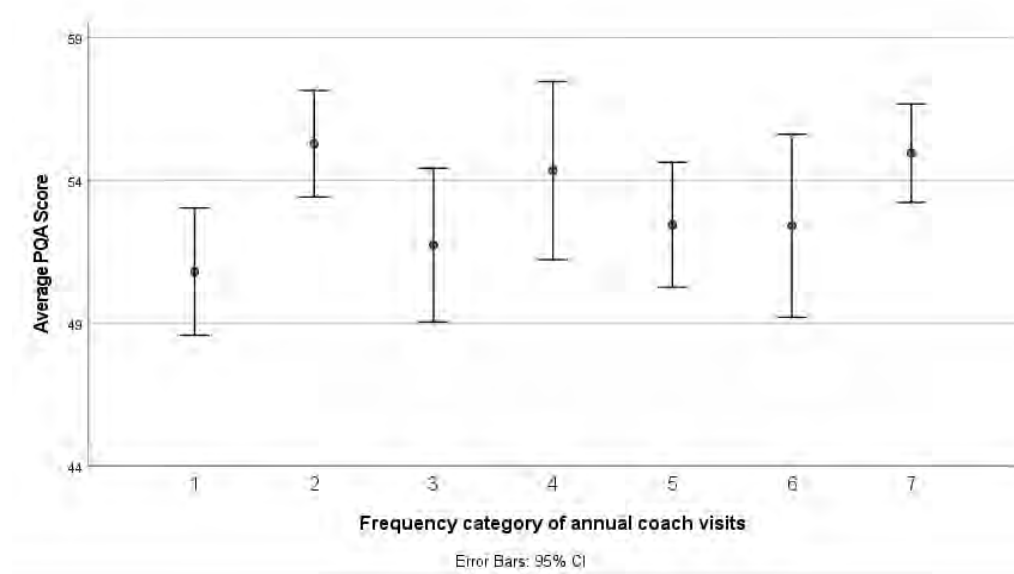
Note: Categories are septiles corresponding to 1=0-1 visits a year; 2=2 visits a year; 3=3 visits a year; 4=4 visits a year; 5=5-6 visits a year; 6=7-10 visits a year; 7=11-24 visits a year.

There was a significant association between the frequency of coach visits in the preceding year and the ELOM gains [F(6,6) = 2.62, p=0.016]. This association was only significant, however, when comparing ELPs with one or fewer visits annually to those receiving 11-24 visits, with the latter group achieving an average of 6.14 additional ELOM points (95% CI = 2.39 - 9.89) compared to the former (Figure 25). These findings suggest that regular coach visits, at least monthly, are necessary if coach visits are to have a measurable effect on ELOM gains.

Figure 25: Mean ELOM gains by frequency category of coach visits to the ELP in the preceding year
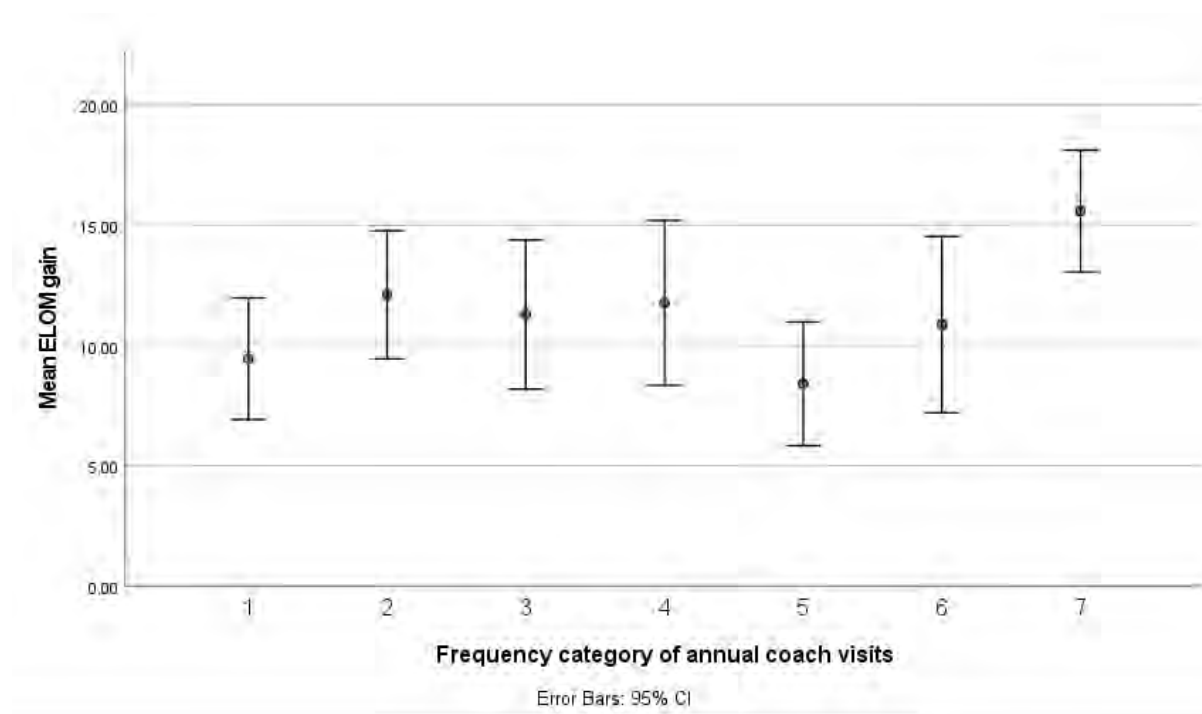
Note: Categories are septiles corresponding to 1=0-1 visits a year; 2=2 visits a year; 3=3 visits a year; 4=4 visits a year; 5=5-6 visits a year; 6=7-10 visits a year; 7=11-24 visits a year.

The quality of coach visits was as important as how often they occurred in influencing ELOM gains. Practitioner satisfaction with coaching effectiveness showed a significant link with ELOM improvements [R^2 = 0.01, $F(1, 274)$ = 4.80, p = .029]. On average, ELPs where coaching was seen as "not at all useful" showed an 8.37 point gain on the ELOM, while those where visits were rated as "extremely useful" saw a higher average gain of 12.45 points.

Logistic regression analyses were further used to explore which aspects of coach support contributed to a practitioner's perception of its "usefulness". In the survey, practitioners rated various dimensions of the SmartStart coaching role on a five-point Likert scale. These dimensions included statements on guidance for age-appropriate activities, techniques for creating stimulating environments, constructive feedback, classroom observation, assistance with child assessments and planning, fostering relationships with parents, organising family days, running parent meetings, and involving families in learning. The findings pinpointed two key attributes significantly associated with positive usefulness ratings: providing practical techniques for stimulating environments and assisting in child assessments and planning, which increased a coach's "usefulness" odds by 145% (95% CI = 1.19 - 5.06) and 89% (95% CI = 1.14 - 3.12), respectively, for each point increase in agreement.

Club Attendance

A significant majority (95%) of practitioners were members of active clubs, with most of these clubs (81%) convening on a monthly basis. In addition, 23% (n=109) of practitioners were club leaders. Practitioners generally viewed these club meetings positively, with a minority (8%) rating them as not at all useful or only slightly useful.

Although club participation was generally viewed positively, our analysis found no meaningful link between it and improvements in ELOM scores. There was also no clear relationship between how frequently practitioners attended meetings, how useful they found them, and any changes in ELOM scores. Additionally, the frequency of club meetings and the practitioners' assessments of their usefulness were not linked to either the overall PQA scores or any of its six subdomains. Overall, while club participation is common and valued, it does not appear to be linked to measurable gains in ELOM or PQA scores.

5 | Recommendations

Capitalise on ELOM and ENM Successes

SmartStart children showed significant and meaningful ELOM gains between the baseline and endline assessments. After eight months, there was a 20% increase in children 'On track' relative to baseline. SmartStart also showed a significantly higher proportion of children 'On track' (65%) compared to the most recent national benchmark, the Thrive by Five 2021 Index data (46%). Furthermore, regression analysis showed SmartStart children scored better than the national benchmark across all domains with the biggest gains in ELL and ENM.

SmartStart children have performed particularly well in terms of ENM. Programme gains as a percentage of overall ELOM gains were notably highest for ENM (64%), and ENM scores of SmartStart children were 24.6 percentage points higher than the national benchmark average. In addition, ENM was the only domain that did not decline in terms of the proportion of children 'On track' between the 2018/2019 evaluation, and the current evaluation.

Given the lack of a statistically significant link between overall PQA scores and ELOM gains, we recommend revisiting how PQA items are aggregated to better reflect their specific contributions to child development outcomes. Enhanced training for coaches is also needed to ensure accurate implementation and scoring, addressing any reluctance to use a 0 score where appropriate. Improved moderation systems, including regular calibration sessions and periodic audits, should be developed to ensure scoring consistency. A detailed analysis of the 34 individual PQA items has also identified those practices most strongly associated with ELOM gains, allowing for targeted interventions. Finally, establishing an ongoing feedback mechanism for coaches will support continuous refinement of training, moderation, and assessment practices, ultimately improving the overall quality and effectiveness of the SmartStart programme and ensuring better developmental outcomes for children.

Strong performance in ENM is associated with specific classroom practices. SmartStarters who are doing particularly well in terms of ENM are, on the whole, significantly more likely to be ensuring there was substantial time set aside for free play (at least 45-55 minutes), ensuring the large group activity lasted for more than 10 minutes, using appropriate

methods to build children's language, allowing children to play and learn at a level and pace which was appropriate to them and, to be reading storybooks effectively.

Notwithstanding these gains, the average improvement in ELOM scores from baseline to endline in the 2023 sample was smaller than in the 2018 sample. In fact, the total ELOM gain in 2023 was nearly half of what was recorded in 2018. The largest declines in gains were seen in the GMD and FMCVMI domains. Note, however, that:

- The 2023 sample still had a greater proportion of children 'On track' at the endline (65% vs 59%), and so the decrease in ELOM gains should not be necessarily regarded as a decrease in programme effectiveness. Given the higher average baseline scores for the 2023 cohort (especially in GMD), it is likely that at least some of the apparent "decline" in ELOM gain relative to five years ago is due to diminishing returns influencing the results.

- It is also critical to note that the sampling approaches in the 2018/2019 evaluation were very different to the sampling approach adopted in the current evaluation, making simple comparisons between the two evaluations potentially misleading. Specifically, the 2018/2019 evaluation only sampled ELPs with a Green PQA rating, whereas the current evaluation used a representative sample of ELPs, and thus included ELPs with Orange and Red PQA ratings. It is thus possible that the differences in ELOM gains can be attributed to differences in programme quality. That said, multivariate regressions were conducted to control for differences in the PQA and ELP characteristics between evaluation samples. These regression findings indicated that even after controlling for many sample differences, across all domains children in the 2023 SmartStart sample still exhibited lower ELOM point gains compared to those in the 2018 sample.

Refine the PQA Tool Usage

One of the important questions SmartStart had for the evaluation team was how the PQA is functioning as a quality assurance tool. Our analysis indicated that several individual PQA items are meaningfully associated with ELOM gains, suggesting that the tool captures important aspects of quality. However, the current methods of aggregating these items - including potentially the use of total PQA score as a benchmark for accreditation - does not align well with actual ELOM outcomes. Likewise, the summed scores across the six PQA subdomains show no correlation with actual ELOM gains. This suggests that the current

approach to grouping and aggregating items into these six subdomains should be carefully evaluated.

Specific individual items on the PQA were significant indicators of ELOM gains. Key observations that appeared to be particularly relevant to gains in total ELOM scores were whether adults allowed children to play and learn at a level and pace which was appropriate to them, whether there had been enough time set aside for children to recall and reflect on their activities, and ensuring that during the session, between 8-10 SmartStart activities were observed. These items should be prioritised in SmartStart monitoring, feedback, and quality assurance based on these reflections. Coaches should also make an effort to provide SmartStarters with more feedback on their performance in these areas, as well as general feedback on performance in the PQA overall.

Emphasise Quality and Quantity of Coach Visits

Coach visits and their quality were significantly associated with both the total ELOM gains and total PQA. However, this positive association only became evident when coach visits were delivered at least monthly. An increased number of coach visits, such as monthly, should be considered an aspirational benchmark for effective coach support and increased ELOM gains.

Coaches should also make efforts to provide practical techniques for stimulating environments and assist in child assessments and planning, as these were the key attributes associated with positive coach ratings and subsequent ELOM gains.

Further Explore the Role of Club Attendance

Clubs meet monthly to provide peer support and share information, helping with tasks like setting fees, managing conflicts, and organising events. While some clubs also offer programme support, such as online training and planning monthly themes, this is secondary and voluntary. Although not primarily intended for programme administration, clubs sometimes handle admin tasks and raise concerns about peers. Many clubs also contribute to sustainability support planning.

Club attendance was highly valued and widespread, and club attendance may perform many important functions around peer support and retention that were not explicitly measured in this evaluation. However, there were no direct significant associations between club attendance variables, PQA and ELOM gains.

It is therefore recommended that SmartStart should:

- Continue to qualitatively unpack the benefit of club attendance and document how club attendance is supporting SmartStart in ways not directly related to the ELOM and/or the PQA.

- Make informed operational decisions to balance the benefits of both club and coach components of the programme.

6 | Conclusion

In closing, this evaluation illustrates how SmartStart's social franchise model can effectively reach large numbers of children with high-quality early learning support, while simultaneously stimulating local employment in under-resourced communities. The findings indicate that rigorous quality assurance, practitioner training and community engagement are central to sustaining meaningful outcomes at scale. They also highlight areas where further refinements could strengthen delivery and deepen impact. As the global debate on expanding ECD services continues, SmartStart's experience provides valuable lessons about balancing reach with quality, mobilising local assets and empowering practitioners. Ultimately, the evidence gathered here underscores the importance of innovative models that bridge persistent gaps in access to - and quality of - early childhood learning provision in developing countries.

The findings indicate that, on average, children participating in the SmartStart programme are achieving better outcomes than those achieved by other children in South Africa. The 2023 SmartStart sample outperformed the 2021 Thrive by Five Index across all domains, even after adjusting for sample differences. In addition, the findings suggest that **SmartStart's programme is reducing disparities in early learning outcomes between children from low-income and higher-income backgrounds.** This result is particularly crucial within the South African context, not only from a rights and equity perspective, but also because achievement across the education system will only improve when every child enters school with a comparable starting-point.

The children in SmartStart's 2023 evaluation experienced lower baseline-to-endline gains in ELOM scores compared to those in the 2018/2019 SmartStart evaluation sample. However, the proportion of children 'On track' at endline was higher in the 2023 sample. These differences in outcomes between the 2018 and 2023 samples are not necessarily a result of the larger network of programmes in 2023, but may be attributable to other sample differences, such as omitted variable bias (some variables present in the 2018 evaluation were not available in the 2023 evaluation) or genuine differences in performance. These findings should thus be interpreted with this limitation in mind.

Specific PQA components were shown to be strongly correlated with gains in ELOM scores. These findings offer valuable insights for refining programme implementation and strengthening both PQA design and practitioner training. For example, the strong positive association between consistent implementation of the SmartStart Routine and higher total

ELOM scores is particularly noteworthy. This finding demonstrates that even in low-resource home- and community-based settings, trained paraprofessionals can make a significant difference for young children when using a carefully designed daily programme.

Interestingly, overall PQA scores or subdomain scores were not significantly linked to ELOM gains. This suggests the need to recalibrate how PQA scores are aggregated and used for quality monitoring and accreditation purposes, ensuring that the tool better aligns with measurable improvements in child outcomes.

The evaluation underscores the critical role of SmartStart coaches in programme success. More frequent and high-quality coach visits were strongly associated with improved PQA scores and greater ELOM gains. This finding suggests that increasing both the frequency and the effectiveness of coach visits could further enhance programme impact. Thus, greater attention to the coach-ELP ratio and enhanced training for coaches on the most impactful types of support - such as providing practical guidance on creating stimulating learning environments and supporting assessment and planning - will be essential in strengthening the SmartStart model as the network expands further.

Although club participation was widespread and highly valued by practitioners, quantitative analysis did not establish a direct link between club attendance and ELOM or PQA improvements. This does not diminish the importance of clubs as a mechanism for professional support and retention, but rather indicates that their benefits may not be easily captured through the specific quantitative measures used in this evaluation. Future research should explore qualitative aspects of club participation to better understand its broader impact on practitioner motivation, professional growth, and long-term engagement with SmartStart.

The findings on the mechanisms that lead to observed gains in child outcomes provide critical validation that features of SmartStart's programme design are effectively improving outcomes for children, even in the most under-resourced communities. Furthermore, these impactful programme features - such as frequent coach visits or the faithful implementation of SmartStart's daily routine - were found to be simple practices, underscoring that the value of a programme is not necessarily determined by expensive facilities and resources, but rather by what the programme offers and how it is supported. This insight can encourage policymakers and those designing interventions to think more innovatively about the potential contributions of paraprofessionals, micro-credentialing, and the core components of a supportive ecosystem.

When these findings are considered in conjunction with the unique design and architecture of the SmartStart platform - which serves as an enabler of an ecosystem involving partners, practitioners, families, and communities - they carry significant implications for strategies aimed at closing the early learning access gap. This is partly because SmartStart leverages resources that already exist in communities, such as homes and under-employed women, making its approach both affordable and rapidly implementable. Therefore, the SmartStart delivery method should inform funding and policy solutions for equitable access, including the development of more inclusive regulatory frameworks, to reach a larger number of vulnerable children in the near term. Moreover, by placing networking and collaboration at the core of the platform, SmartStart is strategically positioned to partner with critical stakeholders, such as the government, to reach every child.

In conclusion, SmartStart's model has demonstrated its ability to deliver high-quality early learning experiences at scale, making a meaningful difference in the lives of young children across South Africa. By continuously refining its approach based on evidence from evaluations such as this one, SmartStart is well-positioned to sustain and expand its impact in the years to come. Furthermore, the study's results indicate that programmes like SmartStart, that offer simple, inexpensive interventions and support mechanisms to early learning practitioners, can effectively improve the quality of early learning in South Africa and beyond.

Technical Appendices

Appendix A

The following tables present the average SmartStart programme for total ELOM score and the domains disaggregated by rural-urban setting and fee category. First looking at urban-rural, Table 12 does not provide strong evidence that programme gain is impacted by whether the ELP is in an urban or rural setting, with the difference in programme gain as a percent of overall gain for total ELOM just three percentage points. Between domains, the differences are similarly modest apart from GMD (15 percentage points) and ELL (21 percentage points).

Table 12: Programme gain, disaggregated by urban-rural setting

Setting		Total	GMD	FMCVMI	ENM	CEF	ELL
Urban	Baseline score (average)	46.0	8.1	11.0	8.8	6.8	11.2
	Endline score (average)	57.4	10.1	13.4	11.4	9.3	13.3
	ELOM point gain	+11.4	+2	+2.3	+2.6	+2.4	+2
	Maturation effect (average)	8.1	1.8	1.8	0.9	2.0	1.6
	Programme gain (ELOM score)	3.1	0.2	0.5	1.6	0.4	0.4
	Programme gain (% of overall gain)	27%	9%	22%	62%	18%	19%

Setting		Total	GMD	FMCVMI	ENM	CEF	ELL
Rural	Baseline score (average)	42.6	8.5	10.8	8.0	5.3	9.9
	Endline score (average)	54.5	10.2	13.1	10.5	7.9	12.8
	ELOM point gain	+12	+1.7	+2.3	+2.5	+2.5	+2.9
	Maturation effect (average)	8.2	1.8	1.8	1.0	2.0	1.7
	Programme gain (ELOM score)	3.6	-0.1	0.5	1.5	0.5	1.2
	Programme gain (% of overall gain)	30%	-6%	21%	61%	21%	40%

Table 13 presents programme gain disaggregated by ELP fee category. The fee category is also a useful proxy for socio-economic status, potentially highlighting differences in programme effectiveness in different socio-economic environments. Overall, there is a slight negative relationship between fee amounts and programme gain, with the correlation coefficient for ELOM Total programme gain and fee amounts statistically significant only at the 10% level. This negative relationship between fees and programme gain stems from ENM (ρ=-0.07; p=0.09) and ELL (ρ=-0.10; p=0.02).

**Table 13: Correlation coefficients between programme gain and ELP fee amount**

Category	Corr. coefficient
ELOM Total	-0.08*
GMD	-0.03
FMCVMI	0.03
ENM	-0.07*
CEF	-0.05
ELL	-0.10**

Notes: *p<0.05, **p<0.01, ***p<0.001

Appendix B

Provider-to-child ratio was significantly associated with ELOM gains (p=0.012). Figure 26 below shows predicted ELOM gains by provider-to-child ratio after adjusting for baseline ELOM values, height-for-age, child age, child gender, the education level of the practitioner, rural/urban location of the ELP, and child attendance.

Figure 26: Predicted ELOM gains by provider-to-child ratio
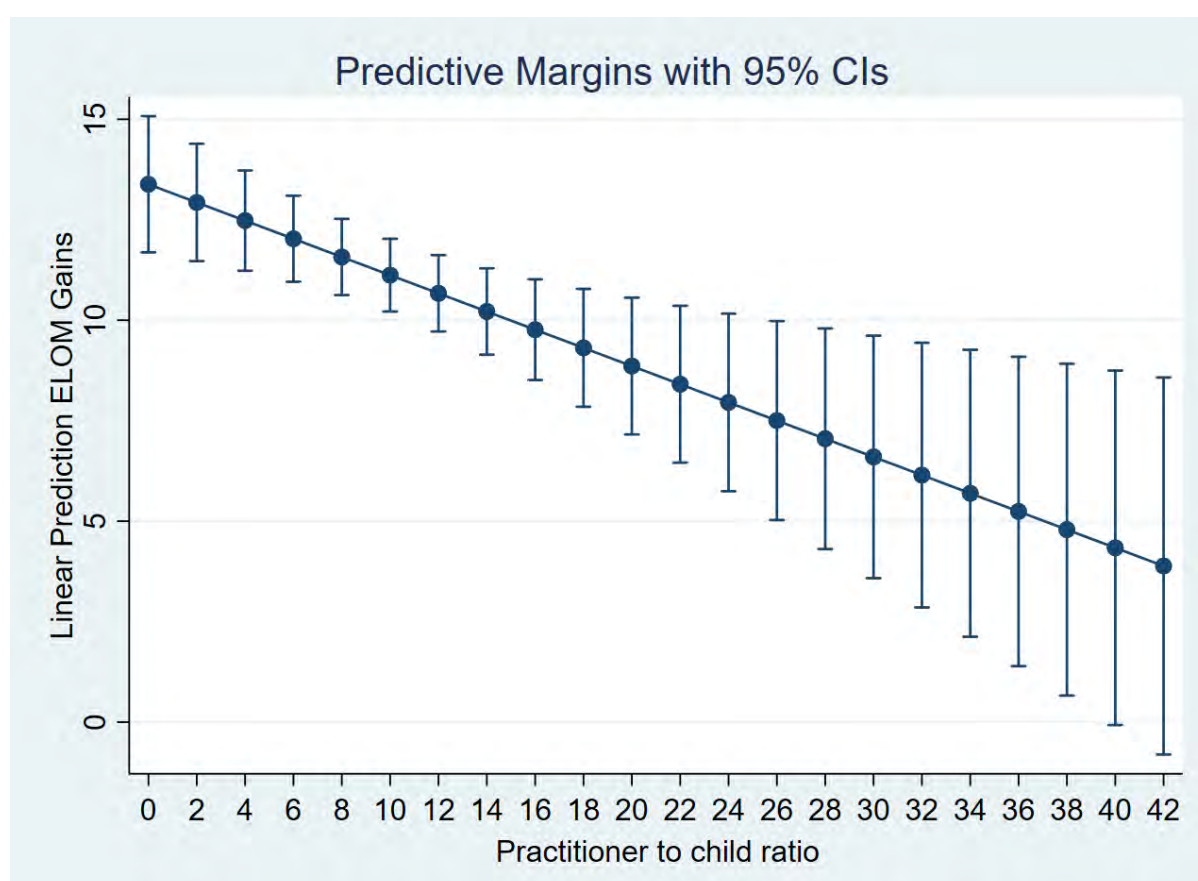
Appendix C

Child attendance is a potentially important factor for measuring SmartStart's impact. Children with higher attendance rates would experience higher 'doses' of the SmartStart programme which would theoretically mean better ELOM scores. The evaluation team had access to two sources of attendance data. The first source was ELP attendance data collected by SmartStart between January to August 2023. The dataset comprised 3797 children attending one of 375 SmartStart ELPs and captured each child's daily attendance over the period. The second source of data was a proxy measure of attendance, where the ELP practitioner was asked during the practitioner interview to estimate how many days a week (i.e. 1-5) the children in our study typically attend the ELP. We performed analysis on both of these sources of data, which is presented below.

Child Daily Attendance Data

Figure 27 below shows that the sample of daily attendance data we were provided with showed that reported attendance rates were high and offered little variation. Over 62% of children had 100% attendance, almost 95% of children had an attendance rate over 90%, and no children had an attendance rate below 60%.

Figure 27: SmartStart children attendance rates: January-August 2023
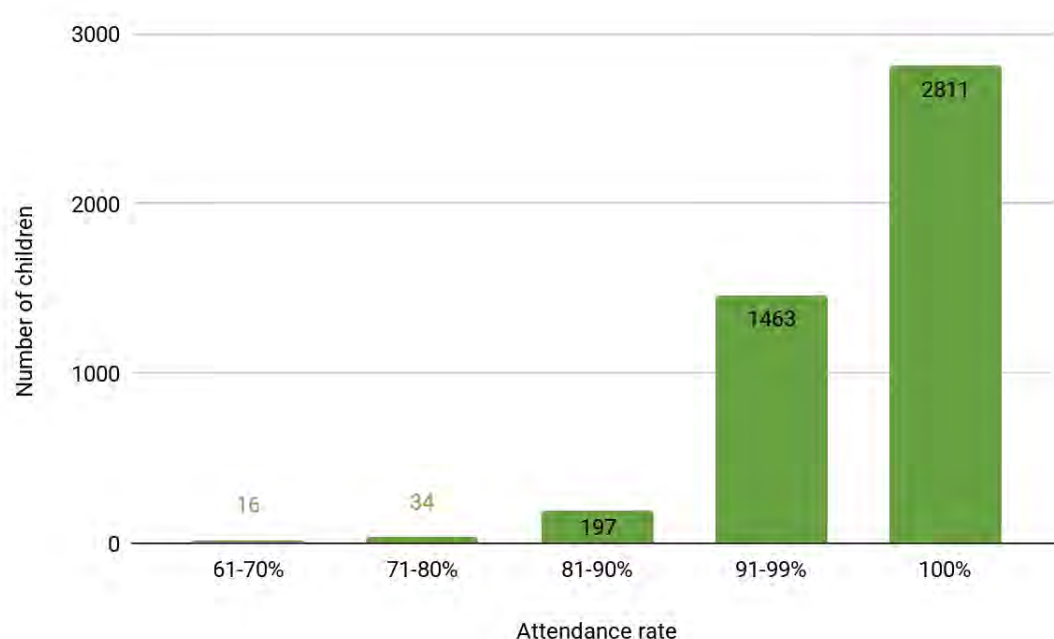
While the attendance data doesn't have an ID variable that would allow us to easily match each child's ELOM data to their attendance data, we can match facilities across the two datasets.[38] Figure 28 below further illustrates the lack of variation in attendance rates within the sample's ELPs. Over 70% of ELPs have attendance rates greater than 95%, while 97% of ELPs have attendance rates over 90%.

Figure 28: SmartStart ELP average attendance rates: January-August 2023
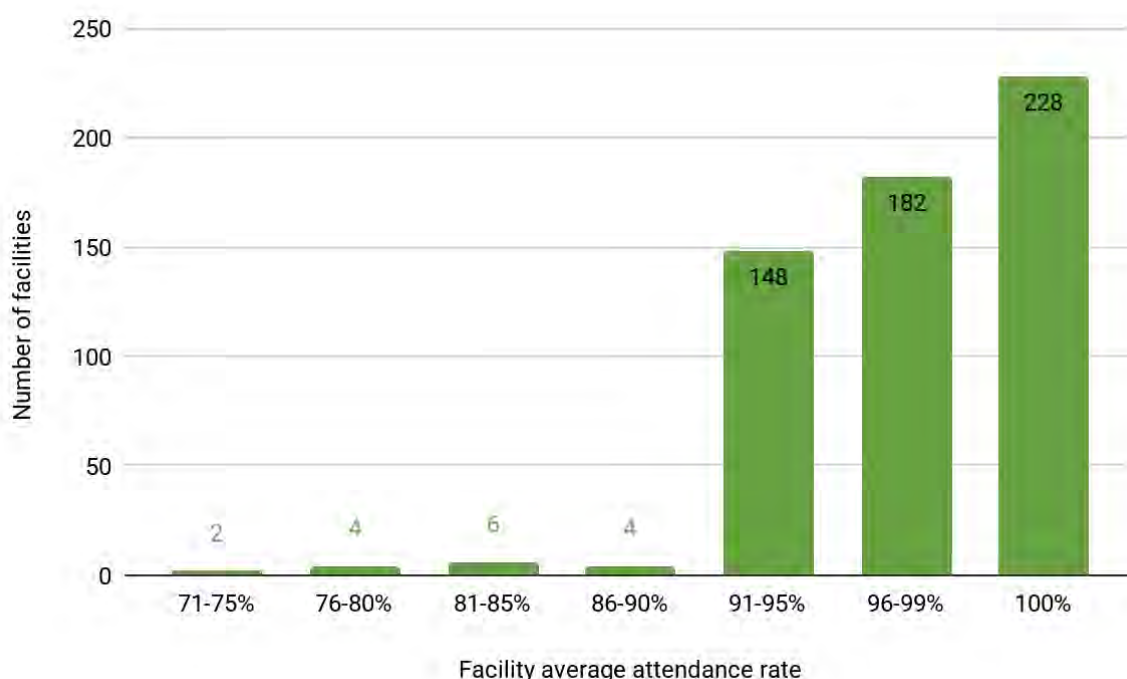
Low variation in ELP attendance rates makes it unlikely we find a statistically significant correlation between attendance and programme gain. The scatter plot in Figure 29, where each dot represents an ELP and the line represents the line-of-best-fit, fails to show a significant relationship between average ELP attendance rate and average ELOM programme gain.[39] Investigating the relationship between average ELP attendance rates and programme gain for each domain yields similarly statistically insignificant results. Resultantly, incorporating this attendance rate data into the statistical models used in this report would be very unlikely to add anything useful.

---

[38] Both the attendance and ELOM data have a 'child name' variable that could theoretically be used for matching but due to spelling differences and missing data, just 16% of children in the ELOM data could be matched to a child in the attendance data.
[39] $\rho(282)=0.03$, $p=0.680$

Figure 29: SmartStart ELP Total ELOM programme gain versus average attendance
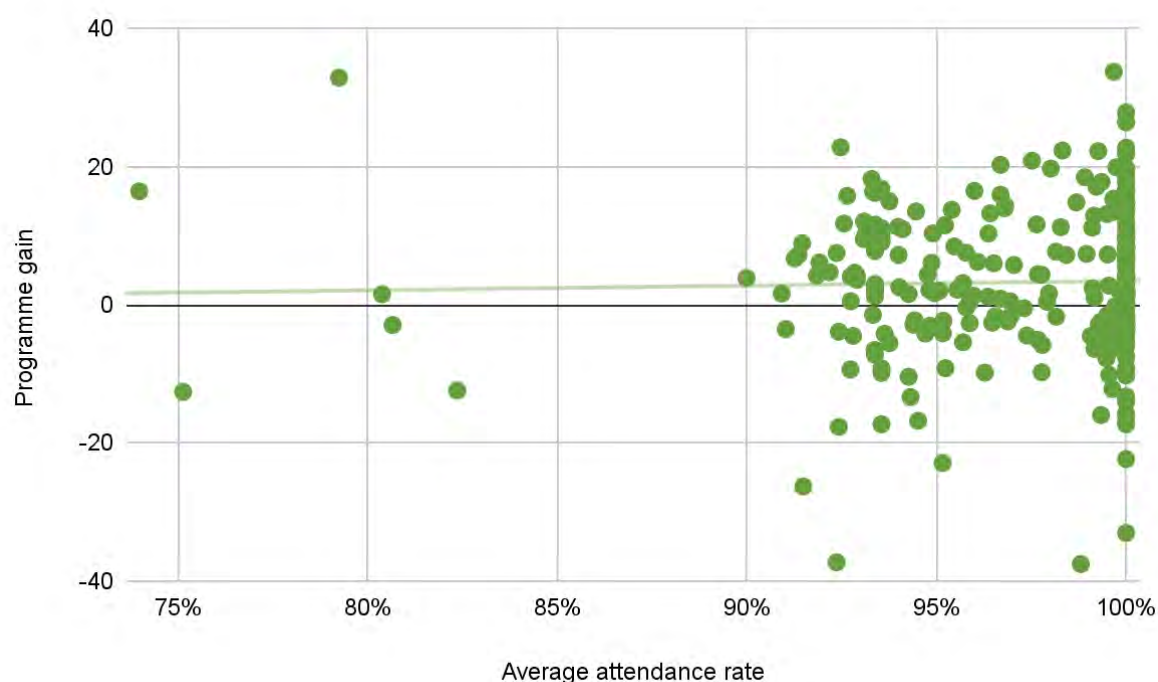
Proxy Attendance Data

Practitioners estimated that 85% (n=470) of the children in our sample attended the ELP five days a week, with 15% (n=38) attending four days a week, 8% (n=12) attending three days a week, 5% (n=29) attending two days a week, and only 0.4% (n=2) attending one day a week. Unlike the daily attendance data, this variable capturing proxy attendance was significantly associated with ELOM gains. After adjusting for baseline ELOM values, height-for-age, child age, child gender, the education level of the practitioner, rural/urban location of the ELP, and adult-child ratio in the ELP, children who attended an ELP more frequently were predicted to have higher ELOM gains (Figure 30). Specifically, children who the practitioner estimated attended one, two or three days a week were predicted to gain 6.21 ELOM points (p=0.007), whereas children who attended four or five days a week were predicted to gain 11.42 ELOM points (p<0.001). As a result of its better statistical properties, this proxy measure of attendance was included in all the multivariate regression models in this report. Thus, the estimates in this report can be said to adjust for attendance as measured by this variable.

Figure 30: Predicted ELOM gains by typical attendance (days a week)


Predictive Margins with 95% CIs
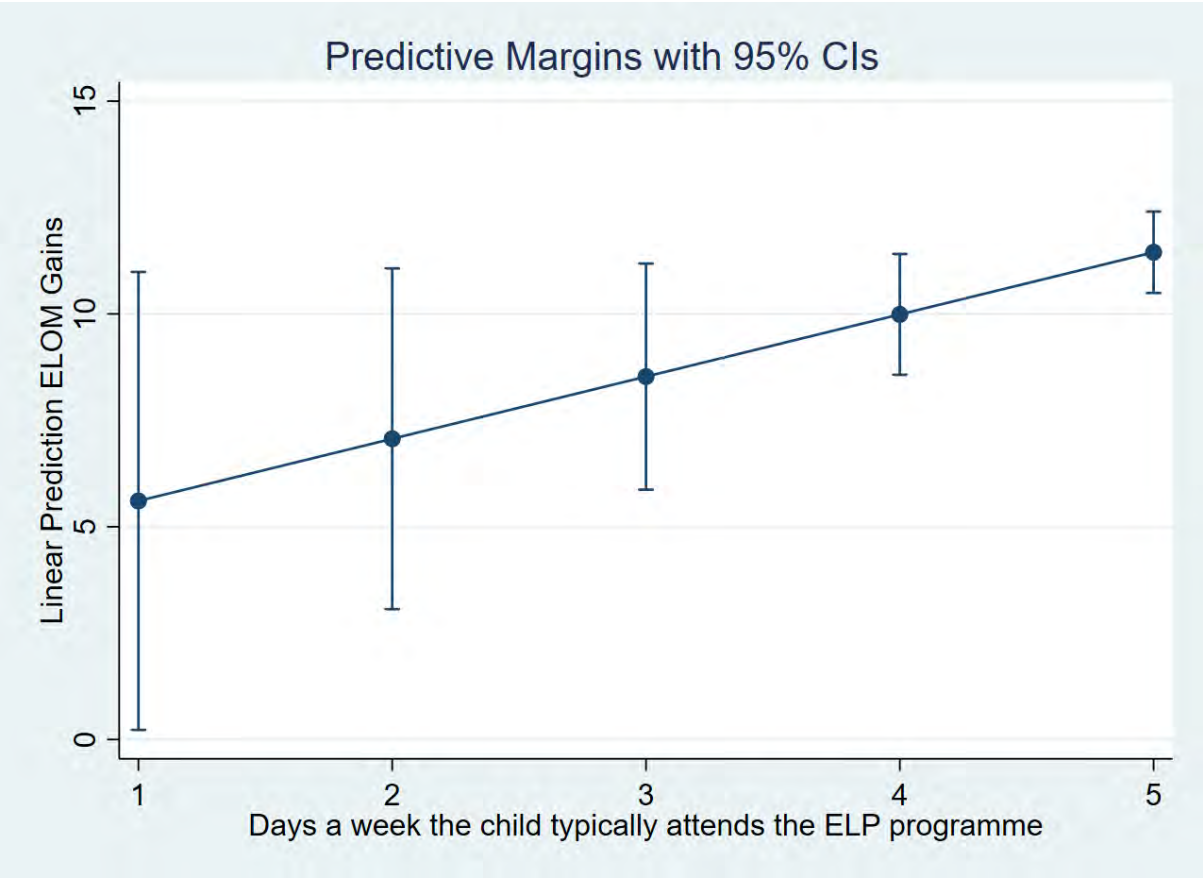
Appendix D

The data indicate a marked improvement in student performance from baseline to endline across all quintiles (Figures 31 and 32). The percentage of students 'On track' has significantly increased, particularly in the lowest quintile (Q1), where the number of students 'On track' jumped from 43% to 64%. Similarly, the middle quintiles (Q2-3) and higher quintiles (Q4-5) also saw improvements in the percentages of students 'On track' and reductions in those 'Falling far behind' or falling behind.

The mean ELOM scores further underscore this positive trend. Quintile 1 saw its mean score rise from 44-55, Quintiles 2-3 from 43-56, and Quintiles 4-5 from 51-60. The overall mean score improved to 56.12 at endline, indicating that students are performing better across the board.

These findings suggest that the interventions and support mechanisms implemented have effectively enhanced student performance, particularly for those in the lower quintiles. The overall improvements in both performance categories and mean scores highlight the success of the initiatives in place, fostering better educational outcomes for the students.

Figure 31: ELOM at baseline per quintile
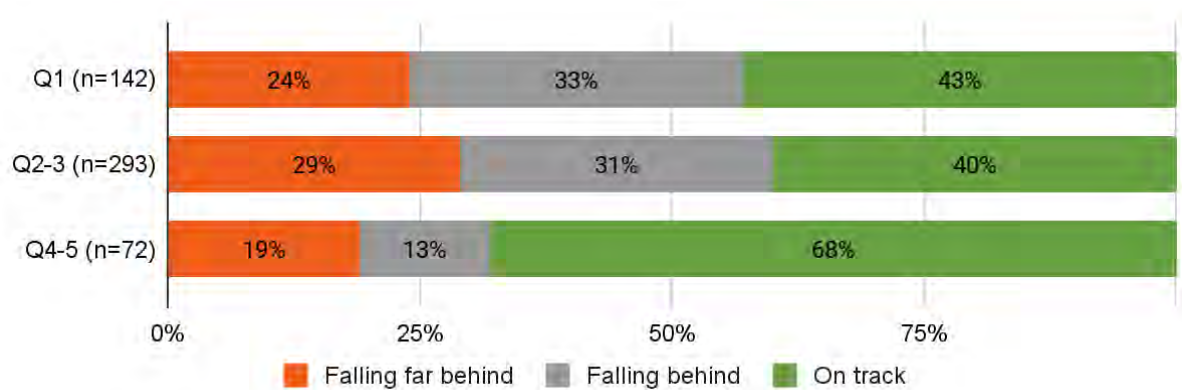
Figure 32: ELOM at endline per quintile