NSW GOVERNMENT

8

Technical report

Pathways of Care Longitudinal Study: Outcomes of Children and Young People in Out-of-Home Care

Measures Manual (Waves 1-4)

Pathways of Care Longitudinal Study: Outcomes of Children and Young People in Out-of-Home Care in NSW

Technical Report No. 8

Measures Manual (Waves 1-4)

Published by
New South Wales Department of Communities and Justice (DCJ)
Insights Analysis and Research
320 Liverpool Road
Ashfield NSW 2131

Phone + 61 2 9716 2222

August 2020

ISBN: 978-0-6485156-9-2

Recommended citation
NSW Department of Communities and Justice (2020). Measures Manual (Waves 1-4). Pathways of Care Longitudinal Study: Outcomes of Children and Young People in Out-of-Home Care. Technical Report Number 8. Sydney. NSW Department of Communities and Justice.

Prepared by
Lucy Tully, Marina Paxman, Albert Zhou, Johanna Watson, Nafisa Asif, Sharon Burke and Merran Butler - NSW Department of Communities and Justice (DCJ).
Mark Bartlett, Tina Navin Cristina and Sarah Baynes - Sax Institute.
Australian Institute of Family Studies scored all the measures for Wave 1.

Pathways of Care Longitudinal Study Clearinghouse
All study publications including research reports, technical reports and evidence to action notes can be found on the study webpage www.facs.nsw.gov.au/resources/research/pathways-of-care

Study design by NSW Department of Communities and Justice Insights, Analysis and Research; Australian Institute of Family Studies; Professor Judy Cashmore, University of Sydney; Professor Paul Delfabbro, University of Adelaide; Professor Ilan Katz, University of NSW; Dr Fred Wulczyn, Chapin Hall Center for Children University of Chicago.

Study data collection by I-view Social Research.

Advisors Expert advice and support has been provided by the CREATE Foundation; Aboriginal Child, Family and Community Care State Secretariat (AbSec); My Forever Family NSW; and program areas.

Acknowledgements We wish to extend our thanks to all the children, young people and caregivers who participated in interviews; childcare teachers, school teachers and caseworkers who participated in online surveys; and the data custodians in the relevant NSW and Commonwealth government departments. Ms Sammy Verma grew up in care and played a key role in the production of the study video for children and stakeholders. Ms Billy Black also grew up in care and designed the study artwork. Ms Sammy Verma and Mr Samuel Eyeson-Annan both did the voiceover for the audio computer-assisted self-interview (ACASI) for the child/young person interview.

Ethics approval by The University of NSW Human Research Ethics Committee (approval number HC10335 & HC16542); Aboriginal Health and Medical Research Council of NSW Ethics Committee (approval number 766/10); NSW Department of Education and Communities State Education Research Approval Process (SERAP, approval number 2012250); NSW Population & Health Services Research Ethics Committee (Ref: HREC/14/CIPHS/74 Cancer Institute NSW: 2014/12/570).

NSW GOVERNMENT Communities & Justice

SaxInstitute

Contents

Preface

The Pathways of Care Longitudinal Study (POCLS) is funded and managed by the New South Wales Department of Communities and Justice (DCJ). It is the first large-scale prospective longitudinal study of children and young people in out-of-home care (OOHC) in Australia. Information on safety, permanency and wellbeing is being collected from various sources. The child developmental domains of interest are physical health, socio-emotional wellbeing and cognitive/learning ability.

The overall aim of this study is to collect detailed information about the life course development of children who enter OOHC for the first time and the factors that influence their development. The POCLS objectives are to:

- describe the characteristics, child protection history, development and wellbeing of children and young people at the time they enter OOHC for the first time
- describe the services, interventions and pathways for children and young people in OOHC, post restoration, post adoption and on leaving care at 18 years
- describe children's and young people's experiences while growing up in OOHC, post restoration, post adoption and on leaving care at 18 years
- understand the factors that influence the outcomes for children and young people who grow up in OOHC, are restored home, are adopted or leave care at 18 years
- inform policy and practice to strengthen the OOHC service system in NSW to improve the outcomes for children and young people in OOHC.

The POCLS is the first study to link data on children's child protection backgrounds, OOHC placements, health, education and offending held by multiple government agencies; and match it to first-hand accounts from children, caregivers, caseworkers and teachers. The POCLS database will allow researchers to track children's trajectories and experiences from birth.

The population cohort is a census of all children and young people who entered OOHC for the first time in NSW over the 18 month period between May 2010 and October 2011 (n=4,126). A subset of those children and young people who went on to receive final Children's Court care and protection orders by 30 April 2013 (2,828) were eligible to participate in the interview component of the study. For more information about the study please visit the study webpage www.facs.nsw.gov.au/resources/research/pathways-of-care

The POCLS acknowledges and honours Aboriginal people as our First Peoples of NSW and is committed to working with the DCJ Aboriginal Outcomes team to ensure that Aboriginal children, young people, families and communities are supported and empowered to improve their life outcomes. The POCLS data asset will be used to improve how services and supports are designed and delivered in partnership with Aboriginal people and communities.

DCJ recognises the importance of Indigenous Data Sovereignty (IDS) and Indigenous Data Governance (IDG) in the design, collection, analysis, dissemination and management of all data related to Aboriginal Australians. The POCLS is subject to ethics approval, including from the Aboriginal Health & Medical Research Council of NSW. The DCJ is currently in the process of scoping the development of IDS and IDG principles that will apply to future Aboriginal data creation, development, stewardship, analysis, dissemination and infrastructure. The POCLS will continue to collaborate with Aboriginal Peoples and will apply the DCJ research governance principles once developed.

Summary of the measures used in the POCLS

One of the key aims of the POCLS is to collect data on child development for a cohort of children and young people who entered OOHC for the first time between May 2010 and October 2011 and examine the factors that are associated with developmental outcomes. A number of standardised developmental and psychosocial wellbeing measures are used in the POCLS caregiver, child and teacher surveys during Waves 1 to 4.

This manual provides:

- Section A: a short description of each measure used in the POCLS
- Section B: detailed information of each measure including respondent type, the age range, availability of norms, scoring and whether used in other studies
- Section C: information about additional questions used in the POCLS that are either project developed or used by other published studies.

The table below provides a list of the standardised measures, activities and questions used in the POCLS during Waves 1 to 4 to examine children's wellbeing, caregiver characteristics and placement characteristics.

Domain	Standardised measures and questions	Scored Wave 1	Scored Wave 2	Scored Wave 3	Scored Wave 4
Child measures					
Physical health and development	Ages and Stages Questionnaire (ASQ3; Squires & Bricker, 2009) – Child age range 9 months-5 years	Yes	Yes	Yes	NA
	Additional questions (either project developed or used by other studies) about the child's health conditions, services received, immunisation, diet, weight, sleep.				
Child socio-emotional development	Abbreviated Temperament Scales adapted from the Revised Infant Temperament Questionnaire (Carey & McDevitt, 1978), the Toddler Temperament Questionnaire (Fullard, McDevitt & Carey, 1978) and the Childhood Temperament Questionnaire (Thomas & Chess, 1977) – Child age range 9 months-7 years	Yes	Yes	Yes	Yes
	School Aged Temperament Inventory (SATI; McClowry,1995) – short form – Child age range 8-17 years	Yes	Yes	Yes	Yes
	Brief Infant Toddler Social Emotional Assessment (BITSEA; Briggs-Gowan et al. 2004) – Child age range 12-35 months	Yes	NA	NA	NA

Domain	Standardised measures and questions	Scored Wave 1	Scored Wave 2	Scored Wave 3	Scored Wave 4
	Child Behaviour Checklist 1.5–5 and 6–18 (CBCL; Achenbach & Rescorla, 2000; 2001) caregiver and teacher (childcare and school) report – Child age range 3-17 years	Yes	Yes	Yes	Yes
	Ages and Stages Questionnaire (ASQ3; Squires & Bricker, 2009) – Child age range 9 months-5 years	Yes	Yes	Yes	NA
	School Problems Scale (Prior, Sanson, Smart & Oberklaid, 2000) – Child age range 12-17 years	Yes	Yes	Yes	Yes
	School Bonding Scale (O'Donnell, Hawkins & Abbott, 1995) – Child age range 12-17 years	Yes	Yes	Yes	Yes
	Short Mood and Feeling Questionnaire 13-item scale (Angold et al. 1995) and additional questions on mood – Child age range 12-17 years	Not asked at W1	Yes	Yes	Yes
	Self Report Delinquency Scale 13-item scale (Moffitt & Silva, 1988) – Child age range 12-17 years	Not asked at W1	Yes	Yes	Yes
	Emotional Responsiveness Scale from the Parenting Style Inventory II, adapted version (PSI-II: Darling & Toyokawa, 1997) - Child age range 7-17 years	Yes	Yes	Yes	Yes
	Adapted Kvebaek Family Sculpture Technique (Kvebaek Family Sculpture Technique; Cromwell, Fournier & Kvebaek, 1980) – Child age range 7-17 years	Yes	Yes	Yes	Yes
	Additional questions (project developed or used in other studies) about services and supports for the child's emotional and behavioural problems, problems at school, psychotropic medication, relationship with peers/ caregivers/ caseworkers.				
Cognitive and language development	Communication and Symbolic Behaviour Scale Infant and Toddler Checklist (CSBS ITC; Wetherby & Prizant, 2003) – Child age range 9-23 months	Yes	NA	NA	NA
	MacArthur Communicative Development Inventories—Short form (Fenson et al. 2000) – Child age range 24-29 months	Yes	Yes	NA	NA

Domain	Standardised measures and questions	Scored Wave 1	Scored Wave 2	Scored Wave 3	Scored Wave 4
	MacArthur-Bates Communicative Developmental Inventories (MCDI-III; Fenson et al. 2007) – Child age range 30-35 months	Yes	Yes	NA	NA
	Peabody Picture Vocabulary Test (PPVT-IV; Dunn & Dunn, 2007) – Child age range 3-17 years	Yes	Yes	Yes	Yes
	Matrix Reasoning Test from Wechsler Intelligence Scale for Children (WISC-IV; Wechsler, 2003) – Child age range 6-16 years	Yes	Yes	Yes	Yes
	Additional questions (project developed or used in other studies) about the child's current schooling (usual grades at school, changes in schools, repeated years, school problems), for young people aged 15 years and older, questions on work, further education, aspirations, life skills and plans for leaving care.				
Caregiver measures					
Caregiver psychological distress	Kessler K10 (Kessler et al. 2003)	Yes	Yes	Yes	Yes
Social cohesion	Social Cohesion and Trust Scale (Sampson, Raudenbush & Earls, 1997)	Yes	Yes	Yes	Yes
Parenting practices/ style/self-efficacy	Parenting – Warmth (Paterson & Sanson, 1999)	Yes	Yes	Yes	Yes
	Parenting – Hostility (Institut de la Statistique du Québec, 2000)	Yes	Yes	Yes	Yes
	Parenting – Monitoring (Goldberg et al. 2001) – Child age range 12-17 years	Yes	Yes	Yes	Yes
	Difficult Behaviour Self-Efficacy Scale (DBSES; Hastings & Brown, 2002)	Yes	Yes	Yes	Yes
	Additional questions (project developed or used in other studies) about the caregiver's socio-demographic characteristics, cultural background and cultural activities, health, relationship with partner and study child, caregiver experience and training, satisfaction with support from services, family activities, support network, satisfaction with caregiving role.				

The measures are standardised meaning they can be used to show how a cohort of children compare with peers in the general population and also how individuals are developing. It is important to take cultural considerations into account when using standardised measures. The standardised measures used in the POCLS were selected in 2010 at which time measures of child development had not been tested for validity with Aboriginal children in Australia. For Aboriginal people in urban settings, non-verbal performance based tests that are less reliant on language skills such as the Matrix Reasoning Test from the Wechsler Intelligence Scale for Children IV (WISC-IV) have been found to be comparable to existing Australian norms[1]. The Child Behaviour Checklist (CBCL) has been tested in a range of diverse cultures but clinical cut-offs may not be uniform across all cultures[2]. The measures may not be sensitive to the influence cultural norms may have on reporting child behaviours and parents' problem ratings. This should be considered when interpreting the data.

[1] Westerman, T. and Wettinger, M. (1997). Psychological Assessment of Aboriginal people. Psychologically Speaking.

[2] Crijnen A.A.M, Achenbach, T.M. and Verhulst, F.C. (1997). Comparisons of Problems Reported by Parents of Children in 12 Cultures: Total Problems, Externalizing, and Internalizing, Journal of the American Academy of Child & Adolescent Psychiatry, Volume 36, Issue 9, 1997, Pages 1269-1277.

Section A: Description of each measure

Section A provides a short description of each measure and is intended to be a guide for researchers and to be used in drafting the methodology and measures sections of research publications. Section B provides more detailed information on each measure.

1. Children's wellbeing

1.1 Physical health and development

Ages and Stages Questionnaire (ASQ-3)

The Ages and Stages Questionnaire (ASQ-3) was designed as a screening tool for infants and young children who are at risk of development delays or disorder (Squires & Bricker, 2009). The ASQ-3 is composed of 21 questionnaires designed to be completed by caregivers of children aged between 1 and 66 months. Each questionnaire contains 30-items across five developmental domains: Communication, Gross Motor, Fine Motor skills, Problem Solving, and Personal-Social. The focus of each domain is:

- Communication addresses parental concerns regarding babbling, vocalising, listening and understanding
- Gross Motor focuses on arm, body and leg movements
- Fine Motor pertains to hand and figure movements
- Problem Solving addresses learning and playing with toys
- Person-Social focuses on social play and play with toys and other children.

For each item, caregivers mark Yes, Sometimes or Not yet to indicate that their child performs, occasionally performs or does not yet perform the behaviour. The responses are scored as Yes = '10', Sometimes = '5' and Not yet = '0'. The ASQ-3 uses empirically driven cut-off scores to determine whether the child is developing appropriately, requires close monitoring or further assessment. In the POCLS, The ASQ-3 was competed by caregivers from Wave1 to 3 for children aged 9-66 months – by Wave 4 the children in the POCLS were no longer in the age range for this measure.

1.2 Child socio-emotional development

Abbreviated Temperament Scales for Infants, Toddlers and Children

For children aged 9 months to 7 years, the abbreviated Temperament Scale for Infants includes 12-items and assesses approach/sociability, irritability and cooperation (Carey & McDevitt, 1978). The Temperament Scale for Toddlers

includes 13-items and assesses approach/sociability, reactivity and persistence (Fullard, McDevitt & Carey, 1978); and the Childhood Temperament Scale includes 12-items and assesses approach/sociability, reactivity and persistence (Thomas & Chess, 1977). These are scored on a 6-point scale. The response categories are: almost never; not often; variable, usually does not; variable, usually does; frequently, and almost always. To score each subscale, the mean is computed and a high score reflects high sociability, reactivity and persistence.

School Aged Temperament Inventory (SATI) – short form

For children and young people aged 8-17 years, the short form of School Aged Temperament Inventory (SATI) was used (McClowry, 1995). It was developed for the Longitudinal Study of Australian Children (LSAC), and assesses three dimensions of temperament: negative reactivity, task persistence, and approach/withdrawal. The short form contains 12-items which are scored on a 5-point scale. Following reverse coding of certain items, the mean of the items in each scale is computed to yield 3 composite scores. Higher scores reflect higher negative reactivity, higher persistence and a higher degree of introversion.

Brief Infant Toddler Social Emotional Assessment (BITSEA)

The Brief Infant Toddler Social Emotional Assessment (BITSEA) is a screening tool that identifies children (age 12-35 months) that may require further assessment to identify clinically significant social/emotional and behavioural problems and/or delays/deficits in social-emotional competence (Briggs-Gowan & Carter, 2006). There are forms to be completed by parents and childcare providers. It is recommended that adoptive/foster carers complete the form after the child has been in their care for at least one month. The forms can be self-complete or as part of a structured interview.

The tool covers two domains: Problem (31-items) and Competence (11-items). There are 42-items and the response categories are; ‘0’ = Not True/Rarely, ‘1’ = Somewhat True/Sometimes, ‘2’ = Very True/Often. Each domain yields a total score. Higher Problem scores indicate greater levels of socio-emotional or behaviour problems and lower Competence scores indicate a possible delay or deficit. The summed raw scores are compared to empirically derived cut-off scores for four age groups and by sex to determine percentile ranks. Problem cut-off scores corresponding to the 25th percentile ranking or less indicate possible socio-emotional or behaviour problems. This means that the child’s Problem score was higher than 75% of children of the same age and sex from the normative sample. Competence cut-off scores corresponding to the 15th percentile ranking or less indicate possible deficit/delay in competence. This means that the child’s Competence score was lower than 85% of children of the same age and sex from the normative sample. These thresholds vary as the prevalence for socio-emotional problems is expected to be greater than for delay/deficits (Briggs-Gowan et al. 2004; Briggs-Gowan & Carter, 2006). It is

important to note that this indicates possible problems or deficits/delay and further assessment is required to determine if these are clinically significant.

In the POCLS the BITSEA was completed by caregivers in Wave 1 only for children aged 12-35 months of age. From Wave 2 onwards caregivers completed the Child Behaviour Checklist for children of all ages.

Child Behaviour Checklist (CBCL)

The Child Behaviour Checklist (CBCL) is a questionnaire used to assess behavioural and emotional problems in children and young people. It is one component of the Achenbach System of Empirically Based Assessments (ASEBA) that enables you to obtain standardised ratings on functioning to identify children that may benefit from additional support (Achenbach & Rescorla, 2000, 2001)

In the POCLS the CBCL was completed by caregivers of children aged 3-17 years from Wave 1 onwards. Versions validated and normed for use for children 18 months to 5 years of age (CBCL/1.5-5) and 6-18 years (CBCL/6-18) of age were used. The CBCL/1.5-5 years contains 100-items and the CBCL/6-18 contains 120 problem items. All items are rated on a scale from '0' = not true, '1' = somewhat or sometimes true and 2 = very true or often true.

The CBCL can be self-administered and yields subscale scores for a range of conditions and competencies. The principal focus for the POCLS were two composite syndrome profiles: internalising and externalising. Internalising includes the anxious-depressed, withdrawn-depressed and somatic complaints syndrome scales. For the CBCL/1.5-5 internalising also includes the withdrawn syndrome. Externalising captures problems relating to external behaviours including the rule breaking and aggressive behaviours scales. The CBCL Total Problems Score is the sum of all items including internalising, externalising, other syndromes and other problems. The CBCL scores can be presented in a raw score format; as standardised T-scores or children can be classified as falling into clinical, borderline and non-clinical ranges. For statistical analysis of Internalising, Externalising and Total Problems, raw and T scores will usually produce similar results (Achenbach & Rescorla, 2001, p. 138).

School Problems Scale

A 4-item short form version of School Problems Scale was completed by young people aged 7-17 years (Prior, Sanson, Smart & Oberklaid, 2000). Respondents were asked how often they: 'Find someone to have lunch with'; 'Understand the work in class'; 'Follow school rules and routines'; and 'Get assignments, projects and homework done' (extra question for 12-17 year olds). Response categories ranged from '1' = Always to '5' = Never. Lower mean scores indicate fewer problems at school.

School Bonding Scale

The School Bonding Scale is a 4-item scale completed by young people aged 7-17 years and measures to what extent they are settling in and forming relationships at school (O'Donnell, Hawkins & Abbott, 1995). Each item: 'Try hard'; 'Get on well with your teachers'; 'Feel it is important to do well at school' (extra question for 12-17 year olds); 'Enjoy being there' is scored on a 5-point scale from Always = '1' to Never = '5'. Lower mean scores across the items indicate better school bonding.

Short Mood and Feeling Questionnaire (SMFQ)

The Short Mood and Feeling Questionnaire (SMFQ) is a brief measure of depression that was self-completed by young people aged 12-17 years from Wave 2 onwards (Angold, Costello, Messer, Pickles, Winder & Silver, 1995). The questionnaire contains 13-items that describe mood or feelings which may have been experienced in the past 2 weeks. The response categories are: Not true = '0'; Sometimes = '1'; True = '2'. Responses are summed and cut-off scores can be used to identify young people with significant depression.

Short Self Report Delinquency Scale (SRDS)

A short form of the Self Report Delinquency Scale (SRDS) comprising 4-items used to assess self-reported levels of anti-social behaviour in young people aged 10-11 years and 12-items used to assess self-reported levels of anti-social behaviour in young people aged 12-17 years from Wave 2 onwards (Moffitt & Silva, 1988). Items 1-11 are scored on a scale from Not at All = '1' to More than 20 times = '9', while items 12 and 13 about whether the respondent used alcohol and/or an illicit drug in the last 12 months have Yes = '1' or No = '2' response options. All items are recoded to yield a final score of either 0 = no/low antisocial behaviour or 1 = highly antisocial behaviour.

Emotional Responsiveness Scale from the Parenting Style Inventory II (PSI-II)

The Emotional Responsiveness Scale from the Parenting Style Inventory II, adapted version (PSI-II: Darling & Toyokawa, 1997) is used in the POCLS from Wave 1 onwards to measure the relationship between young people aged 7-17 years and their caregivers. The version used in the POCLS comprises of 5-items which asks young people how often their caregiver: 'Help you out if you have a problem'; 'Listen to you'; 'Praise you for doing well'; 'Do things with you that are just for fun'; and 'Spend time talking to you'. Each item is scored on a 5-point scale from Always = '1' to Never = '5'. The items are then reverse coded, with a higher score indicating a better parenting style.

Adapted Kvebaek Family Sculpture Technique

The Kvebaek Family Sculpture Technique (KFST) is a symbolic figure placement procedure used in family assessment and research (Cromwell, Fourier & Kvebaek, 1980). The technique was adapted for the POCLS to measure a child's view of how close (special and important) they feel to the people they are living with, and what relatives/kin they feel close to who they are not living with. The physical distance represented on a checkerboard between the child and family members is interpreted as a measure of experienced psychological distance (e.g., closeness, belongingness, cohesion). The POCLS activity adapted from the Kvebaek Family Sculpture Technique was used with children aged 7-17 years from Wave 1. From Wave 2 onwards, similar questions were added to the child Audio Computer Assisted Self Interview (ACASI) for older children who opted not to complete the activity (i.e., viewed as not age appropriate).

1.3 Cognitive and language development

Communication and Symbolic Behaviour Scale Infant-Toddler Checklist (CSBS ITC)

The Infant-Toddler Checklist (ITC) is one component of the Communication and Symbolic Behaviour Scales Developmental Profile (CSBS DP) (Wetherby & Prizant, 2002). The checklist comprises 24-items designed to be completed by the caregivers of children aged 9-23 months to measure language development and symbolic abilities in seven different cluster areas in Wave 1 of the POCLS. These are: demotion and use of eye gaze; use of communication; use of gestures; use of sounds; use of words; understanding of words; and use of objects (Wetherby & Prizant, 2001). Within clusters, the item response ranges from two to four points. These points are totalled to yield individual cluster scores and then the cluster scores are summed to yield three composite scores: social composite; speech composite; and symbolic composite. The three composite scores are then summed to yield a total raw score. Based on the chronological age of the child, the composite scores and the total raw score are converted to a normed score. Cut-off scores, standard scores, and percentile ranks are used to identify potentially concerning scores. In the POCLS, CSBS ITC was used only in Wave 1. It is recommended that the Checklist should only be used to decide that further evaluation on development is needed.

MacArthur Communicative Development Inventories—short form

A short form of the MacArthur Communicative Development Inventories was used in Waves 1 and 2 of the POCLS with children aged 24-29 months of age to assess communication skills (Fenson, Pethick, Renda, Cox, Dale & Reznick, 2000). There are two equivalent word checklists: Form A and Form B, both with 100-items. The number of words marked by the caregiver as being used by the child are summed to give a score between 0 and 100. The number of words used by the child are marked

by the caregiver on a Yes or No response format. No = '0' indicates caregiver have not heard the child used that word and Yes = '1' indicates that the child used the word. The total scores ranges from 0 to 100 and summated raw scores are then standardised.

MacArthur-Bates Communicative Developmental Inventory-IV (MBCDI-IV)

The MacArthur-Bates Communicative Developmental Inventory-IV (MBCDI-IV) was used to assess communication skills in children aged 30-35 months of age in Waves 1 and 2 of the POCLS (Fenson, Marchman, Thal, Dale, Reznick & Bates, 2007). The Longitudinal Study of Australian Children (LSAC) version of the inventory was used which includes 98 vocabulary words and 12 sentence pairs. The number of words in the vocabulary checklist that the caregiver marks are summed to give a score between 0 and 98. Similarly, each of the sentence pairs are scored '1' if caregiver marks that the child uses the more complex sentence, giving a maximum score of 12.

Peabody Picture Vocabulary Test (PPVT-)

The Peabody Picture Vocabulary Test Fourth Edition is an untimed test of receptive vocabulary and is interviewer administered to children in the POCLS aged 3-17 years from Wave 1 onwards (Dunn & Dunn, 2007). Children are shown a set of 4 pictures on a page and asked to point to the picture that matches the word spoken by the interviewer. There are 228-items with different starting points for children of different ages. The test yields raw scores based on correct answers and errors as well as standardised scores (M = 100, SD= 15) for different ages. Scores higher or lower than the reference point of 100 indicate the extent to which the child's vocabulary compares with same age peers (Dunn & Dunn, 2007).

Matrix Reasoning Test from Wechsler Intelligence Scale for Children (WISC-IV)

The Matrix Reasoning Test from the Wechsler Intelligence Scale for Children (WISC-IV) is a measure of logical reasoning or fluid intelligence (Wechsler, 2003). The 35-item test is administered to children in the POCLS aged 6-16 years by an interviewer from Wave 1 onwards. The test has age-based starting points. Children are shown a set of pictures with 1 missing square and asked to choose from 5 options the picture that best fits the missing square. One point is scored for a correct response and zero for an incorrect response or no response. Raw scores are converted to scale scores which range from 1-19. Scores are standardised based on age norms provided in the WISC-IV manual.

2. Caregiver and placement characteristics

2.1 Caregiver psychological distress

Kessler K10

The 10-item Kessler (K10) measure was completed by all caregivers in the POCLS from Wave 1 onwards to measure their own psychological distress (Kessler, Barker, Colpe, Epstein, Gfroerer, Hiripi et al. 2003). Completion of this scale requires caregivers to rate how often a series of statements applied to them in the last 30 days. The response categories were a 5-point scale ranging from All of the time = '1' to None of the time = '5'.

2.2 Social cohesion

Social Cohesion and Trust Scale

The Social Cohesion and Trust Scale is a caregiver-rated scale used to measure neighbourhood social cohesiveness in the POCLS from Wave 1 onwards (Sampson, Raudenbush & Earls, 1997). One item of the scale was not used after it was found to be inappropriate for Aboriginal respondents. The 4 remaining item statements were: 'This is a close-knit neighbourhood'; 'People around here are willing to help their neighbours'; 'People in this neighbourhood generally don't get along with each other'*; and 'People in this neighbourhood can be trusted'. The response categories were scored on a 5-point scale from Strongly agree = '1' to Strongly disagree = '5'. The third item (*) is reverse scored. The scale is scored by summing the 4-items.

2.3 Parenting practices/styles and self-efficacy

Parenting – Warmth

The Parenting – Warmth scale was used in the POCLS from Wave 1 onwards to assess the warmth of parenting practices for children aged 9 months-17 years (Paterson & Sanson, 1999). The scale consists of 4-items which asks caregivers: 'How often do you tell [Study Child] how happy [he/she] makes you?', 'How often do you have warm, close times together with [Study Child]?'; 'How often do you enjoy listening to [Study Child] and doing things with him/her?'; 'How often do you feel close to [Study Child] both when he/she is happy and when he/she is upset?'. The response categories range from Never/almost never = '1' to Always/almost always = '5'.

Parenting – Hostility

The Parenting – Hostility scale was used in the POCLS from Wave 1 onwards to measure hostile/angry parenting for children aged 9 months-17 years (Institut de la

Statistique du Quebec, 2000). The scale is brief and is completed by caregivers. The scale consists of the following 3-items: 'I have been angry with [study child]'; 'When [study child] cries, he/she gets on my nerves': and 'I have lost my temper with [study child]'. The response is scored on a 10-point scale from Not at all = '1' to All the time = '10'.

Parenting – Monitoring

The Parenting – Monitoring scale was used in the POCLS from Wave 1 onwards to assess parenting practices of monitoring and supervising young people aged 12-17 years (Goldberg, Spoth, Meek & Moolgard, 2001). The scale consists of 4-items which asks the caregiver: 'How often do you know where [study child] is in the course of a day?'*; 'How often do you know who [study child] is with when [he/she] is away from home (i.e. their placement)?'*; 'How often do you talk to [study child] about what's going on in [his/her] life?'*; 'How often does the [study child] go out without telling you where [he/she] will be?. The response categories range from Always = '1' to Never = '5'. The first 3-items* are reverse scored and then items are summed to give a total between 4 (low monitoring) and 20 (high monitoring).

Difficult Behaviour Self-Efficacy Scale (DBSES)

The Difficult Behaviour Self-Efficacy Scale (DBSES) measures caregiver self-efficacy when dealing with challenging behaviours (Hastings & Brown, 2002). It was used in the POCLS from Wave 1 onwards with caregivers of children aged 9 months to 17 years. Four items of the scale were used, these asked: 'How difficult do you personally find it to deal with the challenging behaviours of [Study Child]?; 'To what extent do you feel that the way you deal with the challenging behaviours of [Study Child] has a positive effect?'; 'How satisfied are you with the way in which you deal with challenging behaviours of [Study Child]?'; 'To what extent do you feel in control of the challenging behaviours of the [Study Child]?'. Each item was rated on a 7-point scale.

3. Measures completed by childcare and school teachers

3.1 Child socio-emotional development

Caregiver-Teacher Report Form (C-TRF) and Teacher Report Form (TRF) (CBCL)

The Caregiver-Teacher Report Form (C-TRF), like the CBCL, is a component of the ASEBA and a parallel version of the CBCL 1.5-5 years for completion by childcare or preschool teachers to assess behavioural and emotional problems in children. The Teacher Report Form (TRF) is also part of the ASEBA and a parallel version of the CBCL for completion by a school teacher or other school personnel who are familiar with children's functioning in school, such as teacher aides, counsellors, administrators and special educators to assess children aged 6-17 years. The C-TRF and the TRF are the only standardised measures included in the POCLS teacher survey. The C-TRF and TRF are scored similarly to the CBCL and the POCLS focussed on the Internalising, Externalising and Total Problem profiles.

Section B: Detailed information of each measure

1. Children's wellbeing

1.1. Physical health and development

Ages and Stages Questionnaire (ASQ-3)

Summary	In the POCLS, the Ages and Stages Questionnaire (ASQ-3) was used for children aged 9-66 months to identify and monitor development delays in five domains: Communication; Gross Motor; Fine Motor; Problem Solving; and, Personal Social (Squires et al. 2009). Each domain contains 6-items, and for each item the response options are Yes, Sometimes and Not yet. The ASQ-3 uses empirically driven cut-off scores to determine whether the child is developing appropriately, requires close monitoring or further assessment.
Domains	Five development domains: Communication; Gross Motor; Fine Motor; Problem Solving; Personal-Social.
Respondent type	Computer Assisted Personal Interviewing (CAPI), Caregiver.
Waves	Waves 1-4.
Age range	Children aged 9-66 months.
Length	The POCLS used 18 questionnaires depending on the age intervals (8, 9, 10, 12, 14, 16, 18, 20, 22, 24, 27, 30, 33, 36, 42, 48, 54, 60 month) and each questionnaire contains 30-items on five domains.
Publisher/ cost/ permission	Permission granted by publisher for use in the POCLS. Agreement to computerise and use 1,000 copies of the ASQ-3.
Psychometric properties	Validity: -Sensitivity: 82.5-89.2%; Specificity: 77.9-91.3% (Reliability: Test-Retest reliability: over a 2-week interval, intra-class correlations ranged from 0.75 - 0.82. -Inter-observer reliability: intra-class correlations ranged from 0.43 to 0.69.

	-Internal consistency: Cronbach alphas for five development areas for 20 age interval questionnaire range from 0.51 to 0.87. (Squires et al. 2009).
Studies used/ Rationale for use	This scale measures child development across 5 domains of functioning. This measure is widely used and is relatively brief compared to other developmental measures.
Scoring information	Each domain has 6-items and, for each item the following scores are given: Yes = '10' Sometimes = '5' Not yet = '0' The scores are summed to get a total score for each development domain. A domain should not be scored if there are more than 2 missing responses. In that case, an adjusted total score is computed by dividing the total domain score by the number of items answered in that domain. The adjusted score is then added to the other item scores, producing a total score. The ASQ-3 uses empirically determined cut-off scores to identify child developmental delays or disorders. The cut-off points differ by age and domain. Each questionnaire (age interval) has standardised cut-off points for each domain indicating that the child's development is in the normal range (typical development), the child is near the cut-off for further assessment (need monitoring), or that the child is in need for further assessment (need professional support). These cut-off points are listed in the manual. However, the manual does not include the equivalence of the cut-offs for each domain across all age intervals (questionnaires). As a result, raw score for each domain cannot be used in the analysis.
Additional information	Changes were made to wording to make the questions appropriate for the Australian context and to enable administration by an interviewer without the need for any carer testing of child skills. These changes were approved by the publisher.
References	Squires, J., & Bricker, D. (2009). Ages & Stages Questionnaires, Third Edition (ASQ-3). Baltimore, MD: Brookes Publishing. Squires, J., Twombly, E., Bricker, D., Potter, L. (2009). ASQ-3 Ages and *Stages Questionnaires Third Edition, User's Guide*. Baltimore, MD: Paul H Brookes Publishing.

1.2 Child socio-emotional development

Abbreviated Temperament Scales for Infants, Toddlers and Children

Summary	The Abbreviated Temperament Scale measures temperament in infants, toddlers and children by assessing 3 dimensions of temperament. The Infant Temperament Questionnaire (12-items) assesses Approach/Sociability, Irritability and Cooperation; and the Toddler Temperament Questionnaire (13-items) and the Childhood Temperament Questionnaire (12-items) assesses Approach/Sociability, Reactivity and Persistence. The items are answered on a 6 point scale: '1'=Almost never to '6'=Almost always.
Domains	The dimensions of temperament assessed in infants are; approach/sociability, irritability, and cooperation. In toddlerhood and early childhood these are; approach/sociability, reactivity, and persistence.
Respondent type	Computer-Assisted self-interview (CASI), Caregiver.
Waves	Waves 1-4.
Age range	0-11 months (infants); 12-41 months (toddlers); 42-95 months (children).
Length	Infant and Early Childhood – 12-items; Toddler – 13-items.
Publisher/ cost/ permissions	Publicly available, no cost, permission not needed.
Psychometric properties	Reliability: Cronbach's alphas from the Longitudinal Study of Australian Children (LSAC) for: -Infancy dimensions: 0.72 for approach; 0.57 for irritability, and 0.65 for cooperation (B cohort, Wave 1). (Outcome Index Working Group, 2005). - Early childhood dimensions: 0.81 for approach; 0.65 for negative reactivity, and 0.78 for persistence (K cohort, Wave 1) (Sutin et al. 2017, Forbes et al. 2017.)
Studies used/ Rationale for use	The scales were developed by the Australian Temperament Project (ATP) (Prior et al. 2000) and adapted by LSAC (Outcome Index Working Group, 2005). The scales assess the 3 core dimensions of temperament.

Scoring information	<u>For all scales:</u> Response categories: Almost never = '1' Not often = '2' Variable usually does not = 3' Variable usually does = '4' Frequently = '5' Almost always = '6'. <u>Infancy items:</u> This baby is pleasant (smiles, laughs) when first arriving in unfamiliar places. This baby stays still during procedures like hair brushing or nail cutting. This baby makes happy sounds (coos, smiles, laughs) when being changed or dressed. This baby is fretful on waking up and/or going to sleep (frowns, cries). This baby's first reaction (at home) to approach by strangers is acceptance. This baby accepts regular procedures (hair brushing, face washing etc) at any time without protest. ® This baby amuses self for ½ hour or more in cot or playpen (looking at mobile, playing with toy etc). This baby accepts within a few minutes a change in place of bath or person giving the bath. This baby can be distracted from fretting or squirming during a procedure (nail cutting, hair brushing etc) by a game, singing, TV etc. This baby continues to cry in spite of several minutes of soothing. This baby's first reaction to seeing a doctor or infant welfare sister is acceptance (smiles, coos). This baby cries when left to play alone. ® = reversed. Items 1, 5, 8 and 11 comprise the Approach scale; 2, 3, 6 and 9 the Cooperativeness scale; and 4, 7, 10 and 12 the Irritability scale. Scoring: Item 7 is first reversed (®), and then the mean of the items in each scale is computed, yielding 3 composite scores. High scores reflect high

	sociability, irritability and cooperation. No more than 1 item is permitted to be missing on each temperament dimension. If more than 1 is missing, the child is scored as 'missing' on this variable. Toddler items: (adapted by LSAC) Is pleasant (smiles, laughs) when first arriving in unfamiliar places. Plays continuously for more than 10 minutes at a time with a favourite toy. Responds to frustration intensely (e.g. screams or yells). Smiles when an unfamiliar adult plays with [him/her]. Goes back to the same activity after a brief interruption (for example a snack or trip to the toilet). Has moody 'off' days when [he/she] is irritable all day. Is outgoing with adult strangers outside the home. Stays with a routine task (dressing, picking up toys) for 5 minutes or more. Shows much bodily movement (stomps, writhes, swings arms) when upset or crying. ® Is still wary of strangers after 15 minutes. Stops to examine objects thoroughly (e.g. for 5 minutes or more). Reacts strongly (for example cries or screams) when unable to complete a play activity. Practices a new skill (for example throwing, building, or drawing) for 10 or more minutes. ® = recoded. Items 1, 4, 7 and 10 comprise the Approach scale; 2, 5, 8, 11 and 13 the Persistence scale; and 3, 6, 9, and 12 the Reactivity scale. Scoring: Item-10 is first reversed (®), and then the mean of the items in each scale is computed, yielding 3 composite scores. High scores reflect high sociability, reactivity and persistence. No more than 1 item is permitted to be missing on each temperament dimension. If more than 1 is missing, the child is scored as 'missing' on this variable. Early Childhood items: ® This child is shy with strange adults.

	When this child starts a project such as a puzzle or model, he/she works on it without stopping until it is completed, even if it takes a long time. ® If this child wants a toy or a sweet while shopping, he/she will easily accept something else instead. ® This child is shy when first meeting new children. This child likes to complete 1 task or activity before going onto the next. When this child is angry about something, it is difficult to side-track him/her. When in a park or visiting, this child will go up to strange children and join in their play. This child stays with an activity (e.g. puzzle, construction kit, reading) for a long time. When shopping together, if I do not buy what this child wants (e.g. sweets, clothing) he/she cries and yells. When unknown adults visit our home, this child is immediately friendly and approaches them. If this child is upset, it is hard to comfort him/her. ® When a toy or game is difficult, this child quickly turns to another activity. ® = recoded. Items 1, 4, 7 and 10 comprise the Approach scale; 2, 5, 8, and 12 the Persistence scale; and 3, 6, 9, and 11 the Reactivity scale. Scoring: Items 1, 3, 4 and 12 are first reversed (®), and then the mean of the items in each scale is computed, yielding 3 composite scores. High scores reflect high sociability, reactivity and persistence. No more than 1 item is permitted to be missing on each temperament dimension. If more than 1 is missing, the child is scored as 'missing' on this variable.
References	Outcome Index Working Group, The Longitudinal Study of Australian Children. (2005*). Summarising children's wellbeing: the LSAC* Outcome Index (LSAC Technical Paper #2). Melbourne: Australian Institute of Family Studies. Prior, M., Sanson, A., Smart, D. & Oberklaid, F. (2000). Pathways from infancy to adolescence: Australian Temperament Project 1983-2000. Melbourne: Australian Institute of Family Studies.

	Australian Institute of Family Studies. (2020). Longitudinal Study of Australian Children– (Accessed 06/05/2020). Melbourne: Australian Institute of Family Studies. Australian Institute of Family Studies. (2020). Data dictionary Release 7.0, Waves 1-7 – (Accessed 06/05/2020). Melbourne: Australian Institute of Family Studies. Sutin, A. R., Kerr, J. A., & Terracciano, A. (2017). Temperament and body weight from ages 4 to 15 years. International journal of obesity (2005), 41(7), 1056–1061. (Accessed 01/09/2020) Forbes, M. K., Rapee, R. M., Camberis, A. L., & McMahon, C. A. (2017). Unique Associations between Childhood Temperament Characteristics and Subsequent Psychopathology Symptom Trajectories from Childhood to Early Adolescence. Journal of abnormal child psychology, 45(6), 1221–1233. (Accessed 01/09/2020)

School Aged Temperament Inventory (SATI) – short form

Summary	The short form of School Aged Temperament Inventory (SATI) was developed for the Longitudinal Study of Australian Children (LSAC), and assesses three dimensions of temperament in children aged 8-17 years: negative reactivity, task persistence, and approach/withdrawal. The short form contains 12-items which are scored on a 5-point scale. Following reverse coding of certain items, the mean of the items in each scale is computed to yield 3 composite scores. Higher scores reflect higher negative reactivity, higher persistence and a higher degree of introversion.
Domain	The original School Aged Temperament Inventory assessed four temperament dimensions of school-aged children (McClowry, 1995): -Negative reactivity – 'the intensity and frequency with which the child expresses negative affect' -Approach/withdrawal – 'the child's initial response to new people and situations' -Task persistence – 'the degree of self-direction the child exhibits in fulfilling tasks and other responsibilities' -Activity – 'the amount of large motor activity of the child'.
Respondent type	CAPI, Caregiver.

Waves	Waves 1-4.
Age range	8-17 years.
Length	Original scale had 38-items; 12-items used in the POCLS.
Publisher/ cost/ permission	Publicly available, no cost, permission not needed. LSAC has obtained permission to use an abbreviated form of the scale.
Psychometric properties	Reliability: -Internal consistency: Cronbach's alphas ranged from 0.85 to 0.90. -Test-retest reliability from maternal data after 4-6 months was 0.80 to 0.89 (McClowry, 1995).The original instrument was validated with a sample of 435 predominantly middle-class Caucasian mothers and 228 of their spouses whose children were aged 8-11 years (McClowry, 1995). McClowry and colleagues (2003) have shown that the SATI is also suitable for older children/young people. Some slight wording changes are needed to make it suitable for 14-17 year olds. The short form included in the POCLS has been used in LSAC from Wave 3 onwards.
Studies used/ Rationale for use	The SATI is a relatively brief, psychometrically sound scale. The short version used in the POCLS was developed for LSAC. Exploratory factor analysis of ATP data guided the selection of items from the full SATI.
Scale details/ Scoring information	<u>SATI short form items:</u> Negative reactivity: Reacts strongly to a disappointment or failure (complains loudly or cries). When angry, yells or snaps at others. Moody when corrected for misbehaviour. Responds intensely to disapproval (shouts, cries etc). Task Persistence: Does not complete homework unless reminders are given ®. Remembers to do homework without being reminded. Goes back to the task at hand (chores, housework etc) after an interruption. Has difficulty completing assignments (homework, chores etc). ® Approach/Withdrawal:

	Approaches children his/her own age even when he/she doesn't know them (For children aged 14+ years; Approaches people his/her age even *when he/she doesn't know them)*. Is shy with adults he/she doesn't know ®. Is shy when meeting new children ® (For children aged 14+ years; Is shy when meeting new people of his/her age). Seems uncomfortable when at someone's house for the first time. ® = reverse coded. Response categories: Never = '1' Rarely= '2' Half of the time = '3' Frequently = '4' Always = '5'. Scoring: Following the recoding of the specified items, the mean of the items in each scale is computed. This yields 3 composite scores. No more than 1 item is permitted to be missing on each scale. If 2 or more items are missing in a scale, the child is scored as 'missing' on that dimension. Interpretation: High scores reflect higher negative reactivity, higher persistence and a higher degree of introversion.
References	Australian Institute of Family Studies. (2015). Wave 1-6 Integrated Rational document. Australian Institute of Family Studies. Available from: https://growingupinaustralia.gov.au/data-and-documentation/rationale-documents/downloads (accessed 19 June 2020). McClowry, S. G. (1995). The development of the School-Age Temperament Inventory. Merrill-Palmer Quarterly, 41, 271- 285. McClowry, S. G., Halverson, C. F., & Sanson, A. (2003). A re-examination of the validity and reliability of the School-Age Temperament Inventory. Nursing Research, 52(3), 176-182.

Brief Infant Toddler Social Emotional Assessment (BITSEA)

Summary	The Brief Infant-Toddler Social and Emotional Assessment (BITSEA is a screening tool developed by Carter, Briggs-Gowan, Jones, and Little

	(2003) to identify children that may require further professional assessment to diagnose clinically significant social/emotional and behavioural problems and/or delays/deficits in social-emotional competence (Baxter, & Konold, 2006). The Scale consists of 2 domains and yields a Problem and a Competence Total score. The summed raw scores are compared to cut-off scores by age and sex. Percentile ranks corresponding to the 25th percentile ranking or less indicates possible socio-emotional or behaviour problems and 15th percentile ranking or less indicates possible deficit/delay in competence. This suggests further professional assessment is required to determine if clinically significant.
Domains	Social-emotional/behavioural problems (31-items) and social-emotional competencies (11-items).
Respondent type	CASI, Caregiver.
Waves	Wave 1 only – replaced by CBCL in Waves 2-4.
Age range	12-35 months.
Length	42-items. 5-7 minutes for self-completion, 7-10 minutes structured interview.
Publisher/ cost/ permission	Permission granted by publisher for use in the POCLS.
Psychometric properties	Validity: -Criterion-related validity: BITSEA Problem and Child Behaviour Checklist (CBCL Ages 1.5-5) Internalising, Externalising and Total Scale: 0.64, 0.63 and 0.71 (Briggs–Gowan et al. 2004). -BITSEA Competence and Child Behaviour Checklist (CBCL Ages 1.5-5) Internalising, Externalising and Total Scale: -0.23, -0.31 and -0.30. (Briggs–Gowan et al. 2004). -Relative to the CBCL and ITSEA; Sensitivity of Problem Scale: 78 - 93%; Specificity of Problem Scale: 78– 89% (Briggs-Gowan et al. 2004). -Relative to the ITSEA; Sensitivity of Competence Scale: 69%; Specificity of Competence Scale: 78– 95% (Briggs-Gowan et al. 2004). Reliability: -Internal consistency: 0.65 for Competence and 0.79 for Problem Scale. -Inter-Rater (mother/father): ICC 0.61 for Competence and 0.68 for Problem Scale.

	-Test-retest (10 to 45 days): ICC 0.85 for Competence and 0.87 for Problem Scale (Briggs-Gowan et al. 2004).
Studies used/ Rationale for use	For Wave 1, the rationale was that this is one of the only scales to identify possible socio-emotional/behavioural problems in children under the age of 18 months and it has been used in LSAC (short form used; 34-items only). For Wave 2 onwards, the decision to use CBCL (Achenbach & Rescorla, 2000) for all ages was made as it is a more comprehensive measure, likely to be more relevant for children in OOHC, and the majority of the sample would be over 18 months and therefore the CBCL covers the entire age range (see 1.2.4 Child Behaviour Checklist).
Scoring information	For each item, a score is given and then scores are summed to yield a problem and competence score. Response categories: Not true/Rarely = '0' Somewhat True/Sometimes = '1' Very True/Often = '2'. Problem Total score = Sum of items 2 +3 +4 + 6 + 7 + 8 + 9 + 11 +12 + 14 +16 + 17 + 18 + 21 + 23 + 24 + 26 + 27 + 28 + 30 + 32 + 33 + 34 + 35 + 36 + 37 + 38 + 39 + 40 + 41 + 42. Competence Total score = Sum of items 1 + 5 + 10 + 13 + 15 + 19 + 20 + 22 + 25 + 29 + 31. Missing items: If 5 or more problem items are unanswered, the Problem Total Score is not calculated. If 2 are more competence items are unanswered, the Competence Total score is not calculated. There are two items that respondents may circle N to indicate the child has no contact with other children (no opportunity response). These are considered unanswered. Adjustment for prematurity is suggested for children up to 24 months (who were born less than 36 weeks gestation). To determine possible problem and possible deficit/delay status, as well as percentile rankings for age and gender, use figures in the manual for this measure (Briggs-Gowan & Carter, 2006). Variables have been created in the POCLS dataset using the cut-offs and percentile ranks to identify the children that may have a possible socio-emotional problem or delay/deficits.
Reference	Achenbach, T.M, & Rescorla, L. A. (2000). Manual for the ASEBA school-age forms & profiles: an integrated system of multi-informant

	assessment. Burlington, VT: University of Vermont, Research Center for Children, Youth & Families. Briggs-Gowan, M. J., & Carter, A.S. (2006). Brief Infant-Toddler Social and Emotional As*sessment: Examiner's Manual*. Brookes Publishing: Yale University and the University of Massachusetts. Briggs-Gowan, M.J., Carter, A.S., Irwin, J.R., Wachtel, K, Cicchetti, D.V. (2004). The Brief Infant-Toddler Social and Emotional Assessment: Screening for social-emotional problems and delays in competence. Journal of Pediatric Psychology, 29 (2), 143-155. Baxter, A. & Konold, T. R. (2006). Review of the ITSEA-BITSEA Infant-Toddler and Brief Infant Toddler Social and Emotional Assessment in Mental Measurements Yearbook, 17. San Antonio, Texas: Harcourt. Carter, A.S., Briggs-Gowan, M.J., Jones,S.M. and Little T.D. (2003). The Infant Toddler Social Emotional Assessment (ITSEA): Factor Structure, reliability, and validity. Journal of Abnormal Child Psychology, 31, 495-514

Child Behaviour Checklist (CBCL)

Summary	The Child Behaviour Checklist (CBCL) is a questionnaire used to assess behavioural and emotional problems in children and adolescents It was completed by caregivers of children aged 3-17 years. Versions validated and normed for use for 18 months to 5 years of age (CBCL/1.5-5) and 6-18 years of age were used (CBCL/6-18). The CBCL can be self-administered and yields subscale scores for a range of conditions and competencies. The principal focus for the POCLS are the two composite syndrome profiles, Internalising and Externalising, and the Total Problems Score.
Domains	The CBCL measures a range of childhood/adolescent internalising and externalising behaviour problems and interpersonal competencies.
Respondent type	CASI, Caregiver.
Waves	Wave 1 (age 3 years and older). Waves 2-4 (all ages).
Age range	1.5-18 years. There are 2 versions: CBCL for Ages 1.5-5 years (Achenbach & Rescorla, 2000); and CBCL for Ages 6-18 years (Achenbach & Rescorla, 2001).

Length	CBCL for Ages 1.5-5 years: 100-items, estimated to take 10-12 minutes to complete. CBCL for Ages 6-18 years: 120-items (with a few items containing additional parts); estimated to take 15-20minutes to complete.
Publisher/ cost/ permission	Permission granted by publisher for use in the POCLS.
Psychometric properties	<u>CBCL Ages 1.5-5 years:</u> Validity: -Content Validity: all but 1 item (item-61 'Refuses to eat') discriminated significantly ($p<.01$) between referred and non-referred children. -Criterion related validity: all scores were significantly ($p \leq .01$) higher for referred than non-referred children. -Construct validity: the CBCL scales are correlated with DSM criteria, Richman Behaviour Checklist (BCL) and the ITSEA. Reliability: -Cronbach Alphas were 0.89 for Internalising, 0.92 for Externalising, and 0.95 for Total Problems Scales. Alphas for the specific behaviour problem sub-scales range from 0.66 to 0.92 (p. 155, Appendix D, Achenbach 2001). -Test-Retest: correlations over an 8 day period range were 0.90 for Internalising, 0.87 for Externalising and 0.90 for Total Problems. The specific sub-scales range from 0.68 to 0.92 (n=68). Over 12 months, these range from 0.66 for Externalising to 0.76 Internalising and Total Problem scores, and for the specific sub-scales, from 0.53 to 0.64 (n=80). -Inter-Rater: correlations range from 0.59 for Internalising, 0.65 for Total Problems and to 0.67 for the Externalising Scales; and from 0.48 to 0.67 for the behaviour problem sub-scales (n =72). <u>CBCL Ages 6-18 years:</u> Validity: -Content validity: All the problem items scored significantly higher $p<0.01$ for referred than for demographically similar non referred children on CBCL. -Criterion-related validity: All empirically based scales were significantly ($p<0.01$) lower for non-referred than referred children.

	-Construct validity. The CBCL scales are correlated with DSM criteria, the Behavioural Assessment System for Children (BASC Scales), and the Conners scales. Reliability: -Cronbach Alphas for the 0.90 for Internalising, 0.92 for Externalising, and 0.94 for Total Problems. Alphas for the specific behaviour problem sub-scales range from 0.78 to 0.94. (n=3210). The alpha for the Total Competence scale is 0.79, and for its component sub-scales ranges from 0.63 to 0.69. -Test-Retest validity: correlations over an 8 day period were 0.91 for Internalising, 0.92 for Externalising and 0.94 for Total Problems and ranged from 0.82 to 0.92 for the specific syndrome scales. (n=73). -Inter-Rater validity: correlations were 0.72 for Internalising, 0.85 for Externalising and 0.80 for Total Problems Scales. Correlations ranged from 0.65 to 0.85 for the behaviour problem syndrome scales. For the Total Competence scale it is 0.68 and ranged from 0.57 to 0.76 for the sub-scales. (n=297).
Studies used/ Rationale for use	Widely used in community studies of children, both in Australia and internationally. The CBCL is regarded as the instrument of choice for measuring children's behavioural problems. Its length allows it to measure common and rare problem behaviours, and it includes items that are expected to be particularly relevant for children in care. It will allow improvement over time to be captured.
Scale details	The CBCL/1.5-5 years contains 100-items and the CBCL/ 6-18 years contains 120-items. All items are rated on a scale from '0' = not true, '1' = somewhat or sometimes true, and '2' = very true or often true. <u>CBCL Ages 1.5-5 years:</u> Syndrome scales: Emotionally Reactive e.g. disturbed by change, moody, whining. Anxious/Depressed e.g. clings, unhappy, fearful). Somatic Complaints e.g. constipated, headaches, nausea. Withdrawn e.g. avoids eye contact, doesn't answer, little interest. Sleep Problems e.g. trouble sleeping, resists bed, wakes often. Attention Problems e.g. can't concentrate, can't sit still. Aggressive Behaviour e.g. disobedient, hits others, angry moods. Other Problems e.g. jealous, accident prone, picks skin.

	Syndrome Groupings: Internalising includes Emotionally Reactive; Anxious/Depressed; Somatic Complaints; Withdrawn. Externalising includes Attention Problems and Aggressive Behaviour. Total Problems score is computed by summing scores of Internalising, Externalising, Sleep problems and the other problems that are not part of any syndromes. DSM-Oriented Scales include: Affective Problems, Anxiety Problems, Pervasive Developmental Problems, Attention Deficit Hyperactivity Problems, and Oppositional Defiant Problems. <u>For CBCL Ages 6-18:</u> Syndrome scales: Anxious/Depressed e.g. cries a lot, is fearful, feels too guilty, talks or thinks of suicide. Withdrawn/Depressed e.g. would rather be alone, shy/timid, sad. Somatic Complaints e.g. feels dizzy, constipated, headaches, nausea. Social Problems e.g. too dependent, lonely, gets teased, prefers to be with younger kids. Thought Problems e.g. cannot get mind off certain things, sees things, stores things, strange behaviour. Attention Problems e.g. cannot concentrate, daydreams, impulsive, cannot sit still. Rule breaking Behaviour e.g. drinks alcohol, lacks guilt, breaks rules, sets fires, prefers to be with older kids. Aggressive Behaviour e.g. argues a lot, destroys others' things, gets in fights, mood changes, attacks people. Syndrome Groupings: Internalising Problems includes Anxious/Depressed, Withdrawn Depressed and Somatic Complaints Scales). Externalizing Problems includes Rule breaking and Aggressive Behaviour Scales). Total Problems Score is the sum of the 1 and 2 responses on specific items of the CBCL. DSM-Oriented Scales include:

	Affective Problems, Anxiety Problems, Somatic Problems, Attention Deficit/Hyperactivity Problems, Oppositional Defiant Problems and Conduct Problems. Competence scale: A composite of the Activities, Social, and School scales, with a total of 20-items used. <u>POCLS dataset:</u> Variables have been created in the POCLS dataset to identify the children that are in the clinical, borderline and non-clinical range.
References	Achenbach, T.M., & Rescorla, L.A. (2000). Manual for the ASEBA Preschool forms and Profiles. Burlington, VT: University of Vermont Department of Psychiatry. Achenbach, T.M. & Rescorla, L.A. (2001). Manual for the ASEBA School-Age Forms & Profiles. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families

School Problems Scale

Summary	The School Problem Scale is a 4-item short form version of the Prior and colleagues (2000) school problems scale, completed by children and young people. Children and young people were asked how often do you: 'Find someone to have lunch with'; 'Understand the work in class'; 'Follow school rules and routines'; and 'Get assignments, projects and homework done' (extra question for 12-17 year olds). Response categories ranged from '1' = Always to '5' = Never. Lower mean item scores indicate fewer problems at school.
Domains	Social-emotional development – School problems.
Respondent type	ACASI, Youth self-complete.
Waves	Waves 1-4.
Age range	7-17 years.
Length	Originally 8-items; 4-items used in the POCLS.
Publisher/ cost/ permission	Publicly available, no cost, permission not needed.

Psychometric properties	The scale is relatively short, was developed for the Australian context, and is psychometrically sound. It has been found to be predictive of problem outcomes.
Studies used/ Rationale for use	Derived from the academic and social progress at school scale that the Australian Temperament Project (ATP) devised for the 12-13 and 13-14 year data collection. The academically focused items cohere together well to form a school problem scale with a good Cronbach's alpha.
Scoring information	The following 4-item short form is used in the POCLS. At school, how often do you (if not at school, answer for the last school you attended): Find someone to have lunch with? Understand the work in class? Follow school rules and routines? Get assignments, projects and homework done? (12-17-year olds only) Response categories: Always = '1' Often = '2' Sometimes = '3' Rarely = '4' Never = '5'. Scoring: The mean of the items is computed. No more than 2-items are permitted to be missing. If more than 2 are missing, the child is scored as 'missing' on this variable.
References	Prior, M., Sanson, A., Smart, D., & Oberklaid, F. (2000). Pathways from infancy to adolescence: Australian Temperament Project 1983-2000. Melbourne, Australia: Australian Institute of Family Studies.

School Bonding Scale

Summary	The School Bonding Scale is a 4-item scale completed by children and young people to measure the extent to which they are settling in and forming relationships at school. Each item: 'Try hard'; 'Get on well with your teachers'; 'Feel it is important to do well at school' (extra question for 12-17 year olds); 'Enjoy being there' is scored on a 5-point scale from Always = '1' to Never = '5'. Lower mean scores across the items

	indicate better school bonding. The items were adapted from the Australian Temperament Project (ATP).
Domains	Social-emotional development – Engagement with school.
Respondent type	ACASI, Youth self-complete (or CAPI if preferred by child).
Waves	Waves 1-4.
Age range	7-17 years.
Length	Originally 10-items; 4-items used in the POCLS.
Publisher/ cost/ permission	Publicly available, no cost, permission not needed.
Psychometric properties	Validity: The scale is predictive of both positive and problematic outcomes.
Studies used/ Rationale for use	The School Bonding Scale was originally developed by Simons, Whitbeck, Conger & Conger (1991) and used in the Seattle Social Development Project (Hawkins et al. 1992). A short version of the scale was used in the ATP. The scale is brief and has demonstrated associations with positive and problematic outcomes.
Scoring information	The 10-items from the original scale are: I try hard in school. It is important to me to get good grades. I try to do things that will make my teachers proud of me. I work hard to be successful in school. I like my maths teacher. I like my language arts/English teacher. I like my Social Science teacher. I like school. Most mornings I look forward to going to school. I like my classes. The ATP version: This version of the scale slightly reworded some items from the original scale, and combined items 5-7 into 'I am getting on well with my teachers'. Items 1 and 4 were seen as highly overlapping and only item 4 was used in the ATP. The ATP scale was thus 7-items long.

	<u>For the POCLS the following short version is used:</u> At school, how often do you (if not at school, answer for the last school you attended): Try hard. Get on well with your teachers. Feel it is important to do well at school (ages 12-17 only). Enjoy being there. Response categories: Always = '1' Often = '2' Sometimes = '3' Rarely = '4' Never = '5' Scoring: The mean of the items is computed. No more than 2 items are permitted to be missing. If more than 2 are missing, the child is scored as 'missing' on this variable.
Additional information	N/A
References	Simons, R.L., Whitbeck, L.B., Conger, R.D., Conger, K.J. (1991). Parenting factors, social skills, and value commitments as precursors to school failure, involvement with deviant peers, and delinquent behaviour. Journal of Youth Adolescence, 20 (6), 645-664 Hawkins, J.D., Catalano, R.F., Morrison, D.M., O'Donnell, J., Abbott, R.D., & Day, L.E. (1992). The Seattle Social Development Project: Effects of the first four years on protective factors and problem behaviors. In J. McCord & R. Tremblay (Eds.), Preventing antisocial behavior: Interventions from birth through adolescence (pp.139-161). New York: Guilford Press.

Short Mood and Feelings Questionnaire (SMFQ)

Summary	The Short Mood and Feeling Questionnaire (SMFQ) is a brief measure of depression that was self-completed by young people aged 12-17 years (Angold, Costello, Messer, Pickles, Winder & Silver, 1995). The questionnaire contains 13-items that describe mood or feelings which

	may have been experienced in the past 2 weeks. The response categories are: Not true = '0'; Sometimes = '1'; True = '2'. Responses are summed and cut-off scores can be used to identify young people with significant depression.
Domains	Social-emotional development – Depression.
Respondent type Waves	Youth self-complete. Waves 2-4.
Age range	12–17 years.
Length	13-items.
Studies used	US and UK studies (Sharp et al. 2008); ATP and LSAC.
Rationale for use	The SMFQ is a relatively brief measure of depression with sound psychometric properties. It has been used in other Australian longitudinal studies such as LSAC and the ATP, as well as internationally. The authors provide cut-offs by which to identify children experiencing significant levels of depression.
Publisher/ cost/ permission	No Cost. Publicly available, permission not needed.
Psychometric properties	Validity: -Sensitivity ranges from 0.59 to 0.75 across several studies. -Specificity ranges from 0.74 to 0.85 across several studies. -Content validity established by convergence with diagnosis on DISC depression scale/s (correlation of 0.65) and CDI (correlation of 0.67). Reliability: -Cronbach's alpha = 0.85
Scoring information	The scale consists of the following 13-items: Please think about how much you have felt or acted in any of these ways in the past 2 WEEKS. For each one, if this was true about you most of the time, choose true. If it was only sometimes true, choose sometimes. If it was not true about you, choose not true. I felt miserable or unhappy. I didn't enjoy anything at all. I felt so tired I just sat around and did nothing. I was very restless.

	I felt I was no good anymore. I cried a lot. I found it hard to think properly or concentrate. I hated myself. I felt lonely. I felt I was a bad person. I thought nobody really loved me. I thought I could never be as good as other kids. I did everything wrong. Response categories: Not true = '0' Sometimes = '1' True = '2'. Scoring: Responses are summed. Angold and colleagues (1995) suggest that a score of 8 or more achieves a sensitivity of 60% and 85% specificity. This can be used as a cut-off to identify young people with significant depression problems.
Additional information	N/A
References	Angold, A., Costello, E. J., Messer, S. C., Pickles, A., Winder, F., & Silver, D. (1995). Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. International Journal of Methods in Psychiatric Research, 5, 237-249. Sharp, C., Goodyer, I. M., & Croudace, T. J. (2006). The Short Mood and Feelings Questionnaire (SMFQ): A unidimensional item response theory and categorical data factor analysis of self-report ratings from a community sample of 7-through 11-year-old children. Journal of Abnormal Child Psychology, 34, 379–391. Thapar, A., & McGuffin, P. (1998). Validity of the shortened Mood and Feelings Questionnaire in a community sample of children and adolescents: A preliminary research note. Psychiatry Research, 81, 259–268.

Self Report Delinquency Scale (SRDS)

Summary	A short form of the Self Report Delinquency Scale (SRDS) comprising 4-items was used to assess self-reported levels of anti-social behaviour in young people aged 10-11 years and 10-items was used to assess self-reported levels of anti-social behaviour in young people aged 12-17 years (Moffitt & Silva, 1988). Items 1-11 are scored on a scale from Not at All = '1' to More Than 20 Times = '9', while items 12 and 13 are scored by whether the respondent used alcohol and/or an illicit drug in the last 12 months ('Yes = '1' and No = '2'). All items are recoded to yield a final score of either 0 = no/low antisocial behaviour or 1 = highly antisocial behaviour.
Domains	Socio-emotional development – Antisocial behaviour.
Respondent Type Waves	Youth self-complete. Waves 2-4.
Age range	10-17 years.
Length	The original Self Report of Delinquency Scale comprised 58-items. A short form comprising 13-items is used in the POCLS.
Publisher/ cost/ permission	No cost. Publicly available scale.
Psychometric properties	Validity: Convergence was assessed between the SRDS and parent ratings of conduct disorder and socialised aggression (combined to provide a single score) from the Revised Behaviour Problem Checklist (Quay & Peterson, 1983), and teacher ratings on the Rutter Child Scale B antisocial behaviour scale (Rutter, Tizard and Whitmore, 1970). Correlations were 0.45 with parent report and 0.28 with teacher report for the full SRDS, and 0.46 with parent report and 0.27 with teacher report for the Illegal Scale. The short form used in the ATP yields rates of antisocial behaviour that are similar to other studies using lengthier instruments and to other cross-sectional Australian studies. Reliability: -Kuder-Richardson-20 coefficient for full SRDS = 0.90, and for the 29-item Illegal Scale was 0.81. -Test-retest reliability for the original SRDS (over 1 month) = 0.85 (n = 20).

	-Cronbach alpha for a similar short scale used in the ATP = 0.82 (at 15-16 years, n = approx 1,600).
Studies used	The scale was originally developed for use in the Dunedin Multidisciplinary Health and Development Study. A short form has been successfully used in the ATP; and LSAC is using a short form in Wave 5. Thus, the possibility of benchmarking to these community studies exists.
Rationale for use	The SRDS was developed for use with an adolescent population of New Zealand youth, hence has high comparability to Australian youth. The instrument has strong validity, as shown by Moffitt's ground-breaking theoretical work on adolescent antisocial behaviour which is based on the SRDS. A short form has been used in other longitudinal studies such as LSAC and the ATP and has performed well.
Scoring information	The original Self Report of Delinquency Scale contained the following items: 1. Running away from home and staying away overnight. 2. Carrying a weapon in case it is needed in a fight like a knife, chain, piece of wood. 3. Going around with a group of 3 or more damaging property or getting into fights. 4. Damaging something in a public place (e.g. a street, movie theatre, bus, toilet). 5. Purposely damaging or destroying something belonging to your parents. 6. Starting a fire where you should not burn anything. 7. Damaging a parked car (e.g. breaking an aerial, slashing tyres, scratching paint). 8. Raising a false alarm (e.g. dialling 000 or setting off a false fire alarm. 9. Stealing a thing or money worth between $2 and $40. 10. Stealing a thing or money worth more than $40. 11. Breaking into a house, flat or vehicle to try to steal something or just look around. 12. Stealing something from a shop or store (shoplifting). 13. Stealing something out of a parked car. 14. Sealing goods or money from a video machine, public telephone or vending machine. 15. Stealing a bicycle. 16. Taking a car or motorcycle for a drive without permission. 17. Sniffing glue, petrol or other things in order to feel 'high'. 18. Smoking cannabis (pot marijuana, hashish). 19. Using any illegal drugs other than cannabis (heroin, cocaine, speed).

	20. Buying or drinking alcoholic drinks. 21. Drinking alcoholic drinks during school hours or at lunchtime on a school day. 22. Getting suspended or expelled from school. 23. Playing truant from school (skipping school). 24. Hitting one of your parents. 25. Fighting in the street or other place (not at school). 26. Struggling to get away from a policeman. 27. Using force or threats to get money from someone about your age or younger. 28. Using force or threats to get money from someone older than yourself. 29. Using any kind of weapon in a fight (e.g. knife, chain, broken bottle, rock). 30. Cheating in school tests or exams. 31. Going to R rated films without a parent's permission. 32. Driving a car, motorcycle or motor scooter on a public road without a licence. 33. Trespassing anywhere you are not supposed to go (like railway yards, private property, empty houses or building sites). 34. Spending $2 or more on gambling. 35. Making rude telephone calls e.g. ringing someone and saying dirty or threatening things. 36. Swearing loudly in a public place. 37. Purposely littering the street or footpath by smashing bottles or tipping rubbish bins. 38. Purposely damaging property that belongs to a school. 39. Painting or writing graffiti words on a wall in a public place. 40. Doing damage in a park or public garden. 41. Breaking the windows of an empty building. 42. Letting off a fire extinguisher where there is no fire. 43. Placing something that may damage a train on railway tracks. 44. Moving or damaging a traffic sign or road works equipment. 45. Letting down tyres of a car, truck or motorcycle. 46. Stealing a thing or money worth less than $2. 47. Taking money from home without permission. 48. Stealing something from another pupil at school. 49. Stealing school property (e.g. books, sports equipment). 50. Travelling on a bus or going to the movies without paying. 51. Taking a badge or hubcap form a car or truck. 52. Stealing money from milk bottles. 53. Getting drunk on alcoholic drinks.

	54. Stealing beer, wine or spirits from a shop, your parents' home or other place. 55. Getting into trouble with your friends, family or school because you were drinking alcohol. 56. Throwing rocks of objects e.g. rocks or bottles at people or moving cars. 57. Hitting another young person in a serious effort to hurt them (not a brother or sister). 58. Being cruel to an animal so as to injure the animal. The first 29-items form the Illegal Scale and the next-29 comprise the Norm Violation Scale. The 29 Illegal Scale items were selected on police advice that they were: a) illegal for youth under 17 years, and b) the act was viewed as worthy of police intervention. These items also generally constitute more serious acts. In selecting items from the Illegal Scale, items 17-21 were first excluded, as this data is obtainable from the substance use items included in the POCLS. Next, items covering similar content were combined (e.g. 9 & 10; 27 & 28) or the item that received the most numerous positive responses was selected. Additionally, to ensure that the four main categories of antisocial behaviour were covered (authority conflict, violence, vandalism/property damage, and drug use), a small number of items were selected from the Norm Violation Scale. The items used for 10-17 year olds in the POCLS are: Here is a list of things that some people your age have done. In the last 12 MONTHS how often have you done each of the following things... Jigged or wagged school for a whole day? [10-17 YEARS] Got into a fight with someone such as hit or punched them? [10-17 YEARS] Stolen something from a shop? [10-17 YEARS] Tagged or graffitied in public places [ONLY ASK 12 YEARS AND OLDER] Carried a weapon like a knife [ONLY ASK 12 YEARS AND OLDER] Taken a car or motorbike for a drive without permission [ONLY ASK 12 YEARS AND OLDER] Stolen money or other things from someone [ONLY ASK 12 YEARS AND OLDER] Damaged a parked car such as broke an aerial or scratched the paint [ONLY ASK 12 YEARS AND OLDER] Broken into a house, flat or vehicle [ONLY ASK 12 YEARS AND OLDER]

	Broken or damaged something on purpose [10-11 YEARS] Used force or threats to get money or things from someone [ONLY ASK 12 YEARS AND OLDER]. Response categories: Not at all / 1 time / 2 times / 3 times / 4 times / 5 times / 6-10 times / 11-20 times / More than 20 times / Don't know / Pass Have you had an alcoholic drink in the last 12 months? Have you had drugs in the last 12 months? Response categories: Yes = '1'/ No = '2'. Children exhibiting high levels of antisocial behaviours can be identified via the criterion of engagement in 3 or more differing types of antisocial activities during the previous 12 months. This criterion parallels the DSM IV criteria for conduct disorder and is widely used. Scoring: Responses to items 1-11 are first recoded as 1 = 0, and 2-9 = 1. Responses to items 12-13 are recoded as 1 = 1, and 2 = 0. Responses for these new, dichotomous 13 items are then summed to yield a total score (range 0 – 13). Finally, scores of 0-2 are recoded as 0 = no/low antisocial, and 3-13 as 1 = highly antisocial.
Additional information	N/A.
References	Moffit, T.E., & Silva, P.A. (1988). Self-reported delinquency: Results from an instrument for New Zealand. Australian and New Zealand Journal of Criminology, 21, 227-240. Smart, D., Vassallo, S., Sanson A., & Dussuyer, I. (2004). Patterns of antisocial behaviour from early to late adolescence (Trends and Issues in Crime and Criminal Justice No. 290). Canberra, ACT: Australian Institute of Criminology. Vassallo, S., Smart, D., Sanson, A., Dussuyer, I., McKendry, B., Toumbourou, J., Prior, M., & Oberklaid, F. (2002). Patterns and precursors of adolescent antisocial behaviour. Report commissioned by Crime Prevention Victoria.

Emotional Responsiveness Scale from the Parenting Style Inventory II (PSI-II)

Summary	The Emotional Responsiveness Scale from the Parenting Style Inventory (PSI-II) (Darling & Toyokawa, 1977) is used a measure of children and young people's relationships with their caregivers. The POCLS adapted this scale from LSAC and is comprised of 5-items. Each item, which asks children how often their caregiver: 'Help you out if you have a problem', 'Listen to you', 'Praise you for doing well', 'Do things with you that are just for fun', and 'Spend time talking to you' is scored on a 5-point scale.
Domains	Parenting practices/style.
Respondent type	ACASI Youth self-complete (or CAPI if preferred by child).
Waves	Waves 1-4.
Age range	7-17 years.
Length	5-items.
Publisher/ cost/ permission	The original scale (Darling & Toyokawa, 1997) is not formally published as a measure, but has been used quite extensively. The POCLS will continue to reference this scale as it is in the public domain without any restriction to use.
Psychometric properties	Validity: The Responsiveness Scale is predictive of self-esteem, school bonding, school involvement, problem behaviours, substance use, and GPA. Reliability: -Cronbach's alpha = 0.74. (Darling & Toyokawa, 1997).
Studies used/ Rationale for use	The scale was used in LSAC. The Responsiveness Scale provides a brief, psychometrically sound measure of young people's relationships with their caregivers. Alternative scales are generally longer and more diffuse in content.
Scoring information	The version used in LSAC contains the following items. Items 1 and 2 were re-worded so that they were positively, rather than negatively, phrased. Like you to tell them [him/her] when you are worried or have a problem Praise you for doing well

	Spend time just talking with you Do things with you that are just for fun You can count on them [him/her] to help you out if you have a problem <u>For the POCLS, the items have been simplified and re-worded as shown below:</u> Thinking about the adults looking after you, how often do they: Help you out if you have a problem? Listen to you? Praise you for doing well? Do things with you that are just for fun? Spend time just talking with you? Response categories: Always = '1' Often = '2' Sometimes = '3' Rarely = '4' Never = '5' Scoring: All items are reversely scored. The sum of the items is computed. No more than 2 items are permitted to be missing. If more than 2 are missing, the child is scored as 'missing' on this variable.
References	Darling, N., & Toyokawa, T. (1997). Construction and Validation of the Parenting Style Inventory II (PSI-II). Unpublished manuscript, Department of Human Development and Family Studies, The Pennsylvania State University.

Adapted Kvebaek Family Sculpture Technique

Summary	The Kvebaek Family Sculpture Technique (KFST) is a symbolic form of family sculpting to show subjective relationships and is a procedure used in family assessment and research (Cromwell et al. 1980). The person "sculpting" arranges representational figures to show "the way it feels" in their family or group. The technique is simple and easy for small children or those for whom verbal expression is difficult. The KFST is based on both psychoanalytic theory and in family systems theory. The technique was adapted for the POCLS to measure the child's view of how close they feel to others. The physical distance represented on an 8x8 checkerboard between the child and family members is interpreted as a measure of experienced psychological distance (e.g., closeness, belongingness, cohesion). The POCLS activity was used with children aged 7-17 years in Waves 1-4. From Wave 2 onwards, similar questions were added to the child Audio Computer Assisted Self Interview (ACASI) for older children who opted not to complete the activity (i.e., viewed as not age appropriate).
Domains	There are 3 parts for those in placements with families: Part A: Closeness to OOHC family Part B: Closeness to birth family Part C: Three most important people in their life (Wave 1 only). For those living independently: Part A: Closeness to birth family Part B: Closeness to extended family and friends Part C: Three most important people in their life (Wave 1 only).
Respondent type	Interviewer administered to children and young people (excluding not those with birth parents) from Wave 1 onwards. Feedback from interviewers after Wave 1 was that the activity was not age-appropriate for older children. Thus, from Wave 2 onwards, older children answered related questions in the ACASI if they opted not to do the activity.
Waves	Waves 1-4.
Age range	7-17 years.
Length	Administration time is approximately 12 minutes.
Publisher/ cost/ permission	Permission granted by publisher for use an adapted version in the POCLS. The POCLS agreed to acknowledge David Kvebaek and KST Associates for developing the theoretical basis and method under

	girding the Kvebaek instrument that informed the development of the activity exploring family cohesion.
Psychometric properties	Validity: Significant negative correlations between KFST distance scores and Family Adaptability and Cohesion Scale (FACES) III mean cohesion scores (-0.74) indicate convergent validity (Berry et al. 1990).
Scoring information	See the POCLS Adapted Kvebaek Family Sculpture Technique Technical Report Number 11.
References	https://kvebaeksculpting.com/kst/ Cromwell, R. E., Fournier, D. and Kvebaek, D., (1980). The Kvebaek Family Sculpture Technique: A diagnostic and research tool in family therapy. Jonesborough TN: Pilgrimage. Berry, J., Hurley, J., Worthington E. (1990). Empirical validity of the Kvebaek Family Sculpture Technique. American Journal of Family Therapy [serial online], 18(1):19-31.

1.2. Cognitive and language development

Communication and Symbolic Behaviour Scale Infant-Toddler Checklist (CSBS ITC)

Summary	In the POCLS, caregivers of children aged 6-24 months completed the CSBS ICT in Wave 1. The scale comprises 24-items which measure language development and symbolic abilities in children in seven different cluster areas. (Wetherby & Prizant, 2003). Items are scored to yield individual cluster scores and then cluster scores are summed to yield three composite scores: social composite, speech composite, and symbolic composite. A total raw score is obtained by totalling the three composite scores. The composite scores and the total raw score is then converted to a normed score based on the chronological age of the child. Cut-off scores, standard scores and percentile ranks are used to identify potentially concerning scores.
Domains	Infant and toddler communication seven clusters: emotion and eye gaze (4-items); communication (4-items); gestures (5-items); sounds (3-items); words (2-items); understanding (2-items); and use of objects (4-items).
Respondent type	CAPI, Caregiver.

Waves	Used only in Wave 1 as children are all aged above 24 months at Wave 2.
Age range	6-23 months of age.
Length	24-items.
Publisher/ cost/ permission	Permission granted by publisher for use in the POCLS.
Psychometric properties	Validity: Concurrent validity: r = 0.59 to 0.67 (Wetherby et al. 2002) Sensitivity: 76.1% (at 2 year follow-up), 84.2% (at 3 year follow-up) Specificity: 82.2% (at 2 year follow-up), 72.6% (at 3 year follow-up) (Wetherby & Prizant, 2003). Reliability: Test-retest (4 months): r = 0.34 to 0.87 (Wetherby et al. 2002). Internal consistency: Cronbach's alpha for the clusters, composite and total scores range from 0.87 to 0.93 (Wetherby & Prizant, 2003).
Studies used/ Rationale for use	This scale was used in LSAC. There are few other measures suitable for the infant/toddler age range. The one alternative, the Preschool Language Scale (Zimmerman et al. 2002) has a lengthy administration time and significant training requirements.
Scoring Information	There are 24-items on the checklist which are scored on either a 2, 3, or 4-point scale. The points are totalled into seven cluster raw scores, and 3 composite raw scores (Wetherby & Prizant, 2001). Social composite = Emotion and eye gaze + Communication + Gestures Speech composite = Sounds + Words Symbolic composite = Understanding + Object use. Social composite, speech composite and symbolic composite raw scores are summed to give a total raw score, which are then converted to normed scores based on the chronological age of the child. Scores that are considered of concern are: standard scores at or below 6 for the composite scores; standard scores at or below 81 for the total score; and percentile ranks at or below the 10th for the composites and total scores.
Additional information	Department of Communities and Justice (former Department of Family and Community Services) Multicultural Service Unit confirmed that some items are not relevant to all cultures. For example, in relation to Q13 not all cultures nod the head to indicate 'yes' and Q15-18 and Q20 about sounds and words. If the child's first language is not English, they may

	not have developed these skills in English at the same rate as a child from an English-speaking family.
References	Wetherby, A. M., Allen, L., Cleary, J., Kublin, K., & Goldstein, H. (2002). Validity and reliability of the communication and symbolic behaviour scales developmental profile with very young children. Journal of Speech, Language, and Hearing Research, 45(6), 1202-1216. Wetherby, A. & Prizant, B. (2001). Communication and Symbolic Behavior Scales Developmental Profile Infant/Toddler Checklist. Baltimore, MD: Paul H Brookes Publishing. Wetherby, A.M., & Prizant, B.M. (2003). CSBS-DP Infant-Toddler Checklist and Easy-*Score User's Guide.* Baltimore, MD: Paul H Brookes Publishing. Zimmerman, I., Steiner, V. & Pond, R. (2002). Preschool Language Scale, Fourth Edition. San Antonio, TX: The Psychological Corporation.

MacArthur-Bates Communicative Developmental Inventories (MCDI-III)

Summary	The MacArthur-Bates Communicative Developmental Inventories were used to assess communication skills in children aged 30-35 months of age in Waves 1-2. The LSAC version of the inventory was used which includes 98 vocabulary words and 12 sentence pairs. The number of words in the vocabulary checklist that the caregiver marks are summed to give a score between 0 and 98. Similarly, each of the sentence pairs are scored '1' if caregiver marks that the child uses the more complex sentence, giving a maximum score of 12.
Domains	Assesses child communication skills.
Respondent type	CAPI, Caregiver.
Waves	Waves 1-2.
Age range	30-35 months of age.
Length	98 vocabulary words, 12 sentence pairs.
Publisher/ cost/ permission	Permission granted by publisher for use in the POCLS. This agreement recognises that we are using the LSAC version of the tool which is different to the original version.
Psychometric properties	Concurrent Validity:

	For the original version of the scale; 0.57 (with McCarthy Verbal Scale), 0.63 (with PLS-3 Total Score), 0.50 (by PPVT-R) (Fenson et al. 2007).
Studies used/ Rationale for use	Used in the LSAC.
Scoring information	The number of words in the vocabulary checklist that the caregiver marks is summed to give a total score between 0 and 98. Each of the 12 sentence pairs are scored '1' if caregiver marks that the child uses the second (more complex) sentence. There is a maximum score of 12.
Additional information	Differences between LSAC version and original version: 98-item vocabulary list in LSAC version versus 100-items in original list; LSAC version excludes a screening question that occurs before the 12 sentence pairs; LSAC does not include 12 yes/no questions. The Vocabulary Scale that was used in LSAC consisted of 98 words; 95 of the original 100 words and 3 words that were substituted for words that are commonly used in Australia ("kangaroo" for "reindeer", "biscuit" for "cracker", "footpath" for "sidewalk"). This Vocabulary Checklist yielded a total possible raw score of 98. The Grammar Scale consists of 12 sentence pairs (e.g., "Daddy, pick me up" and "Daddy picked me up") and parents are asked to select the sentence that sounds most like the way their child talks. The Grammar Scale yielded a total possible raw score of 12 (LSAC 2010 p.112-113).
References	Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., & Reznick, J. S. (2007). MacArthur-Bates communicative development inventories: *User's guide and technical manual* (2nd Ed) Baltimore, MD: Brookes. Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E., (2007). MacArthur-Bates Communicative Development Inventories (2nd ed.). Baltimore, MD: Paul H. Brookes. Australian Institute of Family Studies. (2011). The Longitudinal Study of Australian Children Annual Statistical Report 2010. Melbourne: Australian Institute of Family Studies. Zimmerman, I., Steiner, V., & Pond, R. (2002). Preschool Language Scale, Fourth Edition. San Antonio, TX: The Psychological Corporation. The Longitudinal Study of Australian Children. Annual statistical report 2010. Chapter 10: Children's Language Development.

MacArthur Communicative Development Inventories—short form

Summary	A short form of the MacArthur Communicative Development Inventories was used in Waves 1-2 with children aged 24-29 months of age to assess communication skills. There are 2 equivalent word checklists, Form A and Form B, both with 100-items. The number of words used by the child are marked by the caregiver on a Yes or No response format. No = '0' indicates caregiver have not heard the child used that word and Yes = '1' indicates that the child used the word. The total scores ranges from 0 to 100 and summated raw scores are then standardised.
Domains	Assesses communication skills.
Respondent type	CASI, Caregiver.
Waves	Waves 1-2 (all children over 29 months by Wave 3).
Age range	24-29 months (toddler version).
Length	100 words plus 1-item about word combinations (2 equivalent word checklists – Form A and Form B).
Publisher/ cost/ permission	Permission granted by publisher for use in the POCLS. Some words changed to be appropriate for the Australian context (with permission).
Psychometric properties	Toddler form correlation to long form: Toddler Form A: 0.99 Toddler Form B: 0.99 *Cronbach's Alpha:* Toddler Form A: 0.99 Toddler Form B: 0.99 Concurrent Validity: Toddler Form A: 0.74 (n=28) Toddler Form B: 0.93 (n=40) (Fenson et al. 2000).
Studies used/ Rationale for use	This instrument has strong psychometric properties and was significantly briefer in administration time than other language measures (e.g., Preschool Language Scale, Zimmerman et al. 2011).
Scoring information	Number of words the caregiver marks is summed to give a total out of 100.]. Raw scores can be converted to percentile ranks (Fenson et al. 2000).

References	Fenson, L., Pethick, S., Renda, C., Cox, J. L., Dale, P. S., & Reznick, J. S. (2000). Short-form versions of the MacArthur Communicative Development Inventories. Applied Psycholinguistics, 21, 95-116. Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E., (2007). MacArthur-Bates Communicative Development Inventories (2nd ed.). Baltimore: Paul H. Brookes. Zimmerman, I. L., Steiner, V. G., & Pond, R. A. (2011). The Preschool Language Scale-5. San Antonio, TX: Pearson.

Peabody Picture Vocabulary Test (PPVT-4)

Summary	This Peabody Picture Vocabulary Test (PPVT-4) test of verbal knowledge can be administered to children from the age of 2½ years up to adults 90 years old. There are 228-items with different starting points for children of different ages. The test yields raw scores based on correct answers and errors as well as standardised scores (M = 100, SD= 15) for different ages. Scores higher or lower than the reference point of 100 indicate the extent to which the child's vocabulary compares with peers.
Domains	Language Development. The PPVT–4 scale measures understanding of the spoken word in standard English (receptive language skills) and can assess vocabulary acquisition over time (PPVT-IV; Dunn & Dunn, 2007).
Respondent type	Interviewer administered test – Form A.
Waves	Waves 1-4.
Age range	3-17 years.
Length	228-items, but covering a wide age range, so children will generally complete a much smaller number of items. It is estimated to take 10-15 minutes to complete.
Publisher/ cost/ permission	Permission granted by publisher for use in the POCLS.
Psychometric properties	Validity: The PPVT–4 was compared with the EVT–2, CASL, CELF- 4, GRADE, and PPVT-III and showed good validity (Dunn & Dunn, 2007). Reliability:

	Internal consistency: 0.94 and 0.95 on each form. Test-retest: 0.92 and 0.96 (very high).
Studies used/ Rationale for use	LSAC uses an adapted shortened m version of the PPVT–III developed by S. Rothman (Australian Council for Educational Research [ACER], 2000)The PPVT is a widely used and highly regarded scale with strong psychometric properties (Dunn & Dunn, 2007). It is predictive of children's later academic performance. It will be an engaging task for children and assist with building rapport. It is can be administered by interviewers following training provided by a psychologist.
Scoring information	The items sample words from 20 content areas (e.g., actions, vegetables, tools) and parts of speech (nouns, verbs, or attributes) across all levels of difficulty. The 2 parallel forms are designed to minimise potential practice effects from repeated administration (Dunn & Dunn, 2007). The 228-items are divided into 19 sets of items. Administration: The training items are administered first, followed by the start item, which is the first item of the age-appropriate item set. After the child has completed 2 training items correctly without help, the start item and the remaining items in its set are presented. The child must meet the basal rule of 0 or 1 errors in the set. If the child fails to meet the basal rule, earlier sets are administered in order of difficulty (from the most to the least difficult) until the rule is met or until Set 1 is completed. Following this, test forward by sets until the ceiling effect - 8 or more errors in a set - is obtained (Dunn & Dunn, 2007). Scoring: The sum of the total number of errors across all sets completed is calculated (from the lowest basal set to the highest ceiling set). The total number of errors is then subtracted from the highest numbered item in the ceiling set. The resulting score is the raw score. The raw scores can be converted to age and grade-based standard scores (M = 100, SD = 15), percentiles, normal curve equivalents, stanines, age and grade equivalents, and Growth Scale Value (incremental vocabulary growth over time), using the norms provided.
References	Dunn, L.M. & Dunn, D. M. (2007). Peabody Picture Vocabulary Test fourth edition, Manual. Minneapolis, MN: Pearson Assessments. Australian Council for Educational Research (ACER). (2000). PPVT-III-LSAC Australian Short-form developed by S Rothman. Camberwell, Vic: ACER.

Matrix Reasoning Test from Wechsler Intelligence Scale for Children (WISC-IV)

Summary	Children aged 6-17 years completed 35 Matrix Reasoning items from the Weschler Intelligence Scale, Fourth Edition (WISC-IV) as a measure of logical reasoning or fluid intelligence. The test is administered by an interviewer. Children are shown a set of pictures with 1 missing square, and asked to choose from 5 options the picture that best fits the missing square. Administration of the test then differs according to whether a correct response is given for the first 2 items (there are age based start points). One point is scored for a correct response and zero for an incorrect response or no response. Raw scores are converted to scale scores which range from 1-19.
Domains	Cognitive development, fluid intelligence.
Respondent type	Interviewer-administered test.
Waves	Waves 1-4.
Age range	6-16 years.
Length	35-items; 5-10 minutes to administer.
Publisher/ cost/ permission	Permission granted by publisher for use in the POCLS.
Psychometric properties	Validity: Convergent and discriminant validity evidence for the WISC-IV is provided by correlational studies with the following instruments: WISC–III, WPPSI–III, WAIS–III, WASI, Wechsler Individual Achievement Test 2nd Edition, Children's Memory Scale, Gifted Rating Scale, Bar-On Emotional Quotient- Inventory, and the Adaptive Behavior Assessment System 2nd Edition (Wechsler et al. 2003). Construct validity evidence for the WISC-IV is demonstrated by a series of exploratory and confirmatory factor-analytic studies and mean comparisons using matched samples of clinical and non-clinical children.
Studies used/ Rationale for use	The Matrix Reasoning was introduced in WISC IV. This sub-test of the WISC-IV has been used in Australian and international studies, for example LSAC.

	The Matrix Reasoning test provides a brief but reliable estimate of general nonverbal intelligence. It is part of the widely used WISC-IV. It is an engaging task for children and assists with building rapport. The test can be administered by interviewers following training provided by a psychologist.
Scale details/ Scoring information	Administration: Children are shown a set of pictures with 1 missing square, and asked to choose from 5 options the picture that best fits the missing square. The Matrix Reason WISC IV subtest starts at predetermined items depending on child's age (age specific starting point), so the first 2 items administered are not the first two items on this subtest If a child does not give correct responses to both the first 2 items administered, the preceding items are administered in reverse order until the child provides correct responses on 2 consecutive items. The test is discontinued after the child gives incorrect responses to 4 consecutive items, or to 4 out of 5 consecutive items. Scoring: One point is scored for a correct response and zero for an incorrect response or no response. Scores can range from 0 to 35. If the child responds correctly to the first 2 items administered, he/she is given scores of 1 for each of the preceding items. If the child does not respond correctly to the first 2 items administered, his/her raw score is the number of items administered that were correctly answered plus scores of 1 for any preceding items that were not administered. Raw scores are converted to scale scores. These range from 1 to 19. Scale score details are provided in the WISC manual (details are not provided here as they are very lengthy) and are age adjusted, with norms provided in 3 monthly blocks (i.e. for 6.0 to 6.3 years; for 6.4 to 6.7 years and so on).
References	Flanagan, Dawn & Alfonso, Vincent & Mascolo, Jennifer & Hale, James. (2010). The Wechsler Intelligence Scale for Children - Fourth Edition in Neuropsychological Practice. Wechsler, D. (2003). WISC-IV technical and interpretive manual. San Antonio, TX: Psychological Corporation. Online: https://images.pearsonclinical.com/images/pdf/wisciv/WISCIVTechReport2.pdf Williams, P.E., Weiss, L.G. & Rolfhus, E. L. (2003). WISC-IV Technical Report #2: Psychometric Properties. San Antonio, TX: Pearson.

	As I did not have access to the manual to verify: https://www.researchgate.net/publication/265601166_The_Wechsler_Intelligence_Scale_for_Children_-_Fourth_Edition_in_Neuropsychological_Practice

2. Caregiver and placement characteristics

2.1. Caregiver psychological distress

Kessler (K10)

Summary	The 10-item Kessler (K10) measure was completed by caregivers to measure psychological distress. Completion of this scale requires participants to rate how often a series of statements applied to them in the last 30 days. The response categories spanned a 5-point scale ranging from None of the time = '1' to All of the time = '5'.
Domains	Psychological distress of caregivers.
Respondent type	ACASI, caregiver.
Waves	Waves 1-4.
Length	10-items.
Publisher/ cost/ permission	Publicly available, no cost, permission not needed.
Psychometric properties	Good precision: 90-99th percentile population range (Kessler et al. 2002). Good discrimination: Under ROC curve = 0.876 (Kessler et al. 2002).
Studies used/ Rationale for use	It is important to assess caregiver psychological distress as this may impact on the quality of care provided to the child and may be a risk factor for poor child outcomes. The K10 is brief, widely used and has strong psychometric properties, with the ability to distinguish between DSM cases and non-cases.
Scoring information	<u>K10 Test items (Kessler et al. 2002):</u> During the last 30 days: about how often did you feel tired out for no good reason? about how often did you feel nervous? about how often did you feel so nervous that nothing could calm you down? about how often did you feel hopeless? about how often did you feel restless or fidgety? about how often did you feel so restless you could not sit still? about how often did you feel depressed?

	about how often did you feel that everything was an effort? about how often did you feel so sad that nothing could cheer you up? about how often did you feel worthless? Response categories: All of the time = '1' Most of the time = '2' Some of the time = '3' A little of the time = '4' None of the time = '5' Scoring: The total score is computed by reverse coding the input items (high scores indicate higher levels of distress). If the response to the preceding question is 'none of the time', items 3, 6 and 8 are not asked but given the same value and scored as '1'. Items are summed to yield a total score between 10 and 50. If more than 1 item is missing the total score is set to missing. The categorisation of the K10 used in the POCLS is: Low distress (10-15), Moderate distress (16-21), High distress (22-29) and Very High distress (30-50).
References	Australian Bureau of Statistics. (2012). Information Paper: Use of the Kessler Psychological Distress Scale in ABS Health Surveys, Australia, 2007-08 [webpage]. Retrieved from [http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/4817.0.55.001Chapter92007-08]. Kessler, R.C., Andrews, G., Colpe, L.J., Hiripi, E., Mroczek, D.K., Normand, S.-L.T., Walters, E.E., & Zaslavsky, A. (2002). Short screening scales to monitor population prevalences and trends in nonspecific psychological distress. Psychological Medicine. 32(6), 959-976.

2.2. Social Cohesion

Social Cohesion and Trust Scale

Summary	The Social Cohesion and Trust Scale was used as a measure of neighbourhood social cohesiveness. One item of the scale was not used after it was found to be inappropriate for Aboriginal respondents. The 4 remaining item statements were: 'This is a close-knit neighbourhood'; 'People around here are willing to help their neighbours'; 'People in this neighbourhood generally don't get along with each other'*; and 'People

	in this neighbourhood can be trusted'. The response categories were scored on a 5-point scale from Strongly agree = '1' to Strongly disagree = '5'. The third item (*) is reverse scored. The scale is scored by summing the 4 items: the higher the score the lower the cohesiveness.
Domains	Neighbourhood social cohesiveness
Respondent type	CAPI, Caregiver.
Waves	Waves 1-4.
Length	4-items used for the POCLS.
Publisher/ cost/ permission	Permission granted by publisher for use in the POCLS.
Psychometric properties	Correlated with informal social control: r = 0.80, p<.0001 (Sampson et al. 1997). Correlated with informal social control: r = 0.68, p<.01 (Xue et al. 2005).
Studies used/ Rationale for use	Used in LSAC and the Project on Human Development in Chicago Neighbourhoods (Earls & Buka, 1997; Sampson et al. 1997). Reliability: Cronbach's alpha between 0.80 and 0.91 for the full scale (Sampson et al. 1997).
Scoring information	<u>The items used in the POCLS are:</u> 'This is a close-knit neighbourhood' 'People around here are willing to help their neighbours' * 'People in this neighbourhood generally don't get along with each other' 'People in this neighbourhood can be trusted' *This item is reversed scored Response categories: Strongly agree ='1' Agree = '2' Neither agree nor disagree = '3' Disagree = '4' Strongly disagree = '5' Scoring:

	The scale is scored by summing the 4-items.
Additional information	One item was deleted for use completely in the POCLS as Aboriginal respondents in the pilot found this item confusing (interpreted the question as being about cultural values). The item deleted was: 'people in this neighbourhood do not share the same values.'
References	Earls, F. & Buka, S.L. (1997). Project on Human Development in Chicago Neighborhoods. Technical Report I. Washington, DC. U.S. Department of Justice, Office of Justice Programs, National Institute of Justice. Sampson, J., Raudenbush, S. W., & Earls, F (1997). Neighbourhoods and violent crime: A multilevel study of collective efficacy, Science, 277, 918–924. Xue, Y., Leventhal, T., Brooks-Gunn, J. & Earls, F.J. (2005). Neighborhood residence and mental health problems of 5 to 11 year olds. Archives of General Psychiatry, 62, 1-10.

2.3. Parenting practices/styles/self-efficacy

Parenting – Warmth

Summary	The Parenting – Warmth scale was used to assess warm parenting practices. The scale consists of 4-items which asks caregivers: 'How often do you tell [Study Child] how happy [he/she] makes you?', 'How often do you have warm, close times together with [Study Child]?'; 'How often do you enjoy listening to [Study Child] and doing things with him/her?'; 'How often do you feel close to [Study Child] both when he/she is happy and when he/she is upset?'. The response categories range from Never/almost never = '1' to Always/almost always = '5'.
Domains	Parenting practices/style – warm parenting.
Respondent type	CASI, Caregiver.
Waves	Waves 1-4.
Length	4-items.
Publisher/ cost/ permission	Publicly available, no cost, permission not needed.
Psychometric properties	Validity:

	The scale is predictive of children's internalising and externalising behaviours. Reliability: As a measure of scale reliability, H-index was calculated in LSAC and H coefficient ranged from 0.92 to 0.96 (excellent) (Zubrick et al. 2014).
Studies used/ Rationale for use	The Warmth scale's advantages are its brevity, development for Australian samples, coverage of core content, and use in LSAC. The items used in the POCLS are items 3-6 of the scale used in LSAC (Zubrick et al. 2014).
Scoring information	<u>The items used in the POCLS are:</u> How often do you tell [Study Child] how happy [he/she] makes you? How often do you have warm, close times together with [Study Child]? How often do you enjoy listening to [Study Child] and doing things with him/her? How often do you feel close to [Study Child] both when he/she is happy and when he/she is upset? Response categories: Never/almost never = '1' Rarely = '2' Sometimes = '3' Often = '4' Always/almost always = '5' Scoring: Items are summed to form a total warmth score ranging from 4 (low warmth) to 20 (high warmth). Missing items are replaced with the mean of completed items for up to 2 missing items. If 3 or more items are missing, respondents are coded as missing on the warmth variable.
Additional information	The scale is derived from the 9-item Warmth scale of the Child Rearing Questionnaire, a tool designed for use with Australian parents of young children (Paterson & Sanson, 1999). The CRQ is not formally published as a measure, but has been used quite extensively. To obtain a reduced item set that retained sound psychometric properties, factor loadings, inter-item correlations and response distributions were examined. Two items with limited response distributions and 1-item that cross-loaded onto other parenting factors were excluded, yielding a 6-item scale in the LSAC version.

References	Zubrick, S.R., Lucas, N., Westrupp, E. M., & Nicholson, J. M. (2014). Parenting measures in the Longitudinal Study of Australian Children: Construct validity and measurement quality, Waves 1 to 4. Canberra: Department of Social Services. Paterson, G. & Sanson, A. (1999). The association of behavioural adjustment to temperament, parenting and family characteristics among 5-year-old children. Social Development, 8, 293-309.

Parenting – Hostility

Summary	The Parenting – Hostility scale was used in Waves 1-4 as a measure of hostile/angry parenting. The scale is brief, consisting of the following 3-items: 'I have been angry with this child'; 'When this child cries, he/she gets on my nerves': and 'I have lost my temper with this child'. The response is scored on a 10-point scale from Not at all = '1' to All the time = '10'. The higher the score, the higher the hostility.
Domains	Parenting practices/style – hostile/angry parenting.
Respondent type	CASI, Caregiver.
Waves	Waves 1-4.
Length	3-items.
Publisher/ cost/ permission	Publicly available, no cost, permission not needed.
Psychometric properties	Reliability: As a measure of scale reliability, H-index was calculated in LSAC and H coefficient ranged from 0.85 to 0.92 (good) (Zubrick et al. 2014).
Studies used/ Rationale for use	The scale is relatively brief but covers key aspects of hostile parenting. It was used in LSAC and the Longitudinal Study of Child Development in Québec.
Scoring information	<u>The items used in the POCLS are:</u> I have been angry with [study child]; When [study child] cries, he/she gets on my nerves I have lost my temper with [study child] Response categories: Not at all 1---2---3---4---5---6---7---8---9---10 All the time

	Scoring: The items are summed to form a total hostility score ranging from 3 (low hostility) to 30 (high hostility). Missing items are replaced with the mean of completed items for 1 missing item. If 2 or more items are missing, respondents are coded as missing on the hostile parenting variable.
Additional information	The LSAC hostility scale uses 5 of the 7 items from the Coercion Subscale of the Parental Perceptions and Behaviours Scale developed for the Longitudinal Study of Child Development in Québec (Institut de la Statistique du Québec, 2000). Two items were excluded as they were regarded as extremely sensitive/ intrusive (assessing frequency of spanking and shaking respectively). For LSAC, some slight re-wording was undertaken. Additionally, response categories were changed from 'never' to 'not at all' and 'always/almost always' to 'all the time', and were reversed so that 'not at all' = 1 and 'all the time' = 10.
References	Zubrick, S.R., Lucas, N., Westrupp, E. M., & Nicholson, J. M. (2014). Parenting measures in the Longitudinal Study of Australian Children: Construct validity and measurement quality, Waves 1 to 4. Canberra: Department of Social Services Institut de la Statistique du Québec (2000). Longitudinal Study of Child Development in Québec (ÉLDEQ 1998-2002): 5-month-old infants, parenting and family relations, Vol 1, Number 10. Québec, Canada: l'Institut de la Statistique du Québec.

Parenting – Monitoring

Summary	The Parenting – Monitoring scale was used to assess parenting practices of monitoring/supervising children. The scale consists of 4-items for the caregiver to answer: 'How often do you know where [study child] is in the course of a day?'*; 'How often do you know who [study child] is with when [he/she] is away from home (i.e. their placement)?'*; 'How often do you talk to [study child] about what's going on in [his/her] life?'*; 'How often does the [study child] go out without telling you where [he/she] will be?. The response categories range from Always = '1' to Never = '5'. The first 3-items* are reverse scored and then items are summed to give a total between 4 (low monitoring) and 20 (high monitoring).
Domains	Parenting practices/style – monitoring/supervision of children.
Respondent type	CASI, Caregiver.

Waves	Waves 1-4.
Age range	Caregivers of study children aged 10-17 years.
Length	4-items.
Publisher/ cost/ permission	Publicly available, no cost, permission not needed.
Psychometric properties	No published information found.
Rationale for use	The scale captures the major aspects of monitoring and is relatively short, it was also used in LSAC.
Scoring information	<u>The items used in the POCLS are:</u> How often do you know where [study child] is in the course of a day* How often do you know who [study child] is with when [he/she] is away from home (i.e. their placement)* How often do you talk to [study child] about what's going on in [his/her] life* How often does the [study child] go out without telling you where [he/she] will be. Response categories: Always = '1' Mostly = '2' Sometimes = '3' Rarely = '4' Never = '5'. Scoring: Items 1, 2 and 3 are reversed scored. The items are summed to form a total monitoring score ranging from 4 (low monitoring) to 20 (high monitoring). Missing items are replaced with the mean of completed items for 1 missing item. If 2 or more items are missing, respondents are coded as missing on the monitoring variable.
Additional information	All items were adapted from the original scale developed for the Iowa Youth and Families project (Whitbeck et al. 1994)
References	Whitbeck, L. B., Lorenz, F. O., Simons, R. L., & Huck, S. (1994). Family origins of personal and social well-being. In R. D. Conger & G. H. Elder (Eds.), Families in troubled times: Adapting to change in rural America (pp. 149-165). New York: Aldine de Gruyter.

Difficult Behaviour Self-Efficacy Scale (DBSES)

Summary	The Difficult Behaviour Self-Efficacy Scale (DBSES) measures caregiver self-efficacy when dealing with challenging behaviours. It was used in this study in Waves 1-4 with caregivers of study children aged 2-17 years. Four items of the scale were used, these asked: 'How difficult do you personally find it to deal with the challenging behaviours of [Study Child]?; 'To what extent do you feel that the way you deal with the challenging behaviours of [Study Child] has a positive effect?'; 'How satisfied are you with the way in which you deal with challenging behaviours of [Study Child]?'; 'To what extent do you feel in control of the challenging behaviours of the [Study Child]? Each item was rated on a 7-point scale ranging from: Very difficult to Not at all difficult; Has no positive effect at all to Has a very positive effect; Not satisfied at all to Very satisfied; and Not in control at all to Very much in control, respectively.
Domains	Measures caregiver self-efficacy when dealing with challenging behaviours.
Respondent type	CASI, Caregiver.
Waves	Waves 1-4.
Age range	Caregivers of study children aged 2-17 years.
Length	4-items.
Publisher/ cost/ permission	Permission granted by creator for use in the POCLS.
Psychometric properties	Cronbach's alpha = 0.94 for mothers and 0.92 for fathers of autistic children (Hastings & Brown, 2002a). Cronbach's alpha = 0.94 for a group of 70 staff members working with children who had intellectual disabilities and/or autism (Hastings & Brown, 2002b).
Studies used/ Rationale for use	This is the only scale that has been developed to assess self-efficacy for dealing with children's challenging or difficult behaviours. The scale has been used in research with foster caregivers (Whenan et al. 2009) and has been found to predict intention to continue providing foster care. High levels of self-efficacy in caregivers may also be a protective factor for children with challenging behaviours, leading to positive outcomes.

Scoring	<u>The 3-items used were:</u> How confident are you in dealing with the challenging behaviours of [Study Child]? Not at all confident 1--2---3---4---5---6---7 Very confident How difficult do you personally find it to deal with the challenging behaviours of [Study Child]? Very difficult 1---2---3---4---5---6---7 Not at all difficult To what extent do you feel in control of the challenging behaviours of [Study Child]? Not in control at all 1—2—3—4—5—6—7 Very much in control The three scores are summed and give a total ranging from 3 to 21. Higher scores indicate the caregiver "perceived a higher level of self-efficacy in managing difficult behaviour" (Whenan et al. 2009)
Additional information	As the scale has a high internal consistency score, some of the items are probably redundant so 2 items were deleted from the scale for use in the POCLS.
References	Hastings, R. P., & Brown, T. (2002a). Behavior problems of children with autism, parental self-efficacy, and mental health. American Journal on Mental Retardation, 107(3), 222-232. Hastings, R. P., & Brown, T. (2002b). Behavioural knowledge, causal beliefs, and self-efficacy as predictors of special educators' emotional reactions to challenging behaviours. Journal of Intellectual Disability Research, 46, 144-150. Whenan, R., Oxlad, M., & Lushington, K. (2009). Factors associated with foster caregiver well-being, satisfaction and intention to continue providing out-of-home care. Children and Youth Services Review, 31, 752-760.

3. Measures completed by childcare and school teachers

3.1. Child socio-emotional development

Caregiver-Teacher Report Form (C-TRF) and Teacher Report Form (TRF)

Summary	The Caregiver-Teacher Report Form (C-TRF) is a parallel version of the CBCL for completion by childcare workers or preschool teachers of children aged 1.5-5 years to assess behavioural and emotional problems in children. The Teacher Report Form (TRF) is a parallel version of the CBCL for completion by a school teacher (TRF) or other school personnel who are familiar with children's functioning in school, such as teacher aides, counsellors, administrators and special educators. The C-TRF and the TRF are the only standardised measures included in the POCLS teacher survey.
Domains	C-TRF The syndrome scales are the same as the CBCL for Ages 1.5-5 years with Sleep Problems not included; Emotionally Reactive, Anxious/Depressed, Somatic Complaints, Withdrawn, Attention Problems, and Aggressive Behaviour and Other Problems. Syndrome groupings: Internalising comprises Emotionally Reactive, Anxious/Depressed, Somatic Complaints, Withdrawn. Externalising includes Attention Problems, and Aggressive Behaviours. Total Problem score is computed by summing scores of Internalising, Externalising, Sleep problems and the other problems that are not part of any syndromes. The DSM-Oriented Scales are: Affective Problems, Anxiety Problems, Pervasive Developmental Problems Attention Deficit Hyperactivity Problems and Oppositional Defiant Problems. TRF The empirically-based syndrome scales are: Anxious/Depressed, Withdrawn/Depressed, Somatic complaints, Social Problems, Thought Problems, Attention Problems, Rule breaking Behaviour, Aggressive Behaviour. Syndrome groupings:

	Internalising Problems includes Anxious/Depressed and Withdrawn/Depressed and Somatic Complaints. Externalising Problems includes Rule breaking and Aggressive Behaviour Scales. The Total Problems Score is the sum of the internalising, externalising, other syndrome and other problems that are not on any of the syndromes. There are also DSM-Oriented Scales which include: Affective Problems, Anxiety Problems, Somatic Problems, Attention Deficit/Hyperactivity Problems, Oppositional Defiant Problems and Conduct Problems.
Respondent type	The teacher questionnaire for the POCLS (which includes C-TRF and TRF) is completed online. Respondents are asked to base their responses on a two month period rather than a six month period used for CBCL.
Waves	Waves 2 and 3. The teacher questionnaire (which includes the TRF and C-TRF) is completed only once. The teacher survey was continued into Wave 3 only for those who did not have a Wave 2 survey, in order to maximise response rates.
Age range	All ages (however, child must attend preschool or school for at least 2 days per week to be eligible for the survey).
Publisher/ cost/ permission	Permission granted by publisher for use in the POCLS.
Psychometric properties	C-TRF Ages 1.5-5 years: Validity: Content Validity: All but 1 item (item 94 unclean personal appearance) discriminated significantly (p<.01) between referred and non-referred children Criterion related validity: All scores were significantly (p<.01) higher for referred than non-referred children. Reliability: Cronbach Alphas were 0.89 for Internalising, 0.96 for Externalising, and 0.97 for Total Problems Scales. Alphas for the specific behaviour problem sub-scales range from 0.52 to 0.89 (p.157, Appendix D, Achenbach 2001). Test-Retest: correlations over an 8-day period were 0.77 for Internalising, 0.89 for Externalising and 0.88 for Total Problems. The

	specific syndrome scales range from 0.68 to 0.91 (N=59). The DSM-oriented scales range from 0.57 to 0.87. Over 3 months, correlations range from 0.40 for Externalising, 0.56 for Total Problems and 0.65 for Internalising The syndrome scales range from 0.22 to 0.71 and for the DSM-Oriented Scales from 0.46 to 0.85 (N=32). Inter-Rater: correlations were 0.64 for Internalising, 0.79 for Externalising and 0.72 Total Problems Scales. The syndrome scales range from 0.21 to 0.78; and the DSM-Oriented scales range from 0.55 to 0.68 (N=102). <u>TRF- Ages 6-18 years:</u> Validity: Criterion related validity: All of the TRF scales show significant discrimination between referred and non-referred children (p<.01). Reliability: Cronbach Alphas were 0.90 for Internalising, 0.95 for Externalising, and 0.97 for Total Problems. Alphas for the specific behaviour problem syndrome scales range from 0.72 to 0.95. (N=3086). Test-Retest: correlations over an 8 day interval were 0.86 for Internalising, 0.89 for Externalising and 0.95 for Total Problems. The empirically-based syndrome scales ranged from 0.60 to 0.96; the DSM-Oriented scales from 0.62 to 0.95; and for the Competence scale from 0.78 to 0.93 (N=44). Inter-Rater: correlations were 0.58 for Internalising, 0.69 for Externalising and 0.55 for Total Problems; and from 0.28 to 0.69 for the empirically-based syndrome scales; from 0.20 to 0.76 for the DSM-Oriented scales; and from 0.37 to 0.58 for the Competence scales (N=88).
References	Achenbach, T.M., & Rescorla, L.A. (2000). Manual for the ASEBA Preschool forms and Profiles. Burlington, VT: University of Vermont Department of Psychiatry. Achenbach, T.M. & Rescorla, L.A. (2001). Manual for the ASEBA School-Age Forms & Profiles. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families

Section C: Source of other questions used in the POCLS

In addition to the standardised measures, the POCLS collects data from children, caregivers, teachers and caseworkers using questions developed by other studies and also questions developed by the POCLS Study Working Group. For more detail, please refer to the questionnaire spreadsheet for question wording and source https://www.facs.nsw.gov.au/resources/research/pathways-of-care/pocls-publication/questionnaires-pocls.

The final column in the questionnaire spreadsheet records if the question source:

- An item from a standardised measure
- An item used by another published study
- Project developed.