



Australian Government
The Treasury

Australian
Centre for
Evaluation

ace

Pre-analysis plan: Impact evaluation of the Home Interaction Program for Parents and Youngsters (HIPPY)

Contents

Summary table	3
Policy context	4
Study aim	4
Treatments	4
Outcome measure(s)	6
Population of interest and sample collection	8
Hypotheses	8
Assumptions	9
Conditional independence assumption (CIA)	9
Common support assumption	11
Method of analysis	11
Primary/secondary outcomes analysis	11
Additional analysis	12
Study threats	13
Internal validity threats	13
External validity threats	14
Interpretation of results	14
Pre-analysis plan commitments	14
Appendix A: Service types recorded for HIPPY in DEX	15
Appendix B: Variables to control the differences between treated and untreated	16
Appendix C: Deviations from the pre-analysis plan since the review session	22

Summary table

Project title	Impact evaluation of the Home Interaction Program for Parents and Youngsters (HIPPY)
Partners	Department of Social Services (DSS)
Evaluator (Institution)	Australian Centre for Evaluation (ACE)
Principal investigator(s)	Vera Newman Ramethaa Pirathiban
Study design	Observational study
Population	Children who commenced school in 2021 and were eligible to access HIPPY in the preceding years
Sample size	Treatment group: children participated in HIPPY Comparison group: similar children who did not participate in HIPPY, identified using linked administrative data.
Primary outcome measure	Physical health and wellbeing Social competence Emotional maturity Language and cognitive skills Communication and general knowledge
Secondary outcome measures	Number of on-track domains Adapting to school Parent actively engaged with the school

This pre-analysis plan was reviewed during a group session on 29 July 2025. Any deviations from the plan in the subsequent analysis are documented in Appendix C, together with their justifications.

Policy context

The Home Interaction Program for Parents and Youngsters (HIPPY) is a 2-year home-based parenting and early childhood program that helps parents and carers to be their child's first teacher. When HIPPY commenced in Australia, the program was offered to children aged 4–5 years. Since 2023, it has been replaced by the *HIPPY Age 3* program, which is offered to children aged 3–4 years. HIPPY engages children aged between 3 and 5 years from locations with high levels of demonstrated vulnerability. The program is funded by the Department of Social Services (DSS) and delivered by the Brotherhood of St Laurence (BSL) with 66 subcontracted providers operating across 100 sites across Australia. Families enrolled in the program aim to spend approximately 15 minutes per weekday over the two-year program doing educational activities together. The program intends to create a positive home learning environment and prepare the child for school. The program also provides the child with structured education-focused early learning, with a view to improving the child's school preparedness, and building the confidence and skills of parents and carers to create a positive home learning environment.

Although HIPPY has been subject to extensive evaluation internationally and there is good evidence for HIPPY's effectiveness in other countries, there remains a question as to whether it is similarly effective in Australia. Only one impact evaluation of HIPPY has been undertaken in Australia.¹ A later journal article reported specifically on the effectiveness (impact) evaluation, and hence the effects of HIPPY on parents and children.² However, the conclusions of these evaluation are limited by the research technique, which has not adequately accounted for the bias in the sample; further, there have been substantial program changes since 2011. As a result, questions regarding the impact of HIPPY on relevant outcomes remain.

Study aim

The aim of this study is to evaluate the impact of HIPPY on several school readiness outcomes for children. Specifically, we are interested in testing whether participation in HIPPY improves these outcomes.

Treatments

The treatment comprises engagement in the HIPPY program. It is a two-year program that runs in step with local school schedules (30 weeks a year), incorporating breaks during end-of-term holidays. Parents/carers enrolled in the program spend approximately 15 minutes per weekday with their children to do play-based educational activities together. HIPPY operates under concept of 'everywhere learning' – looking for learning opportunities in everyday settings. Activity packs, storybooks and other learning resources are delivered to the parents/carers to support the activities.

-
- 1 Liddell, M. (2011). Investing in our future: an evaluation of the national rollout of the Home Interaction Program for Parents and Youngsters (HIPPY): final report to the Department of Education, Employment and Workplace Relations.
 - 2 Barnett, T., Roost, F. D., & McEachran, J. (2012). Evaluating the effectiveness of the Home Interaction Program for Parents and Youngsters (HIPPY). *Family Matters*, (91), 27-37.

Parents and carers who are doing HIPPY with their child in the program years or have completed the program in the recent past are employed as tutors. They are paid and trained to demonstrate the activities and support parents in working through the curriculum with their children. Tutors provide peer support to families in their home or another familiar environment where families feel comfortable. Home visits allow parents to discuss their HIPPY journey one-on-one with their tutor.

Role play is the primary method of delivery to families. HIPPY tutors and parents role play each activity before HIPPY parents and carers work through them with their child.

HIPPY gatherings are also held at a community venue. It is an opportunity for parents and carers to get together, share their experiences of HIPPY, build relationships with each other, enjoy enrichment activities, and do HIPPY in small groups. HIPPY is a free and voluntary program, meaning that families are not obliged to complete any or all of the HIPPY modules throughout their enrolment in the program.

The cohort for this study is children who commenced school in 2021 and are part of the 2021 Australian Early Development Census (AEDC).³ The HIPPY program data is recorded in the Data Exchange (DEX), a platform used by DSS for grants program reporting. The DEX Program-Specific Guidance for 2017 and 2020 will be used to identify the service types that are relevant to children and would therefore be an indicator of program participation. Appendix A shows the service types recorded for HIPPY based on the DEX Program-Specific Guidance for 2017 and 2020 and their descriptions.

For the purpose of this evaluation, we will define program participation as all children who had participated in at least one session of child specific activity, that is, child having enrolled ('intake/ assessment') and attended at least one session of 'family capacity building', 'education and skills training', or 'community capacity building' prior to the completion of their AEDC.⁴ In addition, if a child received a referral to other non-HIPPY services but did not participate in any other HIPPY sessions/activities, this will not equate to participation in HIPPY. This is a more policy-relevant definition for HIPPY participation although it will result in a small decrease in sample size compared to including all children who have enrolled, regardless of whether they commenced or participated in any activities. All children who had participated in HIPPY using these criteria will therefore be coded as treated children. The remainder of the children in the AEDC census will be coded as untreated. This will provide us the intent-to-treat (ITT) effect of the program. This may dilute the real treatment effect, but this is unavoidable given the structure of the data. Since we are considering ITT, dosage of the program will not be addressed in this evaluation.

3 There are few limitations around the choice of AEDC 2021 for this analysis. The 2021 AEDC data captured HIPPY participants before the curriculum changes from targeting 4-5-year-old children to 3-4-year-olds. Additionally, this HIPPY cohort was recruited during the pandemic, which might have affected HIPPY's delivery and outcomes. The 2021 AEDC data is, however, the most currently available because the 2024 AEDC captured very few HIPPY families, due to the change in the program from 4-5 years to 3-4-year-olds. AEDC 2018 was not considered due to 2 reasons. First, there were no sufficiently reliable indicators of HIPPY program participation for this period in the DEX data. Second, the required control variables for this cohort would have needed to be sourced from the 2016 Census of population and housing as it will contain the most current information of parents and families; however, legislative arrangements only permit access to one year of census data.

4 AEDC is typically conducted between May to August every 3 years. The month in which AEDC is completed for each child is available in the AEDC dataset.

Outcome measure(s)

This evaluation will compare HIPPY children (treated) with similar children in the same communities (comparison group), identified through linked administrative data. Data from the Person-level Integrated Data Asset (PLIDA) available through the Australian Bureau of statistics (ABS) will be the primary source for controlling for differences between HIPPY and non-HIPPY children, as well as measuring outcomes for both groups. All analysis will be conducted within the ABS secure DataLab environment where data will be made available in de-identified format. All outputs will be vetted by the ABS before they can be removed from the secure DataLab environment.

The AEDC dataset available within PLIDA will be the primary source for measuring outcomes of children for both the treated and the comparison groups. AEDC is a nationwide data collection of early childhood development at the time children commence their first year of full-time school. It measures five areas or domains of early childhood development: Physical health and wellbeing, Social competence, Emotional maturity, Language and cognitive skills and Communication skills and general knowledge, which we define as the primary outcomes (**Table 1**). The scores for these domains are calculated based on teacher responses to the relevant domain questions for each child. Secondary outcomes aligned with HIPPY's program logic are also considered. These outcomes will also be sourced from AEDC (**Table 1**).

Adapting to school was chosen as a secondary outcome as it is one of the outcomes that HIPPY is trying to achieve in terms of school readiness. It is important to note that HIPPY is a parenting program and therefore parent related outcomes are also of interest. Thus, we have chosen 'parents actively engaged with the school' as another secondary outcome. While the primary outcomes focus on continuous scores across each AEDC domain, the number of on-track domains captures a summarised view of how well a child is doing across multiple domains and is more accessible. This is a relative measure so there is concern given HIPPY's focus on children from disadvantaged backgrounds who are more likely to experience developmental vulnerability. However, since the comparison group will be weighted to ensure comparability with the HIPPY group (see Methods), this reduces the risk of biased interpretation arising from the use of relative thresholds.

Table 1: Description of primary and secondary outcome variables

	Variable	Description
Primary	Physical health and wellbeing	Physical health and wellbeing measures child's physical readiness for the school day, physical independence and gross and fine motor skills. This ranges between 0 and 10, where 0 is most developmentally vulnerable.
Primary	Social competence	Social competence measures children's overall social competence, responsibility and respect, approaches to learning and readiness to explore new things. This ranges between 0 and 10, where 0 is most developmentally vulnerable.
Primary	Emotional maturity	Emotional maturity measures children's pro-social and helping behaviour, anxious and fearful behaviour, aggressive behaviour and hyperactivity and inattention. This ranges between 0 and 10, where 0 is most developmentally vulnerable.
Primary	Language and cognitive skills	Language and cognitive skills measures children's basic literacy, interest in literacy, numeracy and memory, advanced literacy and basic numeracy. This ranges between 0 and 10, where 0 is most developmentally vulnerable.
Primary	Communication skills and general knowledge	Communication skills and general knowledge measures children's communication skills and general knowledge. This ranges between 0 and 10, where 0 is most developmentally vulnerable.
Secondary	Adapting to school	Adapting to school identifies whether the child is making good progress in adapting to the structure and learning environment of the school. This is a categorical variable with 3 levels: not true, somewhat true and very true.
Secondary	Parents actively engaged with school	Parents actively engaged with school identifies whether the child has parent(s)/caregiver(s) who are actively engaged with the school in supporting their child's learning. This is a categorical variable with 3 levels: not true, somewhat true and very true.
Secondary	Number of on-track domains	Number of on-track domains counts the number of domains in which the child is in the 50 to 100th percentile category. This is an integer and ranges between 0 and 5. This is based on cut-offs for each domain. Domain cut-offs are created based on the data from all children who participated in the AEDC and take into account the age variations in the population of children in their first year of schooling. The cut-off percentile for each domain category is based on the baseline set in the 2009 AEDC data collection.

Population of interest and sample collection

Our population of interest is children who commenced school in 2021 and were eligible to access HIPPY in Australia.

To be eligible for the program, a family must meet both of the following eligibility criteria:

- The family has a child aged four in the year of enrolment to HIPPY (prior to program changes implemented in 2022).
- The family lives in an area where HIPPY is being delivered in the year of enrolment.

We will use the AEDC data to extract children who commenced school in 2021 and their ages. In order to obtain information whether the child has lived in an area where HIPPY is delivered, we will use the Core locations dataset. Core locations is a dataset that contains location specific information combined from multiple core data sources by the ABS. This dataset will be used to extract location information of the children to determine whether the child has resided in a Statistical Area Level 3 (SA3) where HIPPY is delivered when they are at the age of 4. Once the data relevant to the eligibility criteria is obtained, the dataset will be reduced to only those who are eligible to be included in the program.

We will then extract variables to control the differences between HIPPY and non-HIPPY children. The key datasets that will be used to extract these variables are Census of Population and Housing 2021,⁵ Data Over Multiple Individual Occurrences (DOMINO), Data Exchange (DEX), Income Tax Return (ITR), Medicare Benefits Schedule (MBS) and Australian Immunisation Register (AIR). These variables will be joined with the reduced dataset that contains the age and the location information of the children.

For parent/carer related variables, we will identify the parent/carer of the children using family id, parent indicator and family/household reference person information available in census. For children living with carers, the family/household reference person will be defined as the carer. For couple parents/carers, we will pool the information across the two parents/carers.

Hypotheses

Our hypotheses flow from the theory of change of HIPPY that HIPPY builds confidence and skills of parents and carers of children in disadvantaged communities to create a positive home learning environment that assists children prepare for school.

Our alternate hypothesis is that participation in HIPPY would lead to better school readiness outcomes for children and greater school engagement for parents, compared with those who did not participate.

This involves comparison of the treatment group against the comparison group. Given our directional hypothesis, we will use a one-sided test.

⁵ Census 2021 data was collected after the treatment period. Despite being post-treatment, the Census provides the best available and most current data for the cohort of interest to account for differences between HIPPY and non-HIPPY children.

We will not apply corrections for multiple comparisons across the primary or the secondary outcomes. This is because we intend to conduct a conjunction test, where the interpretation considers the overall pattern of results rather than significance in individual outcomes. For instance, if only one outcome is statistically significant while others are not, this would not in itself be sufficient to determine whether the program is effective. Instead, we will consider the collective pattern of results. Furthermore, we expect correlation across the outcome domains, meaning the tests are not statistically independent. The same rationale applies to the secondary outcomes as well.

Assumptions

Conditional independence assumption (CIA)

The observational study relies on the conditional independence assumption (CIA). This means that the variables on which the treated and comparison (untreated) groups differ must be observable to the researcher. These variables are called the confounders. After controlling for these variables, the treatment assignment is “as good as random”. The CIA is crucial for correctly identifying the impact of the program, since it ensures that, although treated and untreated groups differ, these differences may be accounted for in order to reduce the selection bias. Unfortunately, CIA is not directly testable but still requires justification.

In the context of HIPPY, we need to justify that selection into the program is based only on observable characteristics of the eligible children, and that after conditioning on the variables influencing participation, the expected outcome in the absence of treatment does not depend on treatment status. This allows the untreated children to be used to construct a counterfactual for the treatment group.

However, the extant literature indicates that selection into HIPPY is not random. It is based on various factors, only some of which are observable. Not adequately controlling for factors that influence program participation will contribute to bias in the results.

In addition to the eligibility criteria, there are several criteria to support priority access to disadvantaged families, including:

- the family holds a health care card
- the HIPPY child is Aboriginal and/or Torres Strait Islander
- the HIPPY child resides in out-of-home care
- the household has no income, or receives a government support payment as their primary source of income
- the family is a single-parent family
- the child resides with a carer (not a parent)
- the main language spoken at home is not English
- child has a developmental delay or disability
- parent or carer lives with a disability or long-term health condition.

In order to better understand the selection process at sites and characteristics of HIPPY participants, we conducted semi-structured interviews with the site coordinators. Findings from these interviews revealed that:

- Families who are willing to participate in HIPPY are more dedicated to their child's development;
- HIPPY providers use discretionary approach to select participants into the program. In particular, HIPPY providers prioritise families with stable livelihoods as they are more likely to commit to the two-year program participation. However, this was particularly the case where site coordinators had a history of oversubscription in their sites (and applies to a lesser extent where sites may be undersubscribed).
- Some families who haven't immunised their children use HIPPY as a substitute for childcare/preschool programs.

While some characteristics of families who participate in HIPPY are observable, site interviews revealed that many other unobservable factors also influence program participation. These unmeasured characteristics are likely to affect the outcomes that HIPPY aims to improve.

It is also important that the variables that will be used to control for the differences between the treated and the untreated groups to be informed by sound theory. Thus, we have conducted a literature review on the motivations and factors influencing parents' decisions to enrol their children in early childhood, home intervention programs such as HIPPY. We have also consulted with the Policy team and the Evaluation hub within DSS for feedback as their knowledge of the program is also essential to control for any differences between the two groups.

To justify specification of the model used to estimate propensity scores (see below in Method of analysis), we will:

- 1) Use eligibility and priority criteria explicitly known to influence selection and draw other potential variables from the below list:
 - variables considered in Liddell et al 2011 to establish a comparison cohort in the Propensity Score Matching approach;
 - variables extracted from a literature review; and
 - variables derived based on the consultations with the HIPPY site coordinators.

Appendix B provides a list of these variables, their data sources and how they have been derived.

- 2) Ensure that the variables included in the model are stable (constant) over time, deterministic with respect to time (for example, age) or measured before participation so that they are not confounded with outcomes or the anticipation of treatment.
- 3) Draw data for the treatment and comparison groups from the same data sources and will construct the variables to control the differences similarly for both groups.
- 4) Avoid using irrelevant variables (that do not influence participation in the program) so that they do not worsen the common support problem (see below) or unnecessarily increase the variance of the estimates.

Common support assumption

Another important assumption for the observational study is the common support or overlap condition. This assumes that the probability of participating in the program lies strictly between 0 and 1—meaning that participation is not perfectly predicted by observed characteristics. In practical terms, this requires that for every combination of values of the control variables (this is, confounders), there are both treated and untreated individuals in the data. This ensures that valid comparisons can be made across similar individuals in both groups.

We will use the estimated propensity scores to achieve this condition. Instead of needing treatment and comparison group units on each combination of values of the confounders, we just need to have treatment and comparison group units across the propensity score range.

It is useful to plot the distributions of the estimated propensity scores for treated and untreated groups to visually check the overlap condition and to see if the distributions are more similar after we have controlled for the differences between treated and untreated groups.

Method of analysis

Primary/secondary outcomes analysis

We will focus on the average treatment effect on the treated (ATT) statistic as the measure of the program's effectiveness for both primary and secondary outcomes. HIPPY was not meant to be available to all children but to a selected group of children with eligibility and priority access to the program. As such, the ATT statistic provides an understanding of whether the program was successful with the targeted group of children. As such, we will conduct a series of analyses to estimate the ATT for both primary and secondary outcomes of interest. We will use inverse probability weighting (IPW) combined with a doubly robust (DR) estimator, which controls for observed differences between treatment and comparison groups using both a propensity score model and an outcome regression model. This approach provides consistent ATT estimates if either the treatment assignment model or the outcome model is correctly specified. The approach is described below.

Firstly, we will estimate propensity scores by fitting a propensity score model using the treatment status as the response variable and the variables to control for the difference between the treated and untreated groups as covariates. We will use logistic regression noting that we will not optimise for accurate prediction of the treatment status but for balancing confounders across treated and untreated groups (which is what the 'true' propensity score is supposed to do).

Secondly, we will check common support by looking at the density plots of the estimated propensity scores for the treated and untreated groups. The observed propensity scores should be strictly between 0 and 1. As we are interested in ATT, we only need the propensity score distribution of the treated to be a subset of the propensity score distribution of the untreated. Note that we do not need both distributions to look the same, but we do need a positive density or weight of the untreated distribution across all the treated group distribution.

In the case of extreme propensity scores, for example less than 0.05 and more than 0.95 trimming of such observations is commonly used to avoid the instabilities and to reduce the variance. However, trimming usually changes the parameter of interest. We will consider methods such as stabilizing the weights or use optimal trimming rules, or a bias corrected trimmed method such as Ma et al 2023.⁶

We will then use IPW approach to estimate the causal effects of the ATT. In order to obtain the DR estimate for ATT, we will fit an outcome regression model for the untreated group only, with the outcome as the response variable and all the variables used to control for the difference as covariates. We will obtain the predicted values from this model for the whole data (both treated and untreated). We will then use the estimated propensity scores and predicted values to weight the observations in order to make the comparison group similar to the treated group and obtain the DR estimate following Mercatanti and Li, 2014.⁷

We will obtain standard errors for the DR estimate using bootstrapping. Bootstrapping has the advantage of taking into account the uncertainty of the whole process.

Additional analysis

We will conduct robustness checks by using a propensity score matching method, for example 1:1 nearest neighbour distance matching, and compare the change in the effect sizes and standard errors. This is to investigate the precision of our findings and to investigate whether there is sensitivity to the method used which can likely create bias in our results. Similar to IPW, we will estimate propensity scores from a logistic regression method. Then we will match each treated group unit to a comparison group unit that has the closest propensity score to it. Any remaining comparison group units are left unmatched are excluded from further analysis. We will then fit a regression of the outcome on the treatment and covariates in the matched sample and estimate the ATT. Including the covariates used in the matching in the effect estimation can provide additional robustness to slight imbalances remaining after matching and can improve precision.

6 Ma, Y., Sant'Anna, P. H., Sasaki, Y., & Ura, T. (2023). Doubly robust estimators with weak overlap. *arXiv preprint arXiv:2304.08974*.

7 Mercatanti, A, and Li F. (2014). Do debit cards increase household spending? Evidence from a semiparametric causal analysis of a survey. *Annals of Applied Statistics*. 8(4), 2405-2508.

We will also consider a placebo test to probe the conditional independence identification and examine whether alternative explanations and identification issues are likely to cause bias in the results. This assesses whether an observed treatment effect is truly causal or if it can be attributed to placebo effects or other factors related to the comparison group. We will randomly assign the ‘treatment’ label to comparison group units (keeping treatment size same) and re-estimate the DR ATT estimate. This is called the placebo estimate. We will repeat this multiple times to look at the distribution of the placebo estimates which will show whether our original effect is far from what random chance would produce. We will also examine the permutation p-value which is the share of placebo estimates that are as extreme as the original estimate. A low p-value suggests that very few random assignments produce an effect as large as ours so that the observed effect is unlikely to be due to chance or spurious assignment. This supports (but does not prove) that our identification strategy is strong.

Study threats

Internal validity threats

Omitted variable bias

As discussed before, we acknowledge that due to the unobserved characteristics of the families who participate in HIPPY it will not be possible to completely eliminate selection bias. Therefore, it is probable that the CIA may not fully hold. We will investigate omitted variable bias by using methods such as Altonji, Elder, and Taber (2005)⁸ and Oster (2015)⁹ to look at how strong should the un-observables be compared to the observables in explaining the outcomes in order to make the ATT zero.

Blinding

The outcomes considered in this analysis are teacher-reported and the teachers do not have information on whether the child had participated in HIPPY or not. This information will only become available to the teachers if the parents discuss this with the teachers. This means that the evaluation effects (such as the Hawthorne effect) are a possibility. However, it is our view that evaluation effects are unlikely to be strong enough to impact school readiness outcomes.

Spillovers/contamination

The chance of meaningful treatment contamination or spillovers are low. This is because families with older children who’ve participated in HIPPY are not prioritised for future rounds. Thus, we deem this kind of interference/contamination as posing little risk to the study.

8 Altonji, J. G., Elder, T. E., & Taber, C. R. (2005). Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools. *Journal of political economy*, 113(1), 151-184.

9 Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, 37(2), 187-204.

External validity threats

We suspect that it will not be possible to link the program participation data in DEX for the most vulnerable cohort with the other datasets in PLIDA. This is due to the data entry practices followed in DEX to protect participants' privacy and the linkage method used. We will perform exploratory analysis using the DEX data for the individuals who are not linked to other datasets and report on their demographics.

Interpretation of results

The likely risk of selection bias described above will limit the strength of the conclusions that can be drawn from this analysis, which should be more accurately described as an analysis of *associations* (rather than causal relationships) between HIPPY and the various school readiness outcomes. Associations derived from an observational approach may still provide some useful insights but would likely require supplementary evidence. To support this, additional evaluation components are planned, including surveys and interviews with HIPPY parents, as well as analysis of program data collected by BSL at the start, midpoint, and end of the HIPPY journey. These elements are not described in this pre-analysis plan, as they fall outside its scope.

Pre-analysis plan commitments

We have not undertaken any analysis prior to the completion of this pre-analysis plan. We will be transparent about, and provide justification for, any deviations (additions or omissions) from this pre-analysis plan.

Appendix A: Service types recorded for HIPPY in DEX

Service Type	Example of service type use within the program activity 2017	Example of service type use within the program activity 2020
Intake and assessment	Family enrolment and/or tutor employment.	Family baseline enrolment journey when the site coordinator meets with a HIPPY parent/carer/child and completes an enrolment form and baseline journey (assessment) for measurement of the child/family and community outcomes.
Information / advice / referral	Group meetings, family contact, individual information provision and/or service connection.	HIPPY parent/carer/child receives general information while attending a group meeting during enrichment. HIPPY parent/carer/child receives individualised specific information or a referral/service connection.
Education and skills training	Tutor Training (HIPPY program content/curriculum) and/or group meetings (where the enrichment topic relates to parents receiving employment and study skills pathways).	HIPPY parent/carer/child receives specific information while attending a group meeting during enrichment, such as computer skills/digital literacy/online safety or new skills and interests for HIPPY parents. HIPPY tutor is undertaking education and skills training activities
Community capacity building	Group meetings (where the enrichment topic relates to events or cultural celebrations such as HIPPY week, harmony day, graduation, seasonal holidays, NAIDOC week), HIPPY events.	HIPPY parent/carer/child is attending group meetings where the enrichment topic relates to community capacity building.
Core component completed	NA	HIPPY participant has completed core components of the program at the time of the session.
Family capacity building	Home visits.	Home visits where a HIPPY parent/carer/child receives delivery of the HIPPY curriculum via the HIPPY activity packs. HIPPY parent/carer/child receives information while attending a group meeting where the enrichment topic relates to family capacity building.
Facilitate employment pathways	Tutor journey (professional development and training).	Group meetings with a HIPPY parent/carer, where the enrichment topic relates to employment and study skills pathways. It is also used when a HIPPY tutor engages in the pathways to possibilities process.
Service review	NA	To establish, through contact/time spent with the client, that they will no longer be considered an active participant, even though core components of the program may have not been completed.

Appendix B: Variables to control the differences between treated and untreated

HIPPY eligibility criteria

Variable	Category	Admin data source	Description
Family has a child aged four in the year of enrolment to HIPPY	Demographic	AEDC 2021: core module	AEDC contains the age of the children based on the child's reported date of birth. This is an integer variable. Children start school at different ages in each state and thus there are children between 4–7 in the dataset. These children were age-eligible for HIPPY between 2018–2021 (age 4 kids in 2021, age 5 kids in 2020, age 6 kids in 2019 and age 7 kids in 2018).
Family lives in an area where HIPPY is being delivered in the year of enrolment	Demographic	Core locations 2017–2021	This dataset contains the SA2 of the individuals and the start and end dates at these SA2s. Recruitment intervals will be defined for intake of kids for HIPPY each year between 2018–2021 based on their age. These intervals will range from last school term of previous year to end of June in the intake year (age 4 kids: 1 Oct 2020 to 30 April 2021, ¹⁰ age 5 kids: 1 Oct 2019 to 30 June 2020, age 6 kids: 1 Oct 2018 to 30 June 2019, age 7 kids: 1 Oct 2017 to 30 June 2018). Core locations data will be used to extract the SA2 of the children when they are at the age of 4. This will then be used to extract the SA3s and code whether the child was living in an SA3 where HIPPY is delivered when they are aged 4.

HIPPY priority criteria

Variable	Category	Admin data source	Description
Family holds a health care card	Economic	DOMINO: base module (2017–2021)	DOMINO contains information on individuals who holds a health care card. As we are interested in pre-treatment variables, we will look at the recruitment interval and check whether the kids are listed as a child on their parent/carer's health care card. We will code this as binary variable where 1 is when the child is attached to a health care card and 0 otherwise.
Child is Aboriginal and/or Torres Strait Islander	Demographic	Census 2021: person module	Census contains Indigenous status (INGP): Non-Indigenous (1), Aboriginal (2), Torres Strait Islander (3) or both (4). We will code this as binary variable where 1 is when the child is an Aboriginal, Torres Strait Islander or both and 0 when they are non-Indigenous.
Child resides in out-of-home care	Child wellbeing	Covered in 'Child resides with a carer'	N/A

¹⁰ Recruitment interval is defined before the AEDC. AEDC was conducted between May to August in 2021.

HIPPY priority criteria (continued)

Variable	Category	Admin data source	Description
Household has no income, or receives a government support payment as their primary source of income	Economic	Income Tax return 2017/18–2020/21	Household is considered the same as dwelling in the census. Using the dwelling id of the children we will obtain information on the people living in the same household as the kids. Income tax return contains information on government benefits received and total income. We will use this information to obtain the total benefits and total income for the households in the financial year that the children are age-eligible for HIPPY (financial year of the recruitment interval). We will then create a variable to check whether the major source of income is from government benefits. This variable will be coded as 1 if total government benefits contribute to more than 50% of the total household income and 0 otherwise.
Family is a single-parent family	Family	Census 2021: family module	Census contains information on family composition (FMCF). Using the family id of the children we will examine the family composition of the children. We will create a binary variable for single-parent family where all families with code starting with 3 will be coded as 1 to denote that they are from single-parent families and the remaining codes will be coded as 0 to indicate otherwise.
Child resides with a carer (not a parent)	Child wellbeing	Census 2021: person module	Census contains information on child type (CTPP): Foster child (4), Otherwise unrelated child – child who is related but not a natural, adopted, or stepchild (5), or Unrelated child (6). We will create a binary variable to code the above categories as 1 and 0 otherwise.
Main language spoken at home is not English	Family	Census 2021: person module	Census contains information on language used at home (LANP). We will create a binary variable to code all languages other than English to 1 and English as 0.
Child has a developmental delay or disability	Child wellbeing	AEDC 2021: core module	AEDC contains information on special needs status where 1 indicating that the child has special needs and 0 otherwise. Special needs status is based on 32 possible diagnoses, which include developmental delay and disability among others.
Parent/carer lives with a disability or long-term health condition	Family	Census 2021: person module	Census has information on count of selected long-term health conditions (CLTHP): None of the conditions (0), One condition (1), Two conditions (2), Three or more conditions (3). Census also contains a variable to indicate core activity need for assistance (ASSNP): Has need for assistance (1). For each parent/carer, we will create a variable combining information available in these variables where 1 indicates that the parent/carer has one or more long-term health conditions or has need for assistance and 0 indicates none of the selected long-term conditions and no need for assistance. For couple parents/carers we will pool this information and denote 0 if both parents/carers do not have any long-term health conditions or do not have any need for assistance, 1 if both parents have long health conditions, and 2 if one of them have long term health conditions or do have need for assistance.

Liddell et al 2011: Included variables (excluding those captured above)

Variable	Category	Admin data source	Description
Location (metropolitan/non-metropolitan; large/small)	Demographic	Remoteness area/Remoteness structure 2021	PLIDA contains information on the remoteness area at the SA2 level: Inner regional, outer regional, major cities. This will be used to construct a variable indicating the remoteness area of the children for the SA2s they were residing at the age of 4.
SEIFA index score of relative disadvantages	Socio-Economic	SA2 SEIFA index 2021	PLIDA contains SEIFA indices at the SA2 level. We will use the index of relative socio-economic disadvantage (ISRD) at the SA2s that the children were residing when they were 4.
% of local population under 5 years of age	Demographic	Census 2021: person module	We will use the census to calculate the percentage of population under 5 years at the SA3 level. This will then be used to construct a variable indicating the % of local population under 5 years for the SA3s that the children were residing at the age of 4.
'Who Am I?' developmental age school readiness score/classification, grouped according to developmental age	Child development	Not available	N/A
Number of people living in the household	Family	Census 2021: dwelling module	Census has information on the number of persons usually resident in dwelling (NPRD): 1–8. Using the dwelling id of the children, we will extract the number of people residing in the households.
Whether the parent was born in an English-speaking country	Family	Census 2021: person module	Census has information on the country of birth of mother (BPFP) and father (BPMP). We will use this information together with a list of English-speaking countries to create a variable where 1 indicates that the parent was born in an English-speaking country or 0 otherwise. For couple parents/carers, we will then pool this information and will denote 0 if both parents/carers were not born in English speaking countries, 1 if both were born in English speaking countries and 2 if one of them were born in an English-speaking country.
Employment status of parent/carer	Economic	Income tax return 2017/18–2020/21	Income tax return data contains information on wages or salary and business income. We will use this information to determine whether the parent/carer is employed in the financial year that the children are age-eligible for HIPPY. If the salary/wages or business income is greater than 0, then the parent/carer is considered employed and hence will be coded as 1 and 0 otherwise. Parents/carers who do not appear in income tax return data will be considered unemployed and hence will be coded as 0. As before, for couple parents/carers this information will be pooled and will be denoted as 0 if both are not employed, 1 if both of them are employed and 2 if one of them are employed.

Liddell et al 2011: Included variables (excluding those captured above) (continued)

Variable	Category	Admin data source	Description
Financial hardship scale	Economic	Covered in Household income	N/A
Home activities scale	Child development	Not available	N/A
Parent/carer's highest level of education	Family	Census 2021: person extract	Census contains information on the level of highest educational attainment (HEAP). This information will be extracted, and the 8 broad categories: postgraduate degree, bachelors degree, advanced diploma, graduate diploma, certificate I & II, certificate III & IV, 10 years and above and 9 years and below will be used. For couple parents/carers this information will be pooled to obtain the lowest education between the two.

Liddell et al 2011: Excluded variables (Round 1)

Variable	Category	Admin data source	Description
Partner's employment status	Family	Covered in parent's employment status	N/A
How well getting along, all things considered	Community	Not available	N/A
Psychological stress for parent/carer	Health	Census 2021: person module	Census contains information whether the person has mental health condition (including depression or anxiety) (HMHCP): yes (091), no (092). This information will be converted into a binary 1–0 variable. Then this information will be pooled for couple parents/carers where 0 indicates both do not have mental health condition, 1 if both of them have mental health condition and 2 if one of them have mental health condition.
Parental warmth	Child development	Not available	N/A
Whether child attending childcare	Child development	AEDC 2021: core module	AEDC contains information on whether the child has attended preschool or daycare: yes (1), no (0). This information will be directly used for this variable.

Liddell et al 2011: Excluded variables (Round 2)

Variable	Category	Admin data source	Description
Primary parent/carer gender	Demographic	Available but there is ambiguity in identifying primary parent/carer	N/A
Parent/carer marital status	Demographic	Census 2021: person module	Census contains information on social marital status (MDCP): registered marriage (1), defector marriage (2), not married (3). This information will be directly used for this variable.
Child gender	Demographic	AEDC 2021: core module	Information on the sex of the child: male (1), female (2) will be directly used for this variable.
Sense of belonging to the community	Community	Not available	N/A

Liddell et al 2011: Excluded variables (Round 3)

Variable	Category	Admin data source	Description
Parent/carer's age	Demographic	Census 2021: person module	Age is available in census (AGEP). We will calculate the average age of the parents/carers in the case of couple parents/carers.
Difficulty of life	Economic/Health	Covered elsewhere	N/A

Variables from literature review (excluding those captured above)

Variable	Category	Admin data source	Description
Parental motivation	Family	DEX: Attendance module 2017–2021	DEX contains attendance records for various grant programs. Of which the activity IDs of the programs that provides early learning and support services to children are identified. These are Communities for children facilitating pathways (10), Children and parent support services (16), Children and Family Intensive Support (17) and Budget based funded program (74). A binary variable will be created for parental motivation and any children who had participated in these programs prior to the end of their age-based recruitment period (defined above) will be considered as having motivated parents and hence will be coded as 1. (This assumes that children who had participated in other early childhood and parenting support programs recorded in DEX had parents who were motivated to support their learning.) Parental motivation for other children will be coded as 0.
Parental employment and resources	Family	Covered in Employment	N/A
Geography	Family	Covered in Location	N/A

Variables from DSS feedback (excluding those captured above)

Variable	Category	Admin data source	Description
Number of siblings	Family	Census 2021: family module	Census contains information on the count of dependent children under 15 in the family (CDCUF): 00-13. Using the family id of the children this information will be extracted.
Housing stability	Economic	Census 2021: dwelling module	Census contains information on household one year mobility indicator (MV1D): all (1), some (2), none (3) had different address. Using the dwelling id of the children this information will be used to construct a binary variable where 1 indicates none had different address (housing is stable) and 0 indicates all or some had different address.
Child physical health	Child wellbeing	MBS 2017–2020	MBS contains information on schedule fee (SCHEDFEE). This information will be used to extract total scheduled fee in the year prior to children being age-eligible for HIPPY.

Variables from site consultations (excluding those captured above)

Variable	Category	Admin data source	Description
Immunisation status	Child wellbeing	Australian Immunisation Register	Data custodian approval was not obtained.

Appendix C: Deviations from the pre-analysis plan since the review session

This pre-analysis plan was reviewed on 29 July 2025 at a session attended by representatives from ACE and DSS, as well as experts with experience in the method and the PLIDA datasets. The following changes were made after the review, based on data exploration and assumption testing, but before examining any effect estimates.

Version (Date)	Changes
v.1.1 (11/08/2025)	Immunisation status was not used to control for differences between the treatment and comparison groups, as ACE did not obtain approval from the data custodian to access AIR data within DataLab. Disclosure of AIR data must meet the purposes outlined in section 10 of the Australian Immunisation Register Act 2015. Research use is restricted to vaccination coverage or uptake and does not extend to the purpose of this study.
v.1.1 (11/08/2025)	The Indigenous status of the child was sourced from the AEDC 2021: Indigenous module rather than the 2021 Census to minimise missing data. The AEDC records whether a child is of Aboriginal or Torres Strait Islander descent and contains fewer missing values compared to the Census, which had a higher proportion of missing responses for this variable. Accordingly, the AEDC was selected as the preferred source.
v.1.1 (11/08/2025)	'Main language spoken at home is not English' was extracted from the AEDC 2021 rather than the 2021 Census to reduce missing values for this variable. The AEDC records whether a child speaks a language other than English at home, and this measure was used directly.
v.1.1 (11/08/2025)	Simple imputation was applied to address missing values in the variables used to control for differences between the treatment and comparison groups. An assessment was conducted to determine whether missingness was related to program participation. As no such relationship was found, the missingness was considered random. Observations with missing outcome data were excluded from the analysis.
v.1.1 (11/08/2025)	'Single-parent family' and '% of local population under 5 years of age' were removed from the propensity score model as these were causing singularities in the model due to multicollinearity. Single-parent family is correlated with social marital status and % of local population under 5 years of age is correlated with a linear combination of other variables.
v.1.1 (11/08/2025)	Extreme propensity scores were observed only at the lower tail and clipped up at 0.01. This was performed after inspecting that there is good common support between the treated and comparison groups. Three different thresholds for trimming are inspected (0.001, 0.01 and 0.05). The threshold of 0.01 was chosen after verifying that the results were robust to this trimming.
v.1.2 (18/08/2025)	Outcome regression models were specified as ordinary least squares regression for all primary outcomes. The 3-level categorical outcomes—adapting to school and parent engagement with the school—contained only a few observations in the 'not true' category, and the 'somewhat true' category is ambiguous and lacks a clear interpretation. To improve interpretability and the reliability of estimates, these outcomes were coded as binary by combining not true and somewhat true into a single category. These were then modelled using a linear probability model.
v.1.3 (28/08/2025)	We have investigated omitted variable bias qualitatively rather than quantitatively. We have not used methods such as Altonji, Elder, and Taber (2005) and Oster (2015) as they are developed for linear regression coefficients (OLS, and R^2 as model fit measure). As we are estimating the ATT via DR IPW using an outcome regression, it is not a regression coefficient in the same sense.