



Social Investment Agency
Toi Hau Tāngata

Foundations for Learning

Descriptive Analysis of Early Childhood Education Participation and
Developmental Outcomes

March 2026

Te Kāwanatanga o Aotearoa
New Zealand Government



Author

Charlotte Rose

Senior Analyst

Data, Evidence and Insights

charlotte.rose@sia.govt.nz

Acknowledgements

Beth Charlton, Penny Mok, Andrew Webber, Stats NZ IDI Output Checking Team

Creative Commons Licence



This work is licensed under the Creative Commons Attribution 4.0 International licence. In essence, you are free to copy, distribute and adapt the work, as long as you attribute the work to the Crown and abide by the other licence terms. Use the wording 'Social Investment Agency' in your attribution, not the Social Investment Agency logo.

To view a copy of this licence, visit creativecommons.org/licenses/by/4.0.

Liability

While all care and diligence has been used in processing, analysing and extracting data and information in this publication, the Social Investment Agency gives no warranty it is error free and will not be liable for any loss or damage suffered by the use directly, or indirectly, of the information in this publication.

Citation

Social Investment Agency 2026. Foundations for Learning: Descriptive Analysis of ECE Participation and Developmental Outcomes. Wellington, New Zealand.

Table of Contents

Executive summary	4
Context	4
Methodology	5
B4 School Check (B4SC)	5
Early Childhood Education (ECE) participation	6
Population definition	6
Validation	6
Exclusions	6
Descriptive analysis	7
Definitions	7
IDI Disclaimer	7
Strengths and Difficulties Questionnaire (SDQ)	8
Demographics	8
Descriptive statistics – Early life experiences	9
Descriptive statistics – ECE participation	9
Children’s scores across the two measures	10
Parental Evaluation of Development Status (PEDS)	11
Demographics	11
Descriptive statistics – Early life experiences	12
Descriptive statistics – ECE participation	12
Summary	13
Next steps - regression methodology	13
Limitations	14

Executive summary

This analysis uses Stats NZ Integrated Data Infrastructure (IDI) data to explore how children's developmental, emotional, and behavioural outcomes at the B4 School Check vary across different groups.

By comparing children based on their Strengths and Difficulties Questionnaire (parent completed) (SDQ) and the Parental Evaluation of Developmental Status (PEDS) results, the descriptive findings highlight some key findings where we see **higher levels of concern** among:

- Children living in greater deprivation
- Māori and Pasifika children
- Male children
- Children with recorded early adversity

We also see that the two measures are independently important for capturing the full picture of parental concerns, with some issues identified in the PEDS but not the SDQ, and vice versa.

Early Childhood Education (ECE) participation shows some variation across groups, where children with higher identified levels of concern are, at age 3, **less likely** to:

- Attend Early Childhood Education*
- Attend Early Childhood Education for over 20 hours per week

The insights presented here provide context for the next stage of work, where regression analysis will examine the independent association between ECE participation and developmental outcomes once background factors are accounted for.

*Note missing data in the IDI means children who are attending Kōhanga Reo or some playgroups will not be recorded as attending.

Context

This analysis was requested by the Ministerial Advisory Group on the ECE funding review to contribute to the existing evidence on the positive impacts of Early Childhood Education (ECE) and to highlight areas where these benefits may not be as immediately visible. While there is strong evidence in the literature supporting the benefits of ECE, there is limited information in New Zealand about its short-term positive impacts. This analysis aims to strengthen understanding of how ECE supports children as they move into school, using New Zealand data within the local education context.

This paper represents the first stage of the analysis. The second stage of analysis will examine the relationship between developmental outcomes and ECE participation, analysing SDQ-P scores and PEDS pathways independently. Regression analysis will control for a range of demographic and socio-economic variables, providing insight into the strength of the association between early childhood education participation and developmental outcomes measured at the B4 School Check.

This paper is organised as follows. The first part outlines the IDI definition of the cohort of children used in the analysis examining ECE participation and scores on both the SDQ and the PEDS at the B4

School Check. SDQ-P is completed by parents or carers to assess the emotional and behavioural wellbeing of children, while PEDS is completed by parents or carers to identify potential developmental, behavioural and social-emotional difficulties in children.

The second part presents initial insights from descriptive analysis, comparing groups of children based on their SDQ-P and PEDS results. The analysis explores differences in demographic characteristics, ECE participation patterns, and recorded early life experiences prior to age four, including observed interactions with health, justice and social support services. Findings are intended to highlight patterns of variation between groups rather than establish causal relationships. The second stage of analysis will measure the strength of the relationships between ECE participation and B4 School Check developmental outcomes.

Methodology

B4 School Check (B4SC)

We used the IDI as our data source for this analysis. See Social Investment Agency¹ (2017) for more information on the IDI. In New Zealand, all four-year-old children are eligible for a B4 School Check (B4SC), which is funded by Te Whatu Ora and mainly delivered through Well Child Tamariki Ora services. With an annual participation rate of around 80 to 90 percent, it provides one of the best indications of how most New Zealand children are progressing before they begin school.

The B4SC assesses children's physical, dental, developmental, and behavioural wellbeing, with the developmental and behavioural aspects being the focus for this analysis.

The developmental and behavioural parts of the assessment are measured by the Strengths and Difficulties Questionnaire (completed by parents) (SDQ), the Strengths and Difficulties Questionnaire (completed by teachers) (SDQ-T) and the Parental Evaluation of Developmental Status (PEDS). Because not every child receives an SDQ-T, the parent completed SDQ and PEDS are the two measures used in this analysis. This information is collected by the Ministry of Health and loaded into IDI.

The B4SC practitioner's handbook² includes the full set of questions and lays out the scoring guidelines for the SDQ and PEDS, splitting the scores into the following categories:

<i>SDQ score</i>	<i>Category</i>
0-13	<i>Normal</i>
14-16	<i>Borderline</i>
17 or more	<i>Concerning</i>

<i>PEDS pathway</i>	<i>Category</i>
D	<i>No concerns</i>
C	<i>One non-significant concern</i>
B	<i>One significant concern</i>
A	<i>Two or more significant concerns</i>

PEDS results also note if there was a parental communication barrier to completing the assessment. Where included in the descriptive analysis, 'Declined' refers to those children who completed one of the measures, but not the other.

¹ [SIA's beginners' Guide to the IDI](#)

² [The B4 School Check - A handbook for Practitioners](#)

Early Childhood Education (ECE) participation

This analysis utilises ECE information collected by the Ministry of Education in the IDI. We examined whether children attended ECE, and their average weekly hours during the years in which they were aged two and three. Attendance status was also recorded for children who did not attend ECE, had unknown attendance or were not resident in New Zealand in the relevant year. In addition, the type of ECE attended for the greatest duration in each child's second and third year was recorded.

Population definition

The study population definition is as follows:

Children born between 2013 and 2019 and resident in New Zealand in their 2nd, 3rd or 4th year who:

- Had a B4 School Check
- Completed either the SDQ-P, PEDS or both
- Are on the IDI spine³

Using this definition, we identified **337,503** children with a recorded SDQ-P score and **339,894** children with a recorded PEDS pathway.

Descriptive statistics and demographics are *as at* the child's 4th birthday.

The SQL code used to define this cohort is available on request and will be uploaded to the Social Investment Agency's [GitHub](#) in due course.

Validation

Our results looking at ECE attendance have not been validated against the Ministry of Education's Participation Intensity Measure (PIM)⁴ for exactness, however the figures produced in the IDI were like the PIM figures for both 2- and 3-year-olds.

Additionally, our analysis found that about 80-90% of children born 2013 to 2019 have had either a full or partial B4 School Check, noting we have not taken PHO enrolment into account (published figures are 90%⁵ among children enrolled with a PHO).

Exclusions

Children who are currently resident in New Zealand but not resident when they were 2, 3 or 4 have been excluded from this analysis. Additionally, children who did not have a B4 School Check, or had a partial check completing neither the SDQ-P nor PEDS, have been excluded.

³ [VHIN NZ \(2023\) IDI Spine](#)

⁴ [Early learning participation | Education Counts](#)

⁵ [How universal are universal preschool health checks? An observational study using routine data from New Zealand's B4 School Check | BMJ Open](#)

Descriptive analysis

We undertook descriptive analysis of the population outlined above, grouped by their SDQ or PEDS scores. Children were grouped according to their screening results to enable comparison across levels of identified concerns.

The analysis examines demographic characteristics, ECE participation, and recorded experiences in the child's life up to their fourth birthday.

The purpose of this analysis is to identify differences between groups across multiple domains, including demographic profile, service contact patterns, and recorded early life experiences. Findings are intended to highlight patterns of variation rather than establish causal relationships.

Definitions

Demographics of children are shown by sex, area deprivation and total response ethnicity. Below is a short description of how each demographic is defined.

Sex: is as recorded in the IDI. In children, this is usually biological sex as assigned at birth.

Area deprivation: is based on the NZ Deprivation index 2023 quintiles, with *very low* being quintiles 1 and 2 (least deprived areas), and *very high* being quintiles 9 and 10 (most deprived areas).

Total response ethnicity: counts every ethnic group a child identifies with, meaning individuals can appear in more than one category.

IDI Disclaimer

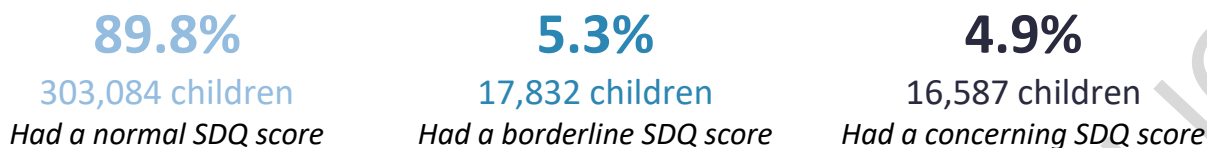
These results are not official statistics. They have been created for research purposes from the Integrated Data Infrastructure (IDI) which is carefully managed by Stats NZ. For more information about the IDI please visit <https://www.stats.govt.nz/integrated-data/>. Access to the data used in this study was provided by Stats NZ under conditions designed to give effect to the security and confidentiality provisions of the Data and Statistics Act 2022.

The results are based in part on tax data supplied by Inland Revenue to Stats NZ under the Tax Administration Act 1994 for statistical purposes. Any discussion of data limitations or weaknesses is in the context of using the IDI for statistical purposes, and is not related to the data's ability to support Inland Revenue's core operational requirements.

Strengths and Difficulties Questionnaire (SDQ)

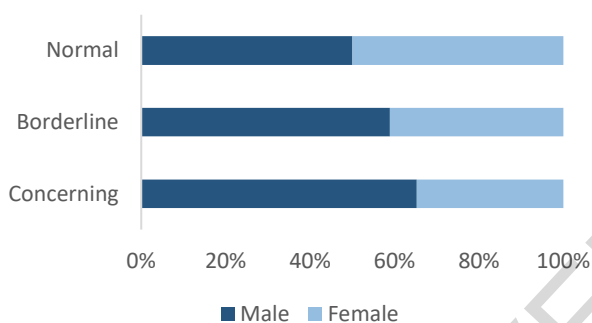
Demographics

Across the study population of **337,503** 4-year-olds who had a completed SDQ:



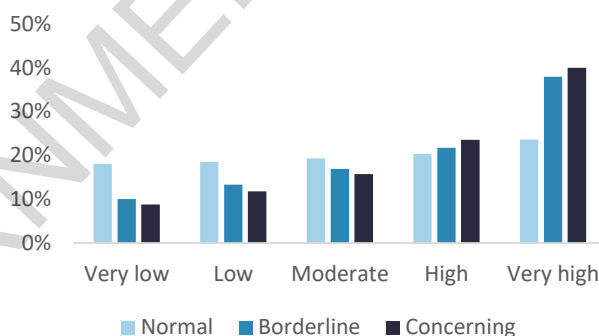
Children with higher SDQ concern scores are concentrated in specific demographic and socioeconomic groups: they include a higher share of males, children living in high-deprivation areas, and over-representation of Māori and Pasifika children in the Borderline and Concerning groups.

Proportion of SDQ score group by sex



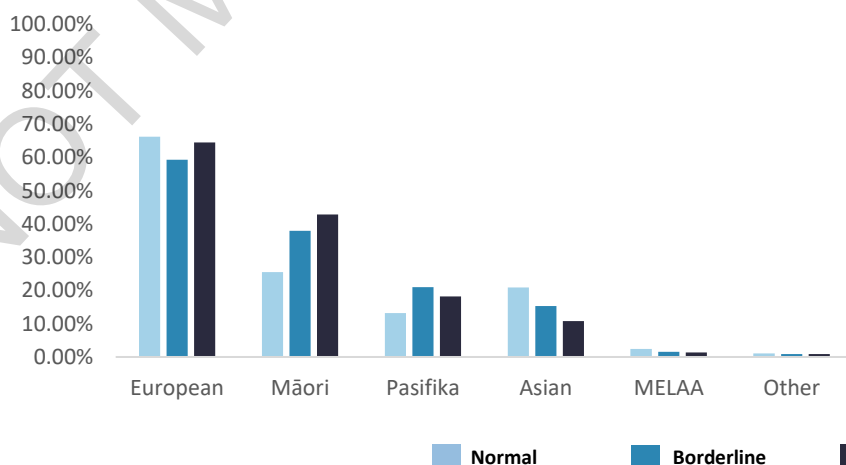
Males make up a notably larger proportion of children with Concerning SDQ scores than those in the Normal range (65% vs 50%), showing a clear step-up in the male proportion as concerns increase.

Distribution of SDQ scores by area deprivation



Children with Concerning SDQ scores are heavily concentrated in more deprived areas. Nearly two-thirds live in high or very high deprivation areas.

Distribution of SDQ scores by total response ethnicity



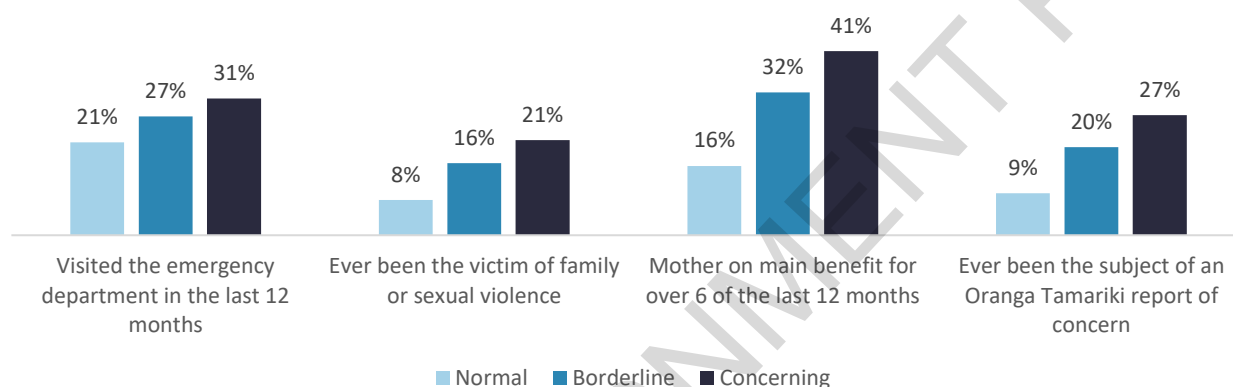
Māori, and to a lesser extent Pasifika children, account for a larger share of Borderline and Concerning SDQ scores compared with Normal.

Note: **Total response ethnicity** means individuals can identify with multiple groups; category totals can exceed 100%.

Descriptive statistics – Early life experiences

Across all four indicators shown, children with Concerning SDQ scores have experienced a substantially higher level of adversity and system contact before age 4. For example, they are more likely to have recently visited the Emergency Department, been a victim of family or sexual violence, had a mother on a main benefit for a sustained period and been the subject of an Oranga Tamariki report of concern. Each indicator shows a stepwise gradient from Normal to Borderline to Concerning, indicating that elevated SDQ scores may be closely intertwined with accumulated early-life disadvantage.

Experiences prior to turning 4

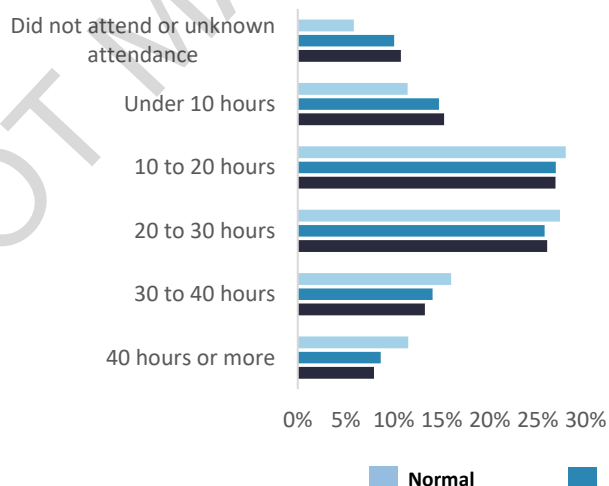


Descriptive statistics – ECE participation

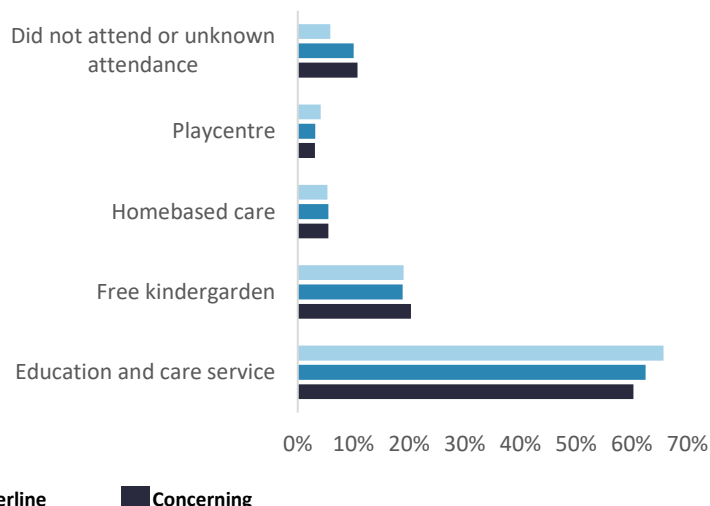
ECE participation patterns, both hours attended and primary ECE type at ages 3, show some differences across SDQ groups. Children with Concerning scores are more likely not to attend ECE or have unknown participation* and less likely to be in ECE for longer hours. Overall distributions are relatively similar. Unlike other early-life factors, ECE participation does not show as strong variations by SDQ category. Data for ECE participation at age 2 was reviewed and patterns at age 3 are consistent, with slightly lower overall hours across the three groups at age 2.

*Note missing data in the IDI means this group likely contains children who are attending Kōhanga Reo or some playgroups.

Average weekly ECE participation at age 3



Primary ECE type at age 3

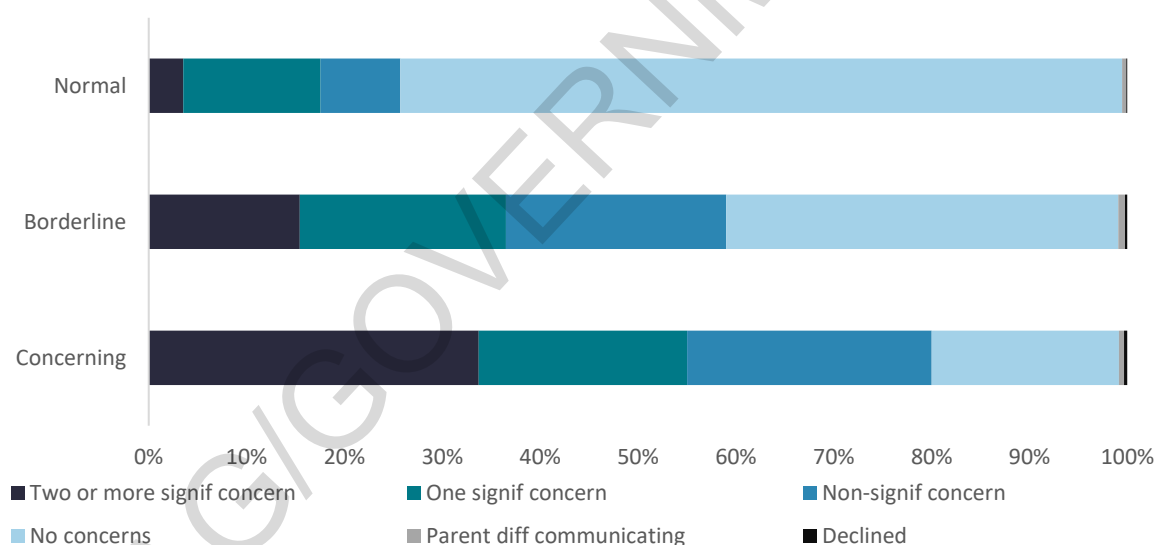


Children's scores across the two measures

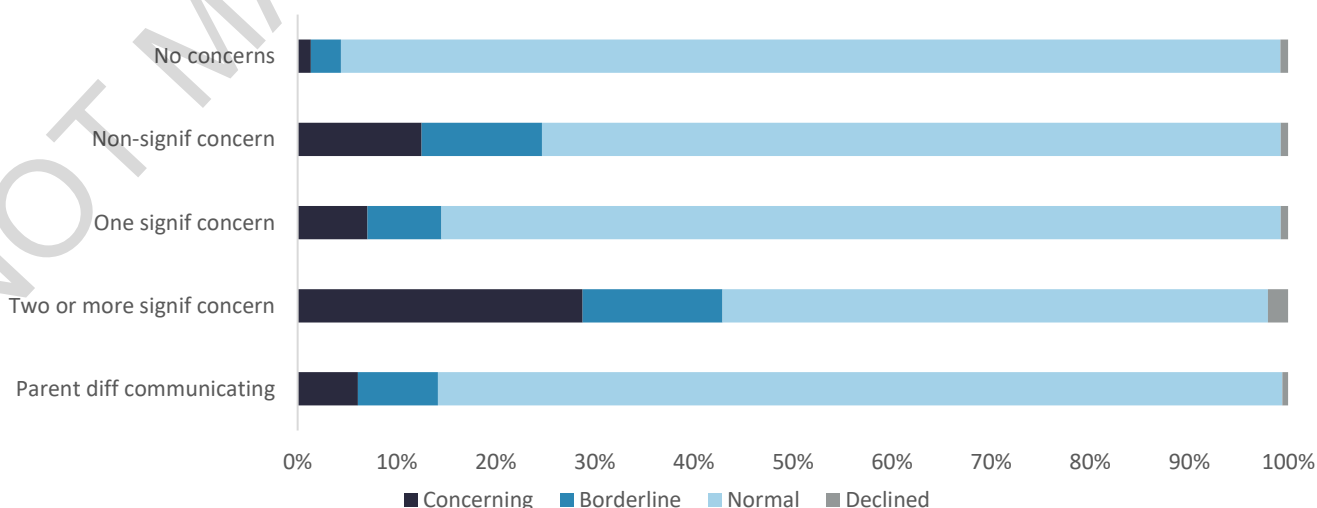
The relationship between PEDS pathways and SDQ scores shows a consistent pattern in that children with Concerning SDQ scores are far more likely to have one or more PEDS concerns (including concerns considered non-significant), while those with Normal SDQ scores predominantly fall into the PEDS “no concerns” group. We see here that a third (33.7%) of children with Concerning SDQ scores have two or more significant PEDS concerns, compared with only 3.5% of those with Normal SDQ scores, and fewer than 2% of children with Concerning scores have no PEDS concerns. However, it is notable that 18% of children with a Normal SDQ score still had at least one significant PEDS concern, indicating that parental concerns can arise even when behavioural/emotional difficulties are not strongly indicated on the SDQ. Overall, this pattern reinforces the alignment between parental concern indicators and measured behavioural/emotional difficulties, showing that children with elevated SDQ scores tend to present with multiple, compounding developmental concerns.

Note: while most children in the overall study group had both SDQ and PEDS completed at their B4SC, a small number in the SDQ group declined the PEDS, and a small number in the PEDS group declined the SDQ.

PEDS pathway score by SDQ score



SDQ score by PEDS pathway



Parental Evaluation of Development Status (PEDS)

Demographics

Across the study population of **339,894** 4-year-olds who had a completed PEDS:

69.3%

235,698 children

Had no concerns raised

9.7%

33,078 children

Had non-significant concerns raised

14.8%

50,172 children

Had one significant concern raised

5.7%

19,461 children

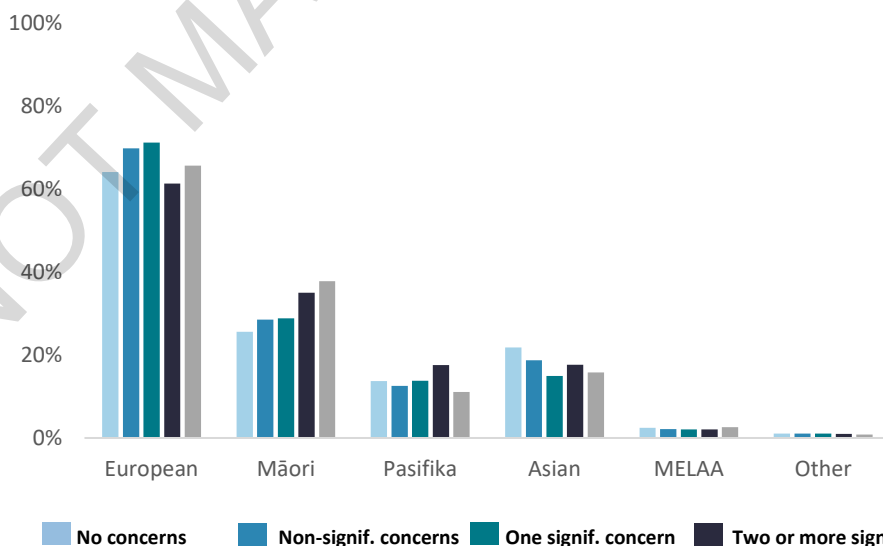
Had two or more significant concerns raised

0.5%

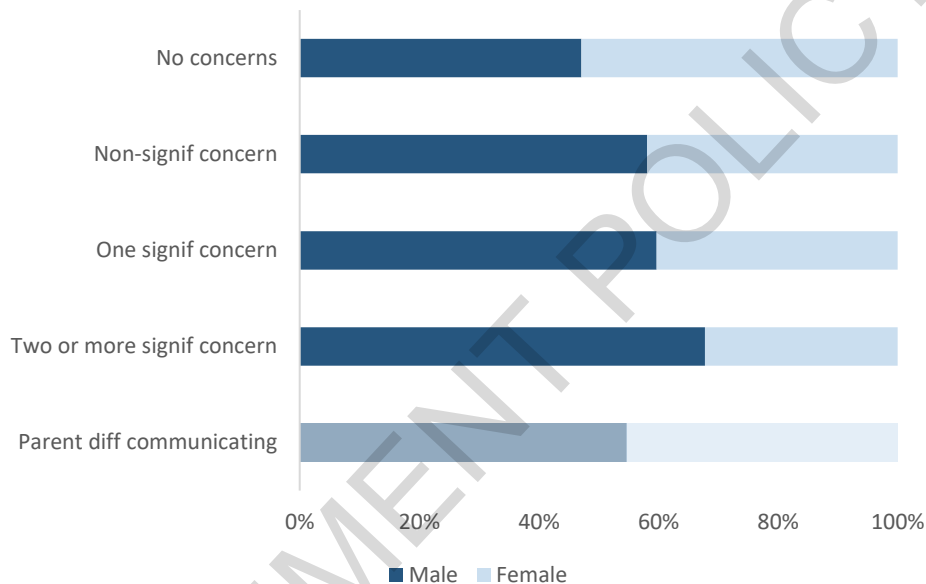
1,485 children

Had a parent with difficulty communicating concerns

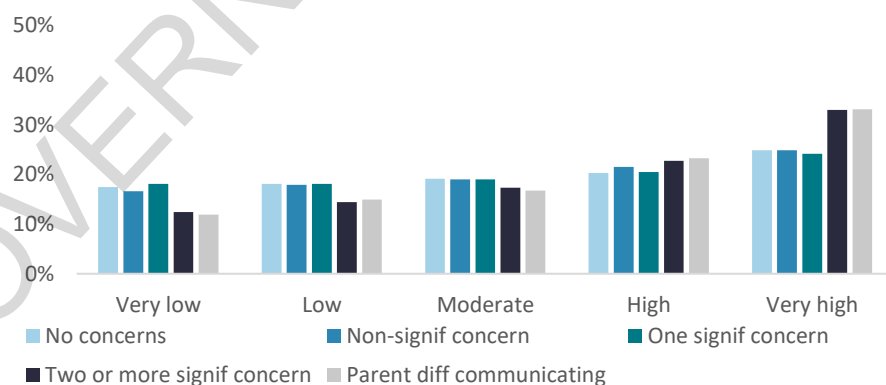
Distribution of PEDS pathways by total response ethnicity



Proportion of PEDS pathway group by sex



Distribution of PEDS pathways by area deprivation

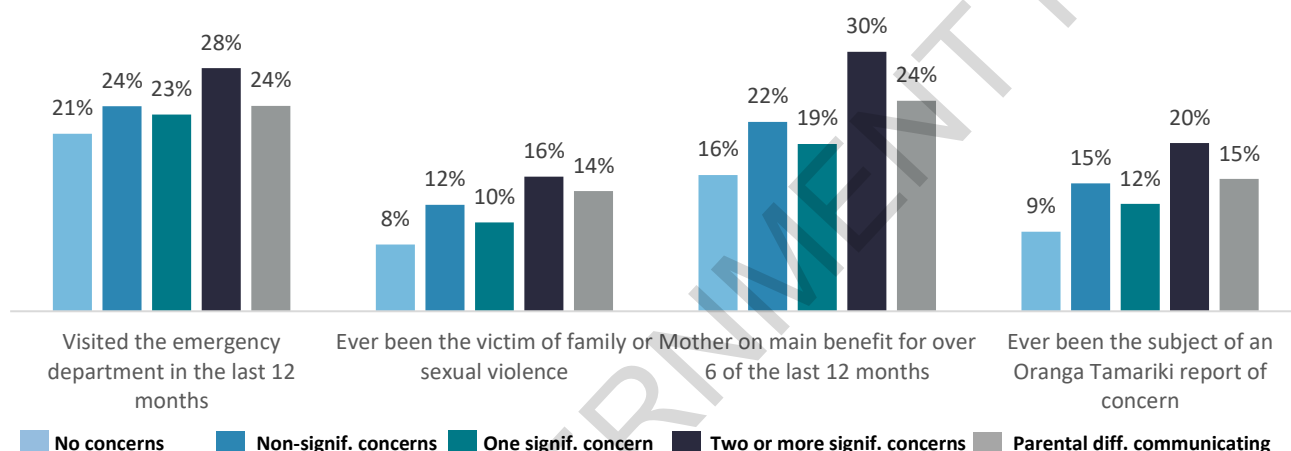


The demographics across the PEDS groups show similar patterns to the SDQ data in that children with one or more concerns are more concentrated in high-deprivation areas, more often Māori or Pasifika, and have a higher male share in the more significant concern categories. Together, the two B4SC measures point to a consistent set of demographic and socioeconomic groups in which developmental and behavioural concerns are more common.

Descriptive statistics – Early life experiences

The PEDS early-life experience data show clear gradients that mirror those found in the SDQ results. Children with one or more PEDS concerns are more likely to have adverse experiences and live in hardship, patterns that align closely with the elevated early-adversity rates seen among children with Concerning SDQ scores. Interestingly however, the non-significant concern group often shows higher levels of adverse experience than the one-significant concern group, suggesting that even lower-level parental concerns can be associated with early-life disadvantage. Taken together, the PEDS and SDQ data provide a consistent picture developmental and behavioural concerns tend to cluster among children with greater exposure to early adversity.

Experiences prior to turning 4

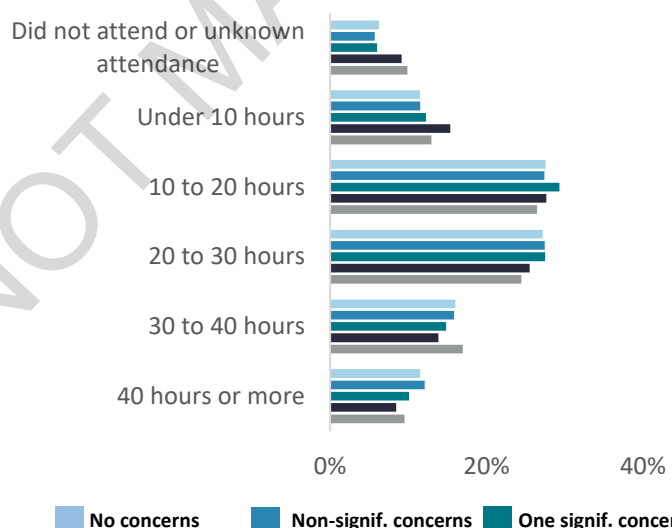


Descriptive statistics – ECE participation

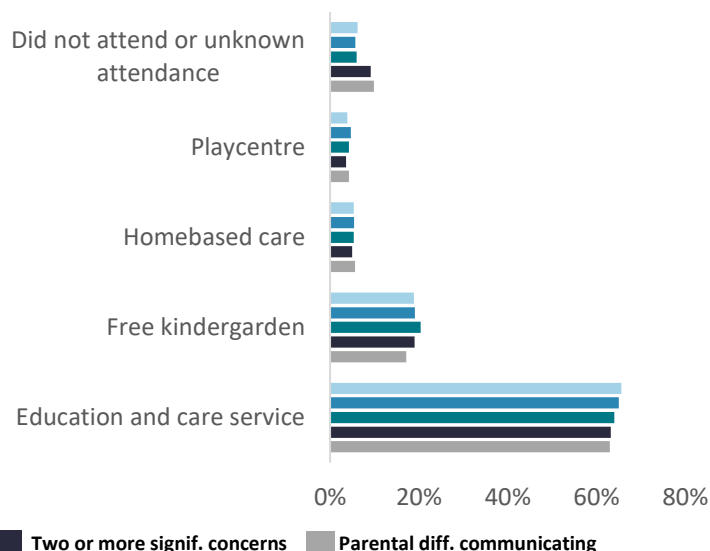
As with the SDQ groups, ECE participation and type at age 3 show only some differences across PEDS groups, with children who have multiple concerns more likely not to attend or have unknown attendance. This mirrors the SDQ findings, where ECE exposure showed some differences across concern levels

*Note missing data in the IDI means this group likely contains children who are attending Kōhanga reo or some playgroups.

Average weekly ECE participation at age 3



Primary ECE type at age 3



Summary

Across the descriptive analysis, a consistent set of demographic and early-life patterns emerges in relation to children's developmental (PEDS) and behavioural/emotional (SDQ) outcomes. Higher concern levels on both measures are more common among children living in high-deprivation areas, those identified as Māori or Pasifika, and those with greater exposure to early adversity and hardship. Sex differences are also evident, with males making up a larger share of children with Concerning SDQ scores and higher-level PEDS concerns.

The alignment between SDQ and PEDS is consistent. Children with Concerning SDQ scores are far more likely to have one or more PEDS concerns, including multiple significant concerns, while children with Normal SDQ scores predominantly fall into the "no concerns" PEDS pathway. However, some children with Normal SDQ scores (18%) still had at least one significant PEDS concern, indicating that parental concerns can capture issues not fully reflected in behavioural/emotional screening alone.

ECE participation, both hours and type of service attended show some variations across concern groups. Children with multiple PEDS concerns or Concerning SDQ scores are more likely not to attend ECE (but possibly attend Kohanga Reo or playgroups) or to attend fewer hours, but these differences are small compared with the socioeconomic and early-life patterns noted above.

Taken together, the descriptive findings suggest that developmental and behavioural concerns tend to cluster within groups of children facing higher levels of socioeconomic disadvantage and early adversity, while ECE participation patterns have fewer but some differences across groups. These patterns provide a clear foundation for the next stage of analysis, where the independent association between ECE participation and development, emotional and behavioural outcomes will be tested once demographic and early-life factors are accounted for.

Next steps - regression methodology

The next part of this analysis will examine the strength of the relationship between ECE participation and SDQ and PEDS scores independently. Given the demographic and early-life patterns observed earlier where higher concern levels clustered among boys, children in high-deprivation areas, Māori and Pasifika children, and those with greater exposure to early adversity it is important to isolate the effect of ECE hours from these confounding influences.

To do this, we will identify non-overlapping covariates by reviewing relevant confounders and testing combinations within linear regression models. The covariates used will include the indicators presented here, additional demographic and socioeconomic variables, and selected parental indicators to account for whānau-level influences.

Based on the stability and significance of these coefficients, we will then select an appropriate set of covariates to include alongside ECE hours and build a regression model.

This approach will allow us to assess whether ECE participation is independently associated with developmental and behavioural outcomes once demographic and early-life differences are considered.

Limitations

As with all IDI analysis, there are limitations to note. These are listed below.

Kōhanga Reo data

ECE attendance data in the IDI is sourced from the Early Learning Information (ELI) database. Although ELI is used by most ECE services, it does not yet have full coverage. In particular, it is not used by Kōhanga reo or some playgroups. This means that some children, particularly tamariki Māori, will be underrepresented in ECE participation statistics.

Single ECE Type Selection

For children attending multiple ECE providers, the provider with the longest duration of attendance over the period has been used to represent their primary ECE type. In cases where multiple providers have the same duration, a single provider was selected at random.

Binning of Attendance Hours

ECE attendance has been grouped into 10-hour wide 'bins' up to 40 hours. Children with 0 or missing attendance hours but marked as 'present' have been assigned to an "unknown hours" category.

Potential Marginal Inaccuracy (Estimate vs. Exact Hours)

Calculated hours may not perfectly reflect actual attendance due to periods when ECEs are closed. Desktop research suggests that many centres align term and holiday periods with primary schools, resulting in approximately 380–390 half days per year across four terms (38.5 weeks used in this analysis). Precise calculations would also need to account for teacher-only days and regional holiday variations (e.g., Auckland Anniversary Day in Q1, Hawke's Bay Anniversary Day in Q4), which may slightly alter the distribution of available hours. Consequently, the reported attendance hours should be considered a close estimate rather than an exact measure of time spent in ECE.

Data Sources and Limitations

Attendance data in the IDI and the Participation Intensity Measure (PIM) are sourced from the Early Learning Information (ELI) databases, which primarily record enrolment and participation. Reconciling these sources is complex and not recommended due to the construction of the PIM. The ELI system is dynamic, and extracts taken at different times may differ.

Comparison with ECE Census Data

IDI-based attendance data will not match Ministry of Education ECE census counts, as the census can double-count children attending multiple ECEs. Additionally, the IDI does not capture all ECE types, including playgroups and Kōhanga reo services (as above).

Impact of COVID-19 Lockdowns

COVID-19 lockdown periods were examined to assess any effect on SDQ or PEDS scores. Minimal differences were observed for this cohort; therefore, lockdown periods were excluded from the analysis.



Social Investment Agency
Toi Hau Tāngata

NOT MAG/GOVERNMENT POLICY

