Australian Government
Department of Social Services
Australian Institute of Family Studies

Growing Up in Australia:
The Longitudinal Study of Australian Children (LSAC)
LSAC Technical Paper No. 30

Agreement between parent and child responses

Jennifer Prattley, Pilar Rioseco, Amanda Vittiglia, Lisa Mundy and Bosco Rowland

August 2023

Acknowledgements

This report makes use of data from *Growing Up in Australia*: The Longitudinal Study of Australian Children (LSAC). LSAC is conducted in a partnership between the Department of Social Services (DSS), the Australian Institute of Family Studies (AIFS) and fieldwork providers (the Australian Bureau of Statistics (ABS) up to and including Wave 9C, and Roy Morgan, from Wave 10 onwards), with advice provided by a consortium of leading researchers.

The authors would like to thank Chris Schilling and the AIFS LSAC team for their contributions to the paper and project. The authors would especially like to thank LSAC participants for their ongoing contribution to the study.

Growing Up in Australia: The Longitudinal Study of Australian Children is conducted in partnership between the Australian Government Department of Social Services, the Australian Institute of Family Studies (AIFS) and the Australian Bureau of Statistics (ABS), with advice provided by a consortium of leading researchers from research institutions and universities throughout Australia growingupinaustralia.gov.au.

Technical paper

The Longitudinal Study of Australian Children: LSAC Technical paper No. 30. Agreement between parent and child responses

Authors: Jennifer Prattley, Pilar Rioseco, Amanda Vittiglia, Lisa Mundy and Bosco Rowland

For more information, write to:
Longitudinal Studies – Research and Methods
Data Strategy and Development| Data and Evaluation
Australian Government Department of Social Services
PO Box 7576
Canberra Business Centre ACT 2610

Email: LongitudinalStudiesDataAccess@dss.gov.au

ISBN 978-1-76016-289-4 (Online)
ISBN 978-1-76016-290-0 (PDF)

LSAC_Parent-child-responses_paper30

Contents

List of tables

List of figures

Glossary

Term	Description
ICC	Intraclass Correlation Coefficient
LSAC	*Growing up in Australia*: The Longitudinal Study of Australian Children
LOA	Limits of agreement
RMA	Reduced major axis
SDQ	Strengths and Difficulties Questionnaire
Parent 1 (P1)	Defined as the parent who knows the study child best; in most cases, this is the child's biological mother
Parent-child dyad	A pair comprised of the study child and their responding parent
PedsQL	Pediatric Quality of Life
Study child	A term used to describe the child as the original subject of the survey

Overview

This technical paper examines the concordance between parents' and children's responses in *Growing Up in Australia*: The Longitudinal Study of Australian Children (LSAC). It provides evidence to inform analysis of LSAC items containing parent and child responses.

A number of measures were collected from both informants (i.e. parent and child) at the same time points between the ages of 10–11 and 16–17 years. The measures analysed in this paper relate to participation in and reasons for the child's paid work; difficulties changing schools; general health and use of prescription medicine; and family cohesion. Scores of the Pediatric Quality of Life Inventory (PedsQL) Physical Health scale and subscale and total scores of the Strength and Difficulties Questionnaire (SDQ) were also examined.

The first aim was to determine the extent of agreement between parents' and children's responses on the above items and measures at different ages. We then established the direction of any disagreement and determined whether parent or child responses led to over- or under-reporting of certain behaviours relative to the other informant's responses at each time point. Finally, factors associated with over- or under-reporting were examined for the PedsQL Physical Health scale and SDQ total score. Factors included the parent-child relationship, family psychosocial characteristics and socio-demographic status. The implications of these findings for analysis are discussed.

Summary of findings

- Participation in paid work was surveyed when children were aged 14–15 years. In 6% of cases, parents responded differently to their child on the question of whether the child had done paid work. There was substantial agreement between parent and child responses on items relating to reasons for paid work.
- Parents and children were asked about the child's experiences in changing schools at age 12–13. Agreement on whether or not the child had experienced difficulties was generally high among those who had changed, although responses were different in 15% of cases. Children citing missing friends and more homework, while parents reporting demanding schoolwork, a larger school, and more subjects and teachers accounted for the majority of disagreements.
- Agreement on family cohesion, asked at age 16–17, was moderate. In 65% of parent-child dyads (i.e. pairings) the parent and child gave different responses from each other. In nearly half of cases where there was disagreement, the child reported the family got along less well than what was reported by the parent.
- There was high concordance on use of prescribed medicine, asked at age 16–17. Ninety-two per cent of parent-child dyads showed agreement, with disagreement mostly attributable to the parent indicating the child currently needed or used prescribed medicine but the child indicating they didn't. If only parent data were used, prevalence on need or use of medicine would be around 2% higher than if child data were used.
- The prevalence of disagreement on general health (asked at ages 14–15 and 16–17) was high at around 55%. At both ages, 38%–39% of children reported their health to be better than what was indicated by their parent, while the parent reported better child health than the child in 16% of cases.
- The prevalence of disagreement on the PedsQL Physical Health scale score (47% at age 14–15 and 45% at age 16–17) and SDQ subscale and total scores (assessed at ages 10–11 through to 16–17 with prevalence ranging from 12%–27%) was such that different results and conclusions could be obtained from parent compared to child responses.
- Scores calculated from parent responses were more likely to under-estimate children's physical health quality of life and over-estimate their social-emotional wellbeing compared to what would be indicated from children's responses.
- Aspects of the parent-child relationship affected the ability of parents and children to assess the child's wellbeing in a similar way. Where there was frequent arguing, parent scores were more likely to indicate higher overall levels of psychosocial problems and lower physical health than what was suggested by child responses. If trust and communication between parents and children was low, scores from parents tended to under-estimate psychosocial difficulties and over-estimate physical health, on average.

Introduction

Cohort studies often collect data from multiple informants. In the case of birth cohorts starting early in life, parent proxies are key informants on the development and wellbeing of the study child. As study children get older, and once the content area is age appropriate, survey data are also collected from the children themselves, with some measures collected simultaneously from both children and parents. Having more than one informant on a particular measure enhances the data by capturing both perspectives. However, it also raises the issue of agreement. Parents and children may provide different accounts of the child's experiences, with consequences for analysis. This paper investigates the extent to which parent and child responses agree on certain measures, and what the implications for researchers might be.

Parent-child agreement matters

Parent-child agreement on survey items has implications for research and analysis (Gray, Scott, Lawrence, & Thomas, 2021). If there are low levels of agreement between parent and child responses on an item, then analysing information from just one source could lead to an incomplete or biased representation of the study child's circumstances (Cleridou, Patalay, & Martin, 2017; Gray et al., 2021; Hogan, Garcia, Tomko, Squeglia, & Flanagan, 2023), especially where it is not clear who the more reliable informant might be. Disagreement could reflect differences in opinion or perception between both parties, and researchers could draw different conclusions depending on whether the responses of both parent and child, only the child or only the parent were analysed. For example, evidence suggests that parents' views often differ from children's in regard to socio-emotional wellbeing and quality of life (see Michels et al., 2013; Sanker et al., 2015). This can occur in longitudinal studies where parent reports of the child are collected from a much younger age, providing longer periods of observation.

The nature and meaning of any disagreement between parent and child responses is important for researchers in evaluating bias that may arise from using only one informant's data. Disagreement could stem from either parents or children under- or over-reporting relative to the other. Furthermore, the tendency to disagree may be associated with a particular parent, child or systemic and environmental characteristics. For example, parents from lower socio-economic backgrounds are more likely to report discordance (Hogan et al., 2023; Rebholz et al., 2014). If analysis was conducted on items from only one informant, the realities of those groups could be misrepresented. Additionally, prevalence statistics inform resource allocation, and under-estimation, in particular, risks under-allocation of necessary resources (Gray et al., 2021).

In addition to the consequences for analysis, it is important to determine concordance on health-related measures as parents are key to children's access and use of health care services (particularly among younger children). Parent-child agreement on the child's problems or symptoms is likely to influence treatment and outcomes (Goolsby et al., 2018; Yeh & Weisz, 2001). *Growing Up in Australia*: The Longitudinal Study of Australian Children (LSAC) is well-placed to examine the role of the parent-child relationship on parents' and children's ability to identify and assess wellbeing in a similar way.

Multi-informant measures in LSAC

LSAC contains a number of items and measures that were asked of both parents and study children, sometimes across multiple waves, and there were several reasons why this happened. This paper investigates concordance on seven indicators. They are introduced here with further details given in later sections.

Participation in paid work, difficulties changing schools and family cohesion

Items on reasons why the study child participated in paid work, difficulties associated with changing schools, and family cohesion were asked of both parents and children because it was expected that each could have different perspectives, and these differences might be insightful and of interest to researchers. For example, studies might investigate whether agreement on reasons why the study child worked was associated with a higher probability of the child remaining in work. Responses to the item on the main reason for working might differ where the child reports working for spending money but the parent perceives it as helping family in the family business, for example.

Use of prescribed medicine

In the case of the study child's use of prescribed medicine, it was not clear who the best person would be to provide this information to the study. Study designers hypothesised that some participants may be more independent of their parents than others in adolescence and attend medical appointments unaccompanied, particularly as LSAC study children got older. In such instances, the child could have more knowledge of their medication than the parent. However, the parent may be at least or more informed than the child when they did attend appointments together. Asking parents and study children this item was likely to provide the most complete picture of the use of prescribed medicine by young people in mid-adolescence.

General health

The Global Health Measure (part of the SF-6; Ware, Nelson, Sherbourne, & Stewart, 1992) was designed to be administered to people over 14 years, and consequently only parents were asked to complete this measure while children were in early to mid-childhood. The scale was included on the study child questionnaire from age 14–15; however, it was deemed prudent to continue to collect the parent responses for comparison, and to enable data users to examine any differences in reporting of perceived general health, particularly during mid-adolescence.

Pediatric Quality of Life (PedsQL) Physical Health scale score

The physical health subscale of the PedsQL (Varni, Seid, & Kurtin, 2001) is a standardised health and wellbeing instrument applicable to children and adolescents aged 2–18. It was specifically designed to collect the same information from more than one respondent. In LSAC, data were collected from both parents and study children when the children were aged 14–15 and 16–17.

Strength and Difficulties Questionnaire

The Strengths and Difficulties Questionnaire (SDQ; Goodman, 2001) is a standardised measure of child wellbeing designed to obtain complimentary perspectives, and the use of both parent and child responses has been recommended at ages 3–16 (e.g. Eiser & Morse, 2001; Goodman, Ford, Simmons, Gatward, & Meltzer, 2003). In LSAC, it was administered to the parent who knew the child best (Parent 1, in most cases the child's biological mother); another adult in the same home with a parental relationship where relevant; parents living elsewhere; and teachers/child care workers involved with the study child. This report uses responses from Parent 1 who reported on their child from age 4–5 until 16–17, and those from children who reported on their own wellbeing from the age of 10–11.

Gray and colleagues (2021) assess concordance between adolescents aged 11–17 and parents on the SDQ using the Second Australian Child and Adolescent Survey of Mental Health and Wellbeing. We do a similar analysis in this paper using LSAC, because of its significance for research into child and family wellbeing in Australia (see e.g. Christian, Mitrou, Cunneen, & Zubrick, 2020; May & Williams, 2022; Sanders, Parker, del Pozo-Cruz, Noetel, & Lonsdale, 2019).

Aims of this paper

This paper examines the level of concordance – in the form of agreement or disagreement – between parent and child responses in LSAC between ages 10–11 and 16–17. The aim is to provide important insight for analysts using the measures of interest. In this context, the present study:

1. Establishes the level of agreement between parent and child responses on selected items and measures where both informants were asked the same questions at the same point in time.
2. Describes disagreement and its direction for those items and measures.
3. Identifies factors associated with parents under- or over-reporting the PedsQL Physical Health score and total SDQ score, including aspects of the parent-child relationship and family socio-demographic characteristics.

The next section describes LSAC, the measures of interest and samples in more detail. The analytical methods used are described after that, followed by results. A summary of findings, discussion of implications and conclusion are given at the end of the paper.

Data and measures

1. The Longitudinal Study of Australian Children

LSAC has followed the development and wellbeing of two cohorts of children since Wave 1 in 2004 (Mohal et al., 2021; Soloff, Lawrence, & Johnstone, 2005). Subsequent Waves were collected every two years. Children in the B ('babies') cohort were born between March 2003 and February 2004 and aged 0–1 at Wave 1. Children in the K ('kindergarten') cohort were born between March 1999 and February 2000 and aged 4–5 at Wave 1. The study is considered broadly representative of Australian children born at these times. For a detailed description of the study design, see Gray and Smart (2009).

The study child is the focus of LSAC. Information is collected from the child (using physical measurement, cognitive testing and, depending on the age of the child, interviews); the parent who knows most about the child (usually the primary carer, known as Parent 1 or P1, in most cases the child's biological mother); any secondary parent in the household (biological, adoptive or step-parents); home-based and centre-based carers for preschool children who are regularly in non-parental care, and teachers for school-aged children. From Wave 2, information was also obtained from parents who live in a separate household from the primary carer but who still have contact with the child. All parent-reported measures in this paper were collected from Parent 1.

The sample from LSAC that was analysed for concordance varied by cohort and age of the study child, according to the availability of items. This is detailed for each item below and summarised in Analytic Samples.

2. Participation in paid work

Paid employment information was collected at age 14–15 from K cohort respondents and their parents. Parents were asked 'In the last 12 months has child worked? (Excludes unpaid work experience, voluntary work, household chores for pocket money)'. Children were asked 'In the last 12 months have you worked?' These two questions were used as filters for the question of interest, which was the study child's main reason for working. Available reasons to select from were:

- For spending money
- To save up for something
- For friendship or to develop social skills
- To improve career prospects
- To help family in family business
- Other.

The reasons for children also included the option 'to supplement family income', which was not included in the parents' item. Only one respondent selected this reason. Therefore, analysis of agreement was still performed, excluding this item.

For consistency in terminology with other measures, cases where the child reported working or where a reason was reported by the child but not the parent were designated as 'child over-reporting' relative to the parent. Instances where the parent reported the child working or a reason when the child did not were labelled as 'parent over-reporting' relative to the child.

3. Difficulties changing schools

At age 12–13, B and K cohort parents and children were asked whether the child had changed schools since the last interview approximately two years prior.

Parents who responded 'yes' were then asked, 'Did the child experience any difficulties in changing schools?' Those who indicated 'yes' to this were asked which of the following the child had difficulty with (response options of 'yes' or 'no' in each case):

- Making new friends
- Missing friends from previous school
- Coping with larger school with more students
- Dealing with more school subjects with different teachers
- Coping with more demanding schoolwork
- Being required to do more homework
- Managing different travel arrangements to/from school
- Other.

Children who indicated they had changed schools were also asked if they had experienced any difficulties ('yes' or 'no'). Those who responded 'yes' were asked what was difficult about the change; they were given the same list of difficulties as above, with 'yes' or 'no' responses in each case.

Child over-reporting compared to their parent on individual difficulty items was identified where they responded 'yes' but the parent responded 'no'. Similarly, parent over-reporting relative to the child occurred where the parent responded 'yes' but the child 'no'.

4. Family cohesion

K cohort respondents at age 16–17, and their parents, were asked 'Sometimes family members may have difficulty getting along with one another. They don't always agree and they may get angry. In general, how would you rate your family's ability to get along with one another?' Response options included 1 = Excellent; 2 = Very good; 3 = Good; 4 = Fair; 5 = Poor. Child over-reporting relative to parents occurred where their response indicated a better ability to get along, for example 'Good' compared to 'Fair'. Parent over-reporting occurred where their response indicated a better ability to get along than the child response.

5. Use of prescribed medicine

K cohort respondents and parents were asked at age 16–17 whether the young person currently needed or used medicines prescribed by a doctor ('Does child/do you currently need or use medicine prescribed by a doctor, other than vitamins?'). Response options were 'yes' and 'no'. Child over-reporting was identified where they indicated 'yes' but the parent said 'no'. Parent over-reporting occurred where parents responded 'yes' but children responded 'no'.

6. General health status

Parents and study children in the K cohort were asked to report on study child health when the child was aged 14–15 and 16–17. The item administered to study children asked 'In general, would you say your own health is …' with response options of 1 = Excellent; 2 = Very good; 3 = Good; 4 = Fair; 5 = Poor. Parents were asked 'In general, how would you say study child's current health is?' with the same response options.

Child over-reporting relative to parents occurred where their response indicated better health compared to what was indicated by the parent response; for example, 'Good' rather than 'Fair'. Parent over-reporting was defined similarly, where the parent response suggested better health than the child response did.

7. Pediatric Quality of Life (PedsQL) Physical Health scale score

The PedsQL Physical Health scale score was calculated from responses to 8 individual items, asked to K cohort parents and children when children were aged 14–15 and 16–17 (Table 1). Study children were asked, 'In the last month, how much of a problem has this been for you ...?' with difficulties as given. Parents were asked, 'In the past month, how often would you say the study child has had a problem with ...?' with the activities listed. There were minor differences in phrasing between the child and parent versions; for example, walking 'more than 100 metres' versus 'more than one block', and 'play sport or do exercise' compared to 'sports activity'. Both parties had the same response options for each item, namely 'Never'; 'Almost never'; 'Sometimes'; 'Often'; and 'Almost always'.

Table 1: Specified activities of the Pediatric Quality of Life Physical Health Scale

Activities specified on child questionnaire	Activities specified on parent questionnaire
It is difficult for me to walk more than 100 metres	Walking more than one block
It is difficult for me to run	Running
It is difficult for me to play sport or do exercise	Participating in sports activity
It is difficult for me to lift something heavy	Lifting something heavy
It is difficult for me to have a bath or shower by myself	Taking a bath or shower by him/herself
It is difficult for me to help around the house	Doing chores around the house
I get aches and pains	Having hurts or aches
I have low energy	Low energy level

This paper examines the extent and direction of concordance between parents and children on PedsQL total scores. Scores were calculated where less than 5 responses from the 8 items were missing. Individual item responses were coded to a numeric value using 'Never' = 100; 'Almost never' = 75; 'Sometimes' = 50; 'Often' = 25 and 'Almost always' = 0. The total score was calculated as the mean of those values, and higher scores indicated better physical health quality of life (due to reverse coding of numeric values).

In line with previous research (White-Koning et al., 2007), disagreement between parent and child PedsQL Physical Health scores was defined as a difference greater than 0.5 standard deviations of the child's score, which is considered a meaningful difference in quality-of-life studies. A parent-child pair (i.e. a dyad) was classified as parent under-reporting the child's quality of life where there was disagreement and the parent score was lower than the child score. It was designated as parent over-reporting where there was disagreement and the parent score was higher than the child score.

Factors associated with disagreement were also analysed. They included aspects of the parent-child relationship, mental health of the parent and child and socio-demographic characteristics. Section 9 has details.

8. The Strengths and Difficulties Questionnaire (SDQ)

The SDQ (Goodman, 2001) comprised 25 items that were organised into five subscales of hyperactivity; emotional symptoms; conduct problems; peer problems; and prosocial behaviour (Table 2). Items were asked to B cohort parents and children when children were aged 10–11 and 12–13, and to the K cohort when children were aged 10–11; 12–13; 14–15 and 16–17. Respondents were asked to indicate for each item which of 'Not true', 'Somewhat true' or 'Certainly true' best described the study child's behaviour. Responses were numerically coded as 0, 1 and 2 respectively, with reverse coding where appropriate.

Note that there were slight variations in the time frame of reporting of behaviours depending on the age of children, the cohort and whether it was the study child or parent responding. When children were aged 10–11, all study children and parents from the K cohort were asked about behaviours over the 'last six months', whereas B cohort parents were asked about 'the last six months or this school year'. At age 12–13, all study children and parents were asked about 'the last six months or this school year'. However, at ages 14–15 and 16–17 children reported on 'the last six months' while parents were asked about 'the last six months or this school year'. Hence, there is the possibility that discrepancies between child and parent responses at ages 10–11, 14–15 and 16–17 could be at least partly explained by differences in the time frame reported on.

Subscale scores

For each of the five subscales, a continuous score was calculated from each of the parent and child responses. It was derived when individuals had less than 3 items missing responses and calculated as the mean value rescaled to an integer between 0 and 10. Higher scores represented more problems, except for on the prosocial behaviour subscale, where a higher score indicated fewer problems.

Subscale classification

Subscale scores were categorised as either indicating 'average' or 'raised/high' ('low' for prosocial) symptoms using cut points for each of the parents and children, given in Table 2. These cut points were derived from United Kingdom samples as published elsewhere (Australian Mental Health Outcomes and Classification Network [AMHOCN], 2005). Child over-reporting relative to parents on the hyperactivity, emotional symptoms, conduct problems and peer problem scales was identified where categorisation based on scores from child responses was 'raised/high' but that from parent responses was 'average'. Parent over-reporting occurred where categorisation based on their responses was 'raised/high' but from child responses was 'average'. On the prosocial scale, child over-reporting occurred where categorisation from child responses was 'average' but from parent scores was 'low'. Parent over-reporting was identified where child responses indicated 'low' but parents indicated 'average'.

Total score

Total difficulties scores for each of parent and child responses were calculated by summing scores for the hyperactivity, emotional symptoms, conduct problems and peer problems subscales. Total scores ranged from 0 to 40 and a higher score indicated more problematic behaviour or emotions. A threshold to establish a meaningful difference between informants has not been established for the SDQ (Youthinmind, n.d.). However, research suggests that each point increase in difficulties scores is associated with increased probability of a clinical diagnosis (Goodman & Goodman, 2009). Given that the estimated population standard deviation for the total difficulties scores is approximately 5–6 points, a difference of 0.5 standard deviations of the child's score (around 3 points) was considered meaningful for the purpose of examining agreement (Youthinmind, n.d.). Thus, parent under-reporting was identified where their SDQ score was lower than 0.5 standard deviations from the child's score, and over-reporting where it was more than 0.5 standard deviations higher.

This paper examined the extent and direction of disagreement between parents and children firstly on scores for each subscale; secondly, on classification of subscale scores into average or raised/high symptoms; and, thirdly, on the total SDQ score. Factors associated with disagreement on the total score were also examined (as described in section 9).

Table 2: Strengths and Difficulties Questionnaire items by subscale

Subscale	Scores for raised/high category (low for prosocial behaviour)^		Items
	Parent	Child	
Hyperactivity	6–10	6–10	Restless, overactive, cannot stay still for long
			Constantly fidgeting or squirming
			Easily distracted, concentration wanders
			Thinks things out before acting
			Good attention span, sees tasks through to the end
Emotional symptoms	4–10	6–10	Often complains of headaches, stomach aches or sickness
			Many worries, often seems worried
			Often unhappy, downhearted or tearful
			Nervous or clingy in new situations, easily loses confidence
			Many fears, easily scared
Conduct problems	3–10	4–10	Often loses temper
			Generally well-behaved, usually does what adults request
			Often fights with other children or bullies them
			Often lies or cheats
			Steals from home, school or elsewhere
Peer problems	3–10	4–10	Rather solitary, tends to play alone
			Has at least one good friend
			Generally liked by other children
			Picked on or bullied by other children
			Gets on better with adults than with other children
Prosocial behaviour [a]	0–5	0–5	Considerate of other people's feelings
			Shares readily with other children (treats, toys, pencils, etc.)
			Helpful if someone is hurt, upset or feeling ill
			Kind to younger children
			Often volunteers to help others (parents, teachers, other children)

Note: [a] Lower scores on the prosocial behaviour subscale indicate lower levels of prosocial behaviour.

9. Factors associated with disagreement

Factors associated with disagreement were examined for the PedsQL Physical Health scale score and SDQ total score. Factors were informed by literature summarised in the 'Introduction' and included the quality of the parent-child relationship, parental and children's mental health and socio-demographic characteristics.

Table 3 presents the measures and response categories for each variable. Frequencies for each at baseline are presented in the Appendix (Table 15), where baseline was defined as Wave 5 for K cohort and Wave 7 for B cohort. Children were aged 12–13 in these waves.

Table 3: Measures examined in analysis of factors associated with agreement

Area	Variable	Description
Socio-demographic characteristics	Sex	Male, female
	Family structure	Single-parent family, two-parent family
	Parents' country of birth	At least one parent born in Australia; at least one parent born in English speaking country (not Australia); other countries of birth
	Child disability	Child has a disability or health condition for 6 months or more
	Parental education	Either parent has a university degree; other levels of education
	Number of siblings	Only child; one sibling; two or more siblings
Parent-child relationship (child report)	Parent and young person disagree and fight	Not at all; a little; sometimes or more often
	Parental trust scale	Very high; moderate/high; lower (based on mean score)
Main caregiver psychological distress	Psychological distress K6 scale	Main caregiver score in K6 scale (continuous)
Child internalising problems	Short Mood & Feelings Questionnaire	Elevated depressive symptoms (score >=8), below depression cut off (score <8)
Child externalising problems	Negative social behaviours	Number of negative social behaviours in the last 12 months (e.g. got into physical fights; stolen something from a shop)

10. Analytic samples

Table 4 summarises, for each item or measure of interest, the age of children in the sample, cohort(s) the data was drawn from, and the Wave and year of data collection. The maximum possible sample size (i.e. the number of parent-child pairings or dyads analysed) for each age group is also given. For example, the sample analysed for concordance on the item 'Did the child experience any difficulties in changing schools?' was comprised of children aged 12–13 drawn from B cohort Wave 7 (data collected in 2014) and K cohort Wave 5 (collected in 2010). The maximum possible number of parent-child dyads for that analysis was N = 7,337, with 3,381 of those coming from B cohort and 3,956 from K cohort.

The final size of the analytic sample for each item/measure was usually lower than the maximum possible; it was determined by the number of parent-child dyads that had missing data on the item of interest and on responses to any relevant filter questions asked in the survey. Further details on sample selection, non-response and missing data for each measure are in the Appendix. Briefly, dyads were retained in the analytic sample if there were valid responses from both parties. Cases were removed if either or both of the study child or parent responses were missing. The number of dyads included in the final analytic sample is stated in the appropriate results section later in this report. To continue the example from above, the final number of dyads used to analyse concordance on the item 'Did the child experience any difficulties in changing schools?' was 4,953, as shown in the Results section, Table 6. Information in the Appendix shows 460 dyads were removed due to missing data, 1,745 were removed due to having responded 'no' to the filtering question that asked whether they had changed schools since the last interview, and 179 were removed due to inconsistent parent and child responses on that same filtering question.

Table 4: Age , cohort and sample size for concordance analysis

Items/measures	Age 10–11	Age 12–13	Age 14–15	Age 16–17
Difficulties changing schools		Pooled cohorts B W7 2014 K W5 2010		
Paid work and reasons for working			K W6 2012	
Family cohesion				K W7 2014
General health status			K W6 2012	K W7 2014
Use of prescribed medicines				K W7 2014
Pediatric Quality of Life (PedsQL)			K W6 2012	K W7 2014
Strengths and Difficulties Questionnaire (SDQ)	Pooled cohorts B W6 2012 K W4 2008	Pooled cohorts B W7 2014 K W5 2010	K W6 2012	K W7 2014
Maximum possible sample size (number of parent-child dyads)	Pooled *N* = 7,933 B W6 *N* = 3,764 K W4 *N* = 4,169	Pooled *N* = 7,337 B W7 *N* = 3,381 K W5 *N* = 3,956	K W6 *N* = 3,537	K W7 *N* = 3,089

Notes: K: K cohort; B: B cohort; W: Wave. For each item/measure analysis was conducted separately for each age group unless otherwise specified in results.

Analytical strategy and methods

Methods used depended on whether the item or measure in question was categorical or continuous. Items relating to education and work, general health, use of prescribed medicine and family cohesion were categorical. Scores for the PedsQL Physical Health scale and the subscale and total scores for the SDQ were continuous. Classification of SDQ subscale scores was categorical.

1. Establishing the level of agreement

Categorical measures

The level of agreement on education and work, general health, use of prescribed medicine and family cohesion was assessed using **prevalence** and **Gwet's Agreement Coefficient (AC)** (Gwet, 2008). This was also done for the classification of SDQ subscale scores into average or raised/high. The prevalence of disagreement was calculated as the percentage of parent-child dyads in the sample where parent and child responses were different (as defined for each measure in the previous section).

Gwet's AC is a measure of chance-corrected agreement robust to bias and prevalence (Xie, 2013). The number of subjects, number of raters and number of categories being analysed were taken into account when establishing the probability distribution of the agreement coefficient (Gwet, 2014). Linear weights were used to estimate this for variables with more than two categories (namely, general health status and family cohesion). Table 5 shows the benchmark scale used to assess agreement.

Table 5: Benchmark criteria for assessing level of agreement based on Gwet's AC

Gwet's AC	Classification
<0.00	Poor
0.00 – 0.20	Slight
0.20 – 0.40	Fair
0.40 – 0.60	Moderate
0.60 – 0.80	Substantial
0.80 – 1.00	Almost perfect

Notes: The colour code used above is applied in all agreement tables for categorical variables presented in the following sections. AC: Agreement coefficient.

Continuous measures

A number of methods were used to evaluate concordance on the PedsQL Physical Health scale score, each of the subscale scores of the SDQ, and the total score of the SDQ, as follows.

Firstly, the **Intraclass Correlation Coefficient** (ICC; Gisev, Bell, & Chen, 2013; Koo & Li, 2016) was calculated. The ICC is an indicator of inter-rater reliability that assesses whether similar conclusions would be reached based on parent compared to child assessments. It considers the variance of scores for the study child rated by themselves and the parent, as well as the variance of the full sample. Higher values of the ICC show greater concordance between parent and child scores; as a guideline, values less than 0.5 indicate low reliability, between 0.5 and 0.75 moderate, and greater than 0.75 good or excellent (Koo & Li, 2016).

A second statistic, the **Absolute Mean Difference**, was calculated as the mean of the difference between the parent's and child's scores from each dyad irrespective of the direction of difference. Finally, the correlation between the difference of parents' and children's scores and the mean of both scores was derived (referred to as '**Correlation difference/mean**'). It indicates whether the difference between informants varied across the distribution of means. Larger correlation indicates whether the difference between parents' and children's scores varied as a function of the mean of those two scores.

Concordance was also assessed visually. **Bland-Altman plots** (Bland & Altman, 2010) indicate systematic differences between parent and child scores. The mean of the parent's and child's score is plotted on the horizontal axis and the difference between the two scores on the vertical. If there is no correlation between the mean and difference of scores, the mean of all differences is shown as a horizontal line. On such graphs, two dotted lines above and below the mean line identify the **Limits of Agreement (LOA)**, which are 1.96 standard deviations of the mean difference. It is expected that 95% of all differences between parent and child scores lie within these limits (Bland & Altman, 2010). If there is non-zero correlation between the mean and difference of scores, a regression model is used to adjust the limits of agreement accordingly, and Bland-Altman plots are presented with trends.

2. Describing disagreement and its direction

After establishing the level of agreement for each item/measure, we assessed the direction of disagreement. Again, the method used depended on whether the measure was in categorical or continuous form.

Categorical measures

For each of the categorical measures (those relating to education and work, general health, use of prescribed medicine, family cohesion, and subscale classifications of the SDQ) we calculated the proportion of parent-child dyads with disagreement where the child had over-reported compared to the parent. The definition of 'child over-report' used for each measure was described in the appropriate measures section above. The equation used was:

$$\textbf{\% Child over report} = \frac{\textit{Number dyads with child over report}}{\textit{Total number of dyads with disagreement}} \times 100$$

The extent of parent over-reporting was the complement of this. 'Parent over-report' for each measure was as defined earlier. Based on that, the following was used:

$$\textbf{\% Parent over report} = \frac{\textit{Number dyads with parent over report}}{\textit{Total number of dyads with disagreement}} \times 100$$

Continuous measures

Direction of agreement on the continuous PedsQL Physical Health scale score, each of the subscale scores and the total score of the SDQ was assessed using the following methods:

The Directional Mean Difference was the mean of the differences between parents' and children's scores, accounting for the direction of the difference. Negative values indicate that children's scores were higher than parents' on average. The **Effect Size** (Cohen's *d*) is the standardised mean difference between scores, using the standard deviation of the children's scores. It gives an indication of how large the difference between two groups is, or the degree to which the difference is present in the population. Effect sizes of 0.2, 0.5 and 0.8 are considered 'small', 'medium' and 'large' respectively (Cohen, 1988).

Bland-Altman plots (Bland & Altman, 2010), as explained above, provide visual evidence of directionality. They show the mean of all differences as a horizontal line. A positive y intercept on that line suggests total scores calculated from parent responses were, on average, higher than those calculated from child responses. A negative y intercept indicates scores derived from parent responses tended to be lower, on average, than those calculated from child responses.

Reduced Major Axis (RMA) plots (Steichen & Cox, 1998) were also used to assess disagreement and direction visually. In RMA plots, children's scores are on the horizontal axis and parent scores on the vertical. A line of perfect concordance is shown, which has a slope of 1 and intercept of 0. A second line, the RMA, is the line of best fit calculated from the data and passes through the intersection of the mean parent and mean child scores. It has slope given by the sign of Pearson's correlation coefficient and the ratio of the standard deviations. Divergence between the line of perfect concordance and RMA indicates proportional bias.

3. Identifying factors associated with disagreement

The third aim of the paper was to examine factors associated with disagreement between parent and child scores on the PedsQL Physical Health scale score and total SDQ score, using children's responses as the reference point.[1]

Once each parent-child dyad was categorised as either having agreement, parent over-report or parent under-report, a series of logistic regression models were constructed. The number of models and sample used varied according to measure. For examining disagreement on the PedsQL Physical Health scale score, six models were fitted in total. Three models examined factors associated with parent over-report relative to agreement, all using data from K cohort; one used data collected when children were aged 14–15, the second used data collected when children were aged 16–17 and a third was estimated from pooled data covering ages 14–17. A second set of three models for examining factors associated with parent under-report (relative to agreement) were fitted to the same samples. Conducting analysis on pooled as well as separate age groups allowed us to quantify any difference in effect according to age and increased the number of smaller population groups in the analytic sample (e.g. single-parent families).

For examining disagreement on the SDQ total score, eight models were fitted. Four models examined factors associated with parent over-report relative to agreement. The first of these used data pooled from the B and K cohorts, collected when children were aged 12–13. The second used K cohort data collected when children were 14–15, while the third used K cohort data from children aged 16–17. The fourth model was estimated from pooled data covering ages 12–17. A further series of four models for determining factors associated with parent under-report (relative to agreement) was estimated from the same samples. Data from children aged 10–11 were not included in any of these samples due to unavailability of key predictor variables.

All logistic regression models – for both the PedsQL Physical Health scale score and total SDQ score – were fitted with robust standard errors and included a random effect for individuals where estimated from pooled data. Factors examined in all models included socio-demographic characteristics, family functioning and parent-child relationship (as described in section 9).

Results

1. Participation in paid work

There was high concordance between parents and children on the filtering question of whether the young person did paid work in the last 12 months, with a Gwet's AC of 0.89 calculated from a sample size of N = 3,303 taken from the K cohort at age 14–15 (Table 6). Ninety-four per cent (n = 3,115) of parent-child dyads showed agreement in that the parent and child gave the same response as each other.

Agreement on reasons for paid work was analysed for N = 1,278 dyads and examined separately for each reason. The two main reasons why young people reported doing paid work were 'for spending money' (45.6%, n = 582) and 'to save up for something' (36.8%, n = 470). Parent-child agreement on these items was moderate (Gwet's

1 In line with previous research (White-Koning et al., 2007) children's scores are used as the reference point, but this does not imply that children's responses are 'correct' and parents' are not.

AC 0.45 and 0.56 respectively; Table 6). The prevalence of disagreement in these reasons appeared high (28% and 24% respectively), but it is possible that they were seen as interchangeable and similar in meaning, which would account for some of the discordance. Among young people who reported they worked 'for spending money' 13.2% of parents stated 'to save up for something'. Among young people who reported 'to save up for something', 29.6% of parents responded 'for spending money'.

Gwet's AC for each of the reasons 'for friendship or to develop social skills', 'to improve career prospects' and 'to help in family business' was between 0.93 and 0.95, indicating high levels of agreement (Table 6). The prevalence of disagreement ranged from 4% to 7%. However, the number of parents or children who selected these reasons was low and the high agreement was due to both parents and children signifying that these statements did not reflect the main reason for the study child's paid work, rather than both indicating that they did. Only 31 children had indicated 'for friendship or to develop social skills', 51 'to improve career prospects' and 58 'to help in family business'.

Table 6: Agreement difficulties changing schools and paid work

	N parent-child dyads	Prevalence of disagreement		Gwet's AC	Discordant parent-child dyads	
		n	%		% Child over-report[a]	% Parent over-report[a]
Paid work at age 14–15						
Paid work in last 12 months	3,303	188	6	0.89	61	39
For spending money	1,278	352	28	0.45	41	59
To save up for something	1,278	312	24	0.56	64	36
For friendship or to develop social skills	1,278	55	4	0.95	36	64
To improve career prospects	1,278	87	7	0.93	44	56
To help in family business	1,278	64	5	0.94	33	67
Difficulties changing schools age 12–13						
Whether difficulties	4,953	744	15	0.80	43	57
Travel to/from school	380	54	14	0.83	56	44
Making new friends	380	98	26	0.49	44	56
Missing friends	380	102	27	0.47	61	39
More homework	380	105	28	0.53	62	38
Demanding schoolwork	380	121	32	0.45	36	64
Larger school	380	123	32	0.44	33	67
More subjects/teachers	380	136	36	0.36	41	59

Notes: [a] Percentage out of the total disagreement. AC: agreement coefficient. Dark green: 'almost perfect' agreement; light green: substantial agreement; blue: moderate agreement; yellow: fair agreement; peach: poor/slight agreement.

Source: Difficulties changing schools: K cohort, Wave 5, and B cohort, Wave 7. Paid work: K cohort, Wave 6

2. Difficulties changing schools

Analysis on difficulties changing schools was conducted on data collected when children were aged 12–13, pooled from K cohort Wave 5 and B cohort Wave 7 (*N* = 4,953 who changed schools). There was substantial agreement between children and parents on the filtering question regarding whether children experienced any difficulties in changing with 85% of parent-child dyads giving the same response (Table 6).

Agreement on specific difficulties was examined for those who had experienced difficulties (*N* = 380, Table 6). Gwet's AC showed high concordance for 'difficulties with travelling to and from school' but moderate for 'making friends' and 'more homework'. All other items showed only 'fair' levels of agreement. The proportion of parent-child dyads that had disagreement ranged from 14% to 36%.

The source of disagreement varied depending on the item. Of parent-child dyads with disagreement, children reported more frequently experiencing difficulties with 'missing friends' (61%) and 'more homework' (62%) than parents. On the contrary, parents reported that children more frequently experienced difficulties with 'demanding schoolwork' (64%) and a 'larger school' (67%) than their child did.

3. Family cohesion

Agreement between parents and children on the measure 'ability of the family to get along' was only moderate with a Gwet's AC of 0.53 (Table 7). This analysis used data from K cohort children aged 16–17 and in 65% of dyads the parent and child gave different responses from each other. Similar proportions of parents and children over-reported problems with family members not getting along, relative to the other informant. Among dyads with disagreement, in 54% the child reported the family had a better ability to get along than what was reported by the parent. The parent reported a better ability to get along in the remaining 46% of cases.

Table 7: Agreement general health, prescribed medicines, family cohesion

	N parent-child dyads	Prevalence of disagreement		Gwet's AC	Discordant parent-child dyads	
		n	%		% Child over-report[a]	% Parent over-report[a]
General health						
General health at age 14–15	3,302	1,817	55	0.67	71	29
General health at age 16–17	2,906	1,561	54	0.67	70	30
Prescribed medicine at age 16–17	2,896	234	8	0.87	38	62
Ability of family to get along age 16–17	2,845	1,841	65	0.53	54	46

Notes: [a] Percentage out of the total disagreement. AC: agreement coefficient. Dark green: 'almost perfect' agreement; light green: substantial agreement; blue: moderate agreement; yellow: fair agreement; peach: poor/slight agreement.

Source: General health and prescribed medicine: K cohort, Waves 6 and 7. Ability of family to get along: K cohort, Wave 7

4. Use of prescribed medicine and general health

Agreement between parents and children on general health and use of prescribed medicine was substantial or higher. Both parents and study children from the K cohort assessed children's general health at ages 14–15 and 16–17 on a 5-point scale from 'excellent' to 'poor'. The agreement between parents' and children's reports was 'substantial' with a Gwet's AC of 0.67 (Table 7). However, despite that result, the prevalence of disagreement was relatively high, at 55% at age 14–15 and 54% at 16–17. This means that in over half of parent-child dyads, the parent gave a different response from the child. In around 70% of these cases the child reported their health to be better than what their parent indicated.

At age 16–17, parents and study children reported on the child's need for and use of medicines prescribed by a doctor. The agreement in this item was 'almost perfect' according to Gwet's AC (Table 7). From a sample of 2,896 dyads, 92% had concordance between the parent's and child's response. Among the dyads where there was disagreement (*N* = 234) 62% had parents reporting use of prescribed medicine where children did not.

5. PedsQL Physical Health scale

Concordance on the PedsQL Physical Health scale score was examined using data from K cohort collected when children were aged 14–15 and 16–17.

Agreement PedsQL Physical Health score

Disagreement between scores derived from parents' and children's responses was defined as a difference greater than 0.5 standard deviations of the child's score. Parent under-reporting was identified where there was disagreement and the score based on parent responses was lower than that calculated from child responses. Over-reporting occurred where there was disagreement, and the parent score was higher than the child score.

Parent and child scores agreed in 46.7% of dyads when children were aged 14–15 and in 44.5% when children were 16–17. This was considered low by ICC values of 0.26 and 0.29 (Table 8), indicating low reliability. Effect sizes at both age points was small to medium (values of 0.47 and 0.48), which translates to some meaningful difference between parent and child scores on average.

Correlation between the difference of scores and mean scores for each parent-child dyad was positive, suggesting that the magnitude of disagreement varied across the quality-of-life scale. Bland-Altman plots (provided in the Appendix) showed that at both ages, agreement between parents' and children's scores was better when the average score of the two was higher (indicating better physical health); or, conversely, disagreement was larger at lower average levels of physical functioning.

Table 8: Measures of agreement PedsQL Physical Health scale score

	Age 14–15	Age 16–17
ICC	0.26 [CI 0.23,0.29]	0.29 [CI 0.26,0.33]
Absolute mean difference	14.17 (*sd* 15.99)	13.33 (*sd* 13.91)
Correlation difference/mean	0.38	0.19
Directional mean difference (P1 – child)	-7.61 (*sd* 19.97)	-7.37 (*sd* 17.80)
Effect size	0.47	0.48
95% LOA	-46.74, 31.53	-42.26, 27.52
N	3,255	2,846

Notes: ICC: Intraclass correlation coefficient. CI: Confidence interval. *sd*: standard deviation. LOA: limits of agreement.
Source: LSAC, K cohort, Waves 6–7

Parental over-reporting was noted in 15.5% of dyads when children were aged 14–15, and parent under-reporting in 37.8%. When children were aged 16–17, 13.4% had parents over-reporting and 42.1% had parent under-reporting. The directional mean difference at both ages was negative (Table 8), confirming that scores calculated on parent responses tended to indicate lower physical health quality of life on average, compared to what was suggested by scores calculated from child responses. RMA plots (given in the Appendix) showed the reduced major axis line deviated from the line of perfect concordance at both ages. It was below the line of perfect concordance and had a steeper gradient, confirming scores from parent responses were on average lower but, additionally, suggesting deviance was less where children had higher scores.

Factors associated with parent over-report of PedsQL Physical Health

Results from three logistic regression models for evaluating parent over-report compared to agreement on the PedsQL Physical Health score are in Table 9. All models were fitted to data from the K cohort; the first used data collected when they were aged 14–15, the second used data collected at age 16–17 and the third model pooled data from ages 14–17. The outcome measure in each model was a binary indicator of whether the dyad was identified as having parent over-report or agreement (as defined above).

Across all three models, factors associated with an increased likelihood of higher scores calculated from parent responses rather than agreement included child age and sex, trust and communication as reported by the child, and children's mental health. Results from the combined age model showed that the probability of parent over-report (compared to agreement) was lower when children were aged 16–17 than 14–15 (OR = 0.77, 95% CI [0.65, 0.93]) and male (OR = 0.69, 95% CI [0.56, 0.84]) but higher in single-parent rather than two-parent households (OR = 1.31, 95% CI [1.02, 1.69]). Note the finding regarding single-parent families was not observed in the separate models for each age group; this could be explained by greater statistical power in the combined age model and the frequency of single-parent families within the combined age group sample compared to the separate age samples.

Scores calculated from parent responses were also likely to be higher on average, rather than agree with those calculated from child responses, where children had reported lower levels of trust and communication with parents (OR = 2.64, 95% CI [2.06, 3.39]), and when children had elevated depressive symptoms (OR = 2.41, 95% CI [1.95, 2.97]). Higher scores indicated better physical health quality of life. These results suggest parents in such circumstances may not be aware of their children's physical health difficulties, as perceived by the child.

Table 9: Factors associated with parent over-report compared to agreement on PedsQL Physical Health scale score

	Age 14–15		Age 16–17		Pooled ages 14–17	
	OR	95% CI	OR	95% CI	OR	95% CI
Demographic characteristics						
Age (ref. 14–15)						
Age 16–17	-	-	-	-	0.77**	[0.65, 0.93]
Sex (ref. female)						
Male	0.78*	[0.62, 0.96]	0.65***	[0.51, 0.83]	0.69***	[0.56, 0.84]
Family structure (ref. two-parent family)						
Single-parent family	1.25	[0.94, 1.66]	1.25	[0.92, 1.70]	1.31*	[1.02, 1.69]
Parent country of birth (ref. at least one parent born Australia)						
At least one parent born English-speaking country (not Australia)	1.23	[0.90, 1.66]	1.13	[0.78, 1.65]	1.22	[0.92, 1.63]
Parents born in other countries	0.79	[0.58, 1.08]	1.08	[0.78, 1.49]	0.89	[0.68, 1.16]
Child has medical condition or disability (ref. yes)						
No	1.00	[0.54, 1.86]	0.59	[0.33, 1.04]	0.77	[0.47, 1.27]
Parental education (ref. no university degree)						
University degree or higher	0.83	[0.67, 1.03]	0.96	[0.75, 1.22]	0.87	[0.72, 1.06]
Number of siblings (ref. only child)						
One	1.01	[0.71, 1.44]	1.19	[0.82, 1.72]	1.11	[0.83, 1.50]
Two or more	1.11	[0.78, 1.58]	1.39	[0.96, 2.02]	1.30	[0.96, 1.76]
Quality of relationships						
Trust/communication with parents (ref. very high)						
Moderate/high	1.88***	[1.45, 2.44]	1.54**	[1.14, 2.10]	1.87***	[1.48, 2.35]
Lower	2.16***	[1.62, 2.88]	2.62***	[1.93, 3.55]	2.64***	[2.06, 3.39]
Frequency child argues with parents (ref. not at all)						
A little	1.05	[0.81, 1.36]	0.77	[0.58, 1.02]	0.91	[0.73, 1.14]
Sometimes/more often	0.88	[0.65, 1.19]	0.85	[0.59, 1.22]	0.85	[0.64, 1.12]
Mental health						
Child depressive symptoms scale (ref. below cut off)						
Above cut off	2.31***	[1.82, 2.94]	2.07***	[1.60, 2.68]	2.41***	[1.95, 2.97]
Parent psychological distress score	0.99	[0.96, 1.02]	1.02	[0.99, 1.06]	1.01	[0.98, 1.04]
Random effects						
lnsig2u	-	-	-	-	0.99	[0.57, 1.73]
Observations	2,020		1,635		3,655	

Notes: 95% confidence intervals in brackets. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. Panel analysis logistic regression with robust standard errors accounting for clustering. Ref: reference category.

Source: LSAC, K cohort, Waves 6 and 7

Factors associated with parent under-report of PedsQL Physical Health

Results from three logistic regression models for evaluating parent under-report on the PedsQL Physical Health score are in Table 10. The samples used were the same as for the models for assessing over-reporting detailed above, namely the K cohort at the two separate ages of 14–15 and 16–17 and pooled for ages 14–17. The outcome measure in each model was a binary indicator of parent under-report or agreement (as defined earlier).

Results from the pooled model showed that the probability of parent under-report, relative to agreement, was higher where children were aged 16–17 rather than 14–15 (OR = 1.21, 95% CI [1.06, 1.38]); families had both parents born in non-English speaking countries (OR = 1.87, 95% CI [1.53, 2.28]); the child had a medical condition or disability (OR = 0.56, 95% CI [0.39, 0.81]); parents did not have university level education (OR = 0.76, 95% CI [0.65, 0.88]); where children reported more frequent arguments with their parents (OR = 2.33, 95% CI [1.89, 2.87]) and among parents with higher levels of psychological distress (OR = 1.08, 95% CI [1.06, 1.11]). The pooled model also showed the probability was higher in single-parent families (OR = 1.24, 95% CI [1.02, 1.51]) but this was not found in the separate models for each age group. This is likely explained by the increased sample size and power in the pooled model.

Table 10: Factors associated with parent under-report compared to agreement on PedsQL Physical Health scale score

	Age 14–15		Age 16–17		Pooled ages 14–17	
	OR	95% CI	OR	95% CI	OR	95% CI
Demographic characteristics						
Age (ref. 14–15)						
Age 16–17	-	-	-	-	1.21**	[1.06, 1.38]
Sex (ref. female)						
Male	0.91	[0.78, 1.06]	1.05	[0.89, 1.24]	0.97	[0.84, 1.13]
Family structure (ref. two-parent family)						
Single-parent family	1.18	[0.95, 1.47]	1.11	[0.89, 1.39]	1.24*	[1.02, 1.51]
Parent country of birth (ref. at least one parent born Australia)						
At least one parent born English-speaking country (not Australia)	1.00	[0.78, 1.28]	1.22	[0.95, 1.58]	1.11	[0.89, 1.40]
Parents born in other countries	1.61***	[1.32, 1.97]	1.73***	[1.40, 2.14]	1.87***	[1.53, 2.28]
Child has medical condition or disability (ref. yes)						
No	0.55**	[0.36, 0.83]	0.63*	[0.42, 0.95]	0.56**	[0.39, 0.81]
Parental education (ref. no university degree)						
University degree or higher	0.74***	[0.63, 0.87]	0.86	[0.73, 1.02]	0.76***	[0.65, 0.88]
Number of siblings (ref. only child)						
One	0.92	[0.71, 1.20]	0.83	[0.65, 1.06]	0.85	[0.68, 1.07]
Two or more	0.82	[0.63, 1.07]	0.88	[0.69, 1.13]	0.82	[0.65, 1.04]
Relationship quality						
Trust/communication with parents (ref. very high)						
Moderate/high	0.91	[0.76, 1.09]	0.82	[0.68, 1.00]	0.86	[0.73, 1.02]
Lower	1.00	[0.81, 1.24]	0.94	[0.76, 1.16]	0.97	[0.80, 1.16]
Frequency child argues with parents (ref. not at all)						
A little	1.34**	[1.11, 1.63]	1.25*	[1.02, 1.52]	1.39***	[1.16, 1.66]
Sometimes/more often	1.75***	[1.40, 2.18]	2.38***	[1.88, 3.01]	2.33***	[1.89, 2.87]
Mental health						
Child depressive symptoms scale (ref. below cut off)						
Above cut off	0.89	[0.73, 1.09]	1.08	[0.90, 1.29]	0.96	[0.81, 1.13]
Parent psychological distress score	1.09***	[1.06, 1.12]	1.05***	[1.03, 1.07]	1.08***	[1.06, 1.11]
Random effects						
lnsig2u	-	-	-	-	1.24	[0.89, 1.73]
Observations	2,737		2,442		5,179	

Notes: 95% confidence intervals in brackets. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. Panel analysis logistic regression with robust standard errors accounting for clustering. Ref: reference category.

Source: LSAC, K cohort, Waves 6 and 7

6. Strengths and Difficulties Questionnaire

This section presents results relating to concordance on the SDQ. We used data collected from study children when they were aged 10–11, 12–13, 14–15 and 16–17, pooled from both the B and K cohorts. We examined the extent and direction of agreement on the score for each of the five subscales of hyperactivity, emotional symptoms, conduct problems, peer problems, and prosocial behaviour; between classification of subscale scores into average or raised/high symptoms (low for prosocial); and on the total SDQ score. Factors associated with disagreement on the total score were also examined.

Hyperactivity subscale

Concordance between parent and child scores on the hyperactivity subscale was relatively low. ICC scores decreased as children got older, ranging from 0.38 at age 10–11 to 0.30 at age 16–17 (Table 11), showing low and worsening reliability between them. The absolute mean difference in scores increased from 2.03 at age 10–11 to 2.38 at 16–17 and effect sizes increased from 0.22 to 0.74, or small to medium.

The directional mean difference in the subscale score across all ages was negative (-0.5–1.64) indicating that scores calculated on parent responses tended to be lower, on average, than those derived from child responses. Thus, parent scores indicated a lower level of difficulties compared to that indicated by scores calculated from child responses.

Table 12 shows the prevalence of disagreement, Gwet's AC and the proportion of child and parent over-reporting for the scores classified as average or raised/high. While Gwet's AC indicated reasonable agreement on this measure, the proportion of dyads with disagreement increased from 22% at age 10–11 to 27% at 16–17. This was driven by increases in child over-reporting, or higher scores being derived from child compared to parent responses. Among dyads with disagreement at age 10–11, 61% were explained by higher child compared to parent scores but by age 16–17, this had increased to 85%.

Emotional symptoms

Agreement on the emotional symptoms scale score tended to be low across ages, albeit with some improvement when children were older (Table 11). ICC between parent and child scores ranged from 0.28 at age 10–11 to 0.39 at age 16–17, indicating low concordance. The mean difference between parent and child scores was substantial, with effect sizes around 0.5; except at age 12–13, where the effect size of 0.32 indicated means were closer. The directional mean difference in scores across all ages was negative. Scores calculated on parent responses tended to be lower on average, indicating a lower level of emotional problems compared to what was suggested by scores derived from child responses.

Gwet's AC values showed substantial agreement between parents and children when scores were classified as either average or raised/higher, although prevalence of disagreement was between 20% and 23% at each age (Table 12). Among dyads where there was disagreement at ages 10–11 and 12–13, 62% and 70% had higher classification on the parent compared to child scores, respectively. In those cases, the classification given by the parent score indicated 'raised/higher' but the child score was 'average'. However, at ages 14–15 and 16–17 disagreement was more equally attributable to both parties, with 54% and 51% respectively designated as parent over-reporting.

Conduct problems

ICCs between parent and child scores on the conduct problems subscale were low, indicating low reliability. Values were 0.36 when children were age 10–11, 0.38 at age 12–13, 0.41 at 14–15 and 0.34 at 16–17 (Table 11). The mean difference at age 10–11 was 1.37, and had changed to 1.27 by age 16–17; at all ages, the directional mean difference was negative, indicating that scores calculated on parent responses were on average lower. They tended to show a lower level of conduct problems than what would be indicated by scores derived from child responses. At each age , effect sizes were small to medium but large enough to be substantively meaningful, ranging from a low of 0.31 at age 12–13 to 0.45 at age 16–17.

Concordance on the classification as average versus raised/higher was substantial or almost perfect at all ages, according to Gwet's AC (Table 12), although prevalence of disagreement ranged from 16% at ages 12–13 and 16–17 to around 20% at 10–11 and 14–15. The directional differences found on the scale score were obscured when classified in the binary measure, with disagreement not clearly attributable to either parent or child reporting. The proportion of dyads with disagreement that had child over-reporting was between 42% and 54%.

Peer problems

The peer problems subscale score showed the highest level of concordance between parents' and children's reports. The ICCs were 0.43 at ages 10–11 and 12–13, 0.45 at 14–15 and 0.42 at 16–17 (Table 11). Absolute mean differences and directional differences were small and effect sizes also minimal, particularly at ages 12–13 and 14–15 (values of 0.01 and 0.03 respectively). Effects were meaningful, albeit still small at 0.26, when children were aged 16–17. Scores calculated from child responses tended to be higher than those derived from parent responses, on average, at this age.

Gwet's AC measure showed that agreement between parents and children on the classification of peer problems as average or raised/higher levels was substantial at each age (Table 12). However, despite that, prevalence of disagreement was around 20%–23% at each point and was driven by parent over-reporting relative to children. Among dyads with disagreement, 35% had child over-reporting at age 10–11 and 38% at age 16–17, but rates at ages between those were 23%–26%.

Note that findings on the subscale classification into average or raised/higher differed from those obtained on the underlying scores, particularly at age 16–17. On the subscale score, disagreement, while minimal, was explained by scores derived from child responses being higher than those from parents on average. The opposite occurred when scores were dichotomized. On that indicator, 62% of dyads with disagreement had children designated as 'raised/higher' when parent scores were used but 'average' when child scores were used; that is, parent over-reporting relative to the child.

Table 11: Agreement on Strength and Difficulties Questionnaire subscale scores

	Age 10–11	Age 12–13	Age 14–15	Age 16–17
Hyperactivity				
ICC [95% CI]	0.38[0.36, 0.40]	0.38 [0.36, 0.39]	0.36 [0.33, 0.38]	0.30 [0.26, 0.32]
Absolute mean difference	2.03	2.07	2.17	2.38
Correlation difference/mean	0.06	-0.01	-0.07	-0.10
Directional mean difference P1 – child (*sd*)	-0.5 (2.56)	-0.73 (2.58)	-1.17 (2.52)	-1.64 (2.49)
Effect size	0.22	0.31	0.51	0.74
95% LOA	-5.51, 4.51	-5.78, 4.33	-6.12, 3.78	-6.52, 3.24
Emotional symptoms				
ICC [95% CI]	0.28 [0.26, 0.30]	0.34 [0.32, 0.36]	0.37 [0.34, 0.39]	0.39 [0.36, 0.41]
Absolute mean difference	2.03	1.82	1.94	2.12
Correlation difference/mean	-0.13	-0.10	-0.22	-0.18
Directional mean difference P1 – child (*sd*)	-1.08 (2.44)	-0.66 (2.33)	-1.06 (2.38)	-1.23 (2.51)
Effect size	0.51	0.32	0.48	0.52
95% LOA	-5.85, 3.70	-5.23, 3.90	-5.73, 3.62	-6.15, 3.70
Conduct problems				
ICC [95% CI]	0.36 [0.34, 0.38]	0.38 [0.36, 0.40]	0.41 [0.38, 0.44]	0.34 [0.31, 0.38]
Absolute mean difference	1.37	1.15	1.15	1.27
Correlation difference/mean	-0.17	-0.08	-010	-0.17
Directional mean difference P1 – child (*sd*)	-0.67 (1.76)	-0.45 (1.62)	-0.53 (1.56)	-0.67 (1.68)
Effect size	0.42	0.31	0.36	0.45
95% LOA	-4.12, 2.78	-3.62, 2.72	-3.60, 2.53	-3.96, 2.62
Peer problems				
ICC [95% CI]	0.43 [0.41, 0.45]	0.43 [0.41, 0.45]	0.45 [0.43, 0.48]	0.42 [0.39, 0.45]
Absolute mean difference	1.34	1.22	1.26	1.36
Correlation difference/mean	-0.03	0.048	-0.003	-0.03
Directional mean difference P1-child (*sd*)	-0.35 (1.80)	-0.01 (1.73)	-0.06 (1.73)	-0.44 (1.79)
Effect size	0.20	0.01	0.03	0.26
95% LOA	-3.88, 3.19	-3.39, 3.37	-3.45, 3.33	-3.95, 3.08

Table continued over page →

	Age 10–11	Age 12–13	Age 14–15	Age 16–17
Prosocial behaviours				
ICC [95% CI]	0.22 [0.19, 0.24]	0.25 [0.23, 0.27]	0.30 [0.27, 0.33]	0.28 [0.25, 0.31]
Absolute mean difference	1.61	1.65	1.65	1.62
Correlation difference/mean	-0.03	-0.01	0.05	0.02
Directional mean difference P1-child (*sd*)	0.56 (2.06)	0.55 (2.12)	0.41 (2.14)	0.44 (2.10)
Effect size	0.34	0.31	0.22	0.25
95% LOA	-3.48, 4.60	-3.59, 4.70	-3.79, 4.60	-3.67, 4.55
N	7,603–7,606	6,935–6,936	3,267–3,270	2,865–2,867

Notes: ICC: intraclass correlation. CI: confidence interval. *sd*: standard deviation. LOA: 95% Limits of Agreement. ICC were also estimated using K cohort respondents only and K cohort balanced panel. Results were similar to those presented above. See the 'Analytical Strategy' section for more details on the measures used. Sample size in each age group varied slightly due to missing values and scale as described in the Appendix. Exact values are given in Table 12.

Source: LSAC, B cohort, Waves 6 and 7, K cohort, Waves 4–7

Table 12: Agreement on Strength and Difficulties Questionnaire subscale classifications into average or raised/high

	N parent-child dyads	Prevalence of disagreement		Gwet's AC	Discordant parent-child dyads	
		n	%		% Child over-report[a]	% Parent over-report[a]
Hyperactivity subscale						
Classification at age 10–11	7,605	1,704	22	0.68	61	19
Classification at age 12–13	6,935	1,511	22	0.69	68	32
Classification at age 14–15	3,267	766	23	0.67	77	23
Classification at age 16–17	2,866	767	27	0.62	85	15
Emotional symptoms						
Classification at age 10–11	7,604	1,673	22	0.70	38	62
Classification at age 12–13	6,936	1,415	20	0.73	30	70
Classification at age 14–15	3,267	653	20	0.72	46	54
Classification at age 16–17	2,861	664	23	0.65	49	51
Conduct problems						
Classification at age 10–11	7,606	1,593	21	0.71	51	49
Classification at age 12–13	6,936	1,092	16	0.80	42	58
Classification at age 14–15	3,269	653	20	0.82	42	58
Classification at age 16–17	2,866	465	16	0.79	54	46
Peer problems						
Classification at age 10–11	7,604	1,593	21	0.70	35	65
Classification at age 12–13	6,936	1,417	20	0.72	23	77
Classification at age 14–15	2,367	466	20	0.69	26	74
Classification at age 16–17	2,866	668	23	0.66	38	62
Prosociality						
Classification at age 10–11	7,603	947	12	0.86	35	65
Classification at age 12–13	6,936	1,087	16	0.81	23	77
Classification at age 14–15	3,270	708	22	0.77	26	74
Classification at age 16–17	2,867	445	16	0.81	38	62

Notes: [a] Percentage out of the total disagreement. AC: agreement coefficient. Dark green: 'almost perfect' agreement; light green: substantial agreement; blue: moderate agreement; yellow: fair agreement; peach: poor/slight agreement.

Source: LSAC, B cohort, Waves 6 and 7, K cohort, Waves 4–7

Prosocial behaviours

ICC values indicated low levels of agreement on the prosocial behaviours scale, ranging from a minimum of 0.22 at age 10–11 to a maximum of 0.30 at age 14–15 (Table 11). However, effect sizes were relatively small and declined as children got older, from 0.34 at age 10–11 to 0.25 at age 16–17. Directional differences were positive, meaning scores derived from parent responses were higher than those derived from child responses, on average. Parent scores, therefore, tended to indicate higher prosociality than what was indicated by child scores.

Agreement was high according to Gwet's AC when parent and child scores were classified into average versus low (Table 12). Prevalence of disagreement within dyads was variable across age groupings and ranged from 12% of dyads at age 10–11 to 22% at age 14–15. Where there was disagreement, it was driven by parent over-reporting relative to children. Child under-reporting occurred in 35% of dyads with disagreement at age 10–11, was 23% at 12–13 and increased to 26% and 38% at 14–15 and 16–17 respectively.

Total difficulties score

There was relatively low agreement between total SDQ scores calculated on parent responses and those derived from child responses. Agreement tended to be lower at older ages (see Table 13 and the Appendix for RMA and Bland-Altman plots). ICC values showed a generally decreasing trend as the children got older, at 0.42 for age 10–11, 0.44 at 12–13 but 0.36 at 16–17. At all ages, values suggested low reliability and it is likely that different conclusions would be drawn from analysis depending on the source of the data used.

Scores calculated from parent responses tended to be lower than those obtained from child responses, on average; directional mean differences were negative at every age , which was reflected in the Bland-Altman plots by a negative intercept on the observed average agreement (see Appendix). The difference was greatest when children were aged 14–15 and 16–17, as shown by directional mean differences of -2.82 and -3.98 respectively. At these ages, effect sizes were medium at 0.50 and 0.70 respectively; large enough to be considered substantively meaningful and likely to exist in the population.

Agreement between parent and child total scores within dyads varied according to the extent of difficulties as well as with age. Bland-Altman plots (given in the Appendix) showed that at every age point, agreement tended to be better when the mean of the two scores was lower. That is, when the average of the two scores indicated a lower level of difficulties, scores calculated from parent responses and scores from child responses were generally closer to each other.

Table 13: Measures of agreement SDQ total difficulties score

	Age 10–11	Age 12–13	Age 14–15	Age 16–17
ICC (95% CI)	0.42 [0.30, 0.52]	0.44 [0.36, 0.50]	0.42 [0.27, 0.53]	0.36 [0.13, 0.53]
Absolute mean difference	5.03	4.65	5.05	5.70
Correlation difference/mean	-0.07	-0.03	-0.11	-0.08
Directional mean difference P1 – child (*sd*)	-2.58 (5.89)	-1.85 (5.73)	-2.82 (5.84)	-3.98 (5.92)
Effect size	0.45	0.33	0.50	0.70
95% LOA	-14.12, 8.95	-13.08, 9.38	-14.26, 8.62	-15.59, 7.63
N	7,605	6,935	3,267	2,865

Notes: ICC: intraclass correlation. CI: confidence interval. *sd*: standard deviation. LOA: 95% Limits of Agreement. ICC were also estimated using K cohort respondents only and K cohort balanced panel. Results were similar to those presented above. See the 'Analytical Strategy' section for more details on the measures used.

Source: LSAC, B cohort, Waves 6 and 7, K cohort, Waves 4–7

Factors associated with disagreement

Disagreement on total scores was identified where the total score calculated from parent responses was more than 0.5 standard deviations from the score calculated on the child responses, with agreement otherwise. Each dyad with disagreement was denoted as either parent over-report, where the score from parent responses was higher than that from child responses, or under-report, where the score from parent responses was lower. This section first examines factors associated with parent over-report compared to agreement, and second, examines factors associated with parent under-report, compared to agreement.

Analysis was conducted on four samples[2] using data pooled at three different ages, when study children were 12–13 (sourced from B and K cohort), 14–15 (K cohort only), and 16–17 (K cohort only). Then, models were fitted to each of those three age groups separately.

Factors associated with parent over-report of SDQ total difficulty

Table 14 contains results of four logistic regression models fitted to the samples described above. The outcome in each case was a binary indicator of whether a dyad was designated as parent over-report or agreement (dyads denoting parent under-report were not included in this analysis). Results indicated that socio-demographic characteristics, relationship quality and parental mental health were associated with higher scores calculated from parent responses rather than agreement with scores derived from the child's report. However, effects varied with the age of the child.

The model fitted to pooled data showed that total scores calculated on parent responses were more likely to over-report rather than agree when children were aged 12–13 and 14–15 compared to 16–17 (for reference category of age 12–13, age category 14–15 had estimated OR = 0.86, 95% CI [0.74, 1.00]; age category 16–17 had estimated OR = 0.64, 95% CI [0.53, 0.77]). Over-report was also more likely than agreement among males (OR = 1.43, 95% CI [1.23, 1.65]) and when parents were born outside of English-speaking countries (OR = 1.29, 95% CI [1.07, 1.54]). Similarly, a higher frequency of arguments with parents as reported by the study child (OR = 2.98, 95% CI [2.42, 3.66]) and higher levels of parental psychological distress (OR = 1.09, 95% CI [1.07, 1.11]) were associated with higher total scores on parent responses rather than agreement. On the other hand, total scores based on responses from parents who were university educated (OR = 0.80, 95% CI [0.69, 0.93]) and who had children with no ongoing medical conditions or disabilities (OR = 0.41, 95% CI [0.30, 0.55]) were less likely to be higher than agree with those calculated from child responses.

The models estimated on separate age groups showed that the significant factors that predicted parent over-report (rather than agreement) varied according to the age of the child when data were collected (Table 14). For example, being male was a significant factor at ages 12–13 (OR = 1.37, 95% CI [1.19, 1.57]) and 14–15 (OR = 1.50, 95% CI [1.19, 1.88]) but not 16–17 (OR = 1.11, 95% CI [0.84, 1.46]), while a single parent was only significant at age 12–13 (OR 1.37, 95% CI [1.13, 1.65]). Parental education, in turn, was significant at older ages but not 12–13 (at age 12–13 OR = 0.90, 95% CI [0.78, 1.03]; age 14–15 OR = 0.78, 95% CI [0.62, 0.98]; age 16–17 OR = 0.75, 95% CI [0.56, 0.99]).

Factors associated with parent under-report of SDQ total difficulty

Table 15 contains results from the series of models for parent under-reporting on total scores, compared to agreement. The outcome in each case was a binary indicator of whether a dyad was designated as parent under-report or agreement (dyads denoting parent over-report were not included in this analysis). The model using pooled data showed that scores based on parent responses were more likely to be lower than to agree with those based on child responses when children were aged 14–15 (OR = 1.17, 95% CI [1.04, 1.31]) and 16–17 (OR = 1.70, 95% CI [1.49, 1.94]) rather than 12–13; when children had no ongoing medical conditions or disabilities (OR = 1.37, 95% CI [1.02, 1.83]); and where the child reported lower levels of trust and communication (OR = 2.47, 95% CI [2.12, 2.88]). Higher levels of depressive symptoms (OR = 2.60, 95% CI [2.27, 2.97]) and negative social behaviours as reported by the study child (OR = 1.04, 95% CI [1.02, 1.07]) were also associated with lower scores derived from parent responses, rather than agreement.

Total difficulty scores calculated from parent responses were less likely to be lower than to agree with scores calculated from child responses where study children were male (Table 15; OR = 0.75, 95% CI [0.66, 0.83]); and children reported more frequent arguments with parents (OR = 0.51, 95% CI = [0.44, 0.60). Parental mental health was also a significant factor; scores calculated on responses from parents with higher levels of psychological distress were less likely to be lower than to agree with child scores (OR=0.96, 95% CI [0.94, 0.97]).

Most of the findings outlined above were consistent across the models fitted to different age groups, with a few exceptions. First, male sex was not a significant factor at age 12–13 (OR = 0.93, 95% CI [0.83, 1.04]) but it was at 14–15 (OR = 0.60, 95% CI [0.51, 0.71]) and 16–17 (OR=0.66, 95% CI [0.55, 0.79]). Parental education was only a significant factor at age 16–17 (OR = 0.83, 95% CI [0.69, 1.00]), while negative social behaviours were not associated with lower difficulty scores arising from parent responses rather than agreement when children were 12–13 (OR = 1.02, 95% CI [1.00, 1.04]).

2 Due to the availability of relevant covariates (i.e. child mental health and negative social behaviours), this analysis is restricted to age 12–13 onwards.

Table 14: Factors associated with over-reporting of SDQ total difficulty scores calculated on parent compared to child responses

	Pooled ages 12–17		Age 12–13		Age 14–15		Age 16–17	
	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
Demographic characteristics								
Age (ref. 12–13)								
Age 14–15	0.86	[0.74, 1.00]						
Age 16–17	0.64***	[0.53, 0.77]						
Sex (ref. female)								
Male	1.43***	[1.23, 1.65]	1.37***	[1.19, 1.57]	1.50***	[1.19, 1.88]	1.11	[0.84, 1.46]
Family structure (ref. two-parent family)								
Single-parent family	1.20	[0.99, 1.46]	1.37**	[1.13, 1.65]	1.03	[0.75, 1.41]	0.69	[0.47, 1.01]
Parent country of birth (ref. at least one parent born Australia)								
At least one parent born English-speaking country (not Australia)	1.08	[0.86, 1.35]	1.17	[0.94, 1.45]	0.83	[0.58, 1.19]	1.08	[0.69, 1.67]
Parents born in other countries	1.29**	[1.07, 1.54]	1.18	[0.99, 1.41]	1.31	[0.99, 1.73]	1.35	[0.97, 1.88]
Child has medical condition or disability (ref. yes)								
No	0.41***	[0.30, 0.55]	0.50***	[0.37, 0.67]	0.57*	[0.35, 0.94]	0.37***	[0.22, 0.63]
Parental education (ref. no university degree)								
University degree or higher	0.80**	[0.69, 0.93]	0.90	[0.78, 1.03]	0.78*	[0.62, 0.98]	0.75*	[0.56, 0.99]
Number of siblings (ref. none)								
One	1.00	[0.79, 1.28]	1.02	[0.78, 1.31]	0.98	[0.68, 1.43]	1.01	[0.68, 1.49]
Two or more	0.91	[0.71, 1.17]	0.98	[0.76, 1.27]	0.89	[0.61, 1.30]	0.80	[0.54, 1.20]
Relationship quality								
Trust/communication with parents (ref. very high)								
Moderate/high	0.78**	[0.67, 0.92]	0.85	[0.72, 1.00]	0.74*	[0.57, 0.96]	0.76	[0.54, 1.06]
Lower	0.88	[0.71, 1.08]	0.93	[0.74, 1.17]	0.89	[0.66, 1.21]	0.94	[0.66, 1.34]
Frequency child argues with parents (ref. not at all)								
A little	1.37**	[1.13, 1.65]	1.21	[1.00, 1.46]	1.64**	[1.21, 2.22]	1.26	[0.87, 1.84]
Sometimes/more often	2.98***	[2.42, 3.66]	2.16***	[1.77, 2.64]	3.17***	[2.30, 4.35]	3.24***	[2.18, 4.81]
Mental health								
Child depressive symptoms scale (ref. below cut-off)								
Above cut-off	0.83	[0.68, 1.02]	0.95	[0.76, 1.18]	0.85	[0.61, 1.19]	0.81	[0.59, 1.12]
Parent psychological distress score	1.09***	[1.07, 1.11]	1.08***	[1.06, 1.10]	1.06***	[1.03, 1.09]	1.08***	[1.05, 1.12]
Number of negative social behaviours	0.99	[0.96, 1.02]	1.00	[0.92, 1.07]	1.02	[0.95, 1.09]	0.99	[0.96, 1.02]
Random effects								
lnsig2u	1.41*	[1.04, 1.90]	-	-	-	-	-	-
Observations (person-period)	6,508		3,826		1,575		1,107	

Notes: Exponentiated coefficients; 95% confidence intervals in brackets. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.
Source: LSAC, K cohort, Wave 5 (age 12–13), Wave 6 (age 14–15) and Wave 7 (age 16–17) and B cohort, Wave 7 (age 12–13)

Table 15: Factors associated with under-reporting of SDQ total difficulty scores calculated on parent compared to agreement with child responses

	Combined ages 12–17		Age 12–13		Age 14–15		Age 16–17	
	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
Demographic characteristics								
Age (ref. 12–13)								
Age 14–15	1.17**	[1.04, 1.31]	-	-	-	-	-	-
Age 16–17	1.70***	[1.49, 1.94]	-	-	-	-	-	-
Sex (ref. female)								
Male	0.74***	[0.66, 0.83]	0.93	[0.83, 1.04]	0.60***	[0.51, 0.71]	0.66***	[0.55, 0.79]
Family structure (ref. two-parent family)								
Single-parent family	1.09	[0.94, 1.28]	1.17	[0.99, 1.38]	1.02	[0.80, 1.30]	0.91	[0.72, 1.15]
Parent country of birth (ref. at least one parent born Australia)								
At least one parent born English speaking country (not Australia)	1.05	[0.88, 1.26]	1.11	[0.93, 1.32]	1.07	[0.82, 1.38]	0.92	[0.70, 1.21]
Parents born in other countries	0.90	[0.78, 1.05]	0.89	[0.77, 1.04]	1.02	[0.82, 1.27]	0.96	[0.76, 1.20]
Child has medical condition or disability (ref. yes)								
No	1.37*	[1.02, 1.83]	1.20	[0.86, 1.68]	1.57	[0.99, 2.51]	1.17	[0.74, 1.86]
Parental education (ref. no university degree)								
University degree or higher	0.90	[0.81, 1.01]	1.00	[0.89, 1.12]	0.89	[0.75, 1.05]	0.83*	[0.69, 1.00]
Number of siblings (ref. none)								
One	0.92	[0.76, 1.10]	0.87	[0.71, 1.06]	0.89	[0.67, 1.18]	1.12	[0.86, 1.44]
Two or more	0.91	[0.76, 1.10]	0.92	[0.75, 1.13]	0.85	[0.64, 1.13]	1.01	[0.78, 1.31]
Relationship quality								
Trust/communication with parents (ref. very high)								
Moderate/high	2.01***	[1.78, 2.26]	1.98***	[1.74, 2.24]	1.75***	[1.45, 2.11]	1.47***	[1.19, 1.82]
Lower	2.47***	[2.12, 2.88]	2.40***	[2.02, 2.86]	2.08***	[1.65, 2.61]	1.67***	[1.32, 2.10]
Frequency child argues with parents (ref. not at all)								
A little	0.79***	[0.69, 0.90]	0.75***	[0.65, 0.86]	0.94	[0.78, 1.15]	0.97	[0.78, 1.19]
Sometimes/more often	0.51***	[0.44, 0.60]	0.53***	[0.45, 0.62]	0.67***	[0.53, 0.85]	0.64***	[0.50, 0.83]
Mental health								
Child depressive symptoms scale (ref. below cut-off)								
Above cut-off	2.60***	[2.27, 2.97]	2.38***	[2.03, 2.79]	2.45***	[1.97, 3.04]	1.88***	[1.54, 2.29]
Parent psychological distress score	0.96***	[0.94, 0.97]	0.96***	[0.95, 0.98]	0.96**	[0.94, 0.98]	0.97**	[0.94, 0.99]
Number of negative social behaviours	1.04***	[1.02, 1.07]	1.02	[1.00, 1.04]	1.10*	[1.02, 1.18]	1.08**	[1.03, 1.13]
Random effects								
lnsig2u	1.38**	[1.12, 1.69]						
Observations (person-period)	10,695		5,513		2,705		2,477	

Notes: Exponentiated coefficients; 95% confidence intervals in brackets. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.
Source: LSAC, K cohort, Wave 5 (age 12–13), Wave 6 (age 14–15) and Wave 7 (age 16–17) and B cohort, Wave 7 (age 12–13)

Summary and implications

Participation in paid work

The 'Introduction' explained that items on reasons why the study child participated in paid work were asked of both parents and study children because differences in perspectives could be of interest to researchers and form the basis of research questions, when the items were asked at age 14–15. However, analysis of agreement and concordance showed minimal differences in responses. In most cases, parents and children either both agreed that a given reason was relevant, or both agreed that it wasn't; there were minimal differences in perception to analyse. Where disagreement was high, namely for the reasons 'for spending money' and 'to save up for something', it was likely attributable to the reasons having similar underlying meaning and interpretation and responses being divided across the two.

Given the above, the most likely use of the participation in paid work items would be to analyse associations between work, reasons for work and other measures of interest relating to the study child, rather than differences between parent and child perceptions of that participation. In such cases, the filtering item on whether the child had done paid work would be used to identify the sample of interest. Six per cent (N = 188) of parent-child dyads had parents give a different response to this question from children. If researchers used child responses to identify those who had worked, then the selected sample would include around 115 more individuals than if parent responses were used. From the parents' point of view, these young people hadn't worked. Identifying a sample from parent responses would include 73 young people who, according to their own responses, had not worked.

Having identified the sample of individuals who worked, data on reasons for working should come from the same informant as per the filtering question. That is, if parent responses were used to identify the sample of those who worked, then parent responses to reasons for working should also be used. However, given similar underlying interpretations and high prevalence of disagreement between parents and children in the two reasons 'for spending money' and 'to save up for something', analysts should consider aggregating information from both. Practically, this might mean creating a variable for child responses that would show whether the young person indicated either reason, and then repeating for parent responses if applicable.

Difficulties changing schools

Items on difficulties associated with changing schools were asked of parents and study children at age 12–13 to enable research into their different perspectives and points of view. Among cases where the child had changed schools, agreement on whether or not they had experienced difficulties was generally high. However, parent and child responses were different in 15% of cases (N = 744). In dyads where there was disagreement, around 43% (N = 320) had the study child reporting that they had experienced difficulties but the parent reporting they had not. In the remaining pairs (57%, N = 424) the parent reported difficulties while the study child indicated none.

There is sufficient variability within parent-child dyads for future research to investigate why members might have responded differently; for example, why parents indicated their child had no difficulties in changing schools when the child's own response suggested otherwise. This could perhaps be attributable to aspects of the parent-child relationship and have consequences for the mental health and wellbeing of children where difficulties went unrecognised or unsupported. Research could be informed by examining discordance in reasons that parents and children gave for difficulties. Evidence in this paper suggested children citing missing friends and more homework, while parents citing demanding schoolwork, a larger school, and more subjects/teachers accounted for the majority of disagreement.

Family cohesion

Parents and study children at age 16–17 were both asked to report on the ability of the family to get along. Overall, agreement in responses was moderate but in 65% (N = 1,841) of parent-child dyads, the parent and child gave different responses from each other. As with participation in paid work and difficulties changing schools, the family cohesion item was asked of both parents and children as different responses could form the basis of interesting and insightful research. Results in this paper showed that in nearly half (46%) of cases where there was disagreement, the child reported the family got along less well than what had been perceived by the parent. An appropriate research project relating to that, therefore, could compare child health outcomes in families where disharmony was identified by the child but not the parent, and consider strategies for mediating or mitigating any negative impacts.

Use of prescribed medicine

Data on use of prescribed medicine was collected from both parents and study children because survey designers were unsure who might be the best person to provide the information. The item was administered to the K cohort when children were aged 16–17 and could be expected to have varying degrees of independence from their parents. Findings from analysing a sample of 2,896 parent-child dyads showed high concordance, with the majority (92%) having agreement between the parent and child. The remaining 8% (*N* = 234), however, disagreed. Disagreement was mostly attributable to the parent indicating the child currently needed or used prescribed medicine but the child indicating that they didn't; this occurred in 145 dyads, or 5% of the sample overall. Three per cent of cases had the opposite occur, where the child reported they needed or used prescribed medicine but the parent reported they didn't.

Given the above, the question of who the best informant might be depends on tolerance for error and the availability of other data sources to cross-check survey results with. If only parent data were collected and used, then prevalence of need or use of prescribed medicine in the sample would be around 2% higher compared to if child responses were collected and used. Conversely, if only child data were used, prevalence estimates would be approximately 2% lower than if parent data were used. Users could opt to access information on prescription medicine use available in the Pharmaceutical Benefits Schedule (see growingupinaustralia.gov.au/data-and-documentation/lsac-data-linkage-administration) to help inform their decisions.

General health

The general health measure was administered to study children and their parents at ages 14–15 and 16–17 to enable comparisons between the two and analysis of differences in reporting of perceived health during mid-adolescence. This paper showed that while concordance was assessed as substantial, the prevalence of disagreement is high enough at each age group (54%–55%) to support future research. Around 3,300 parent-child dyads were examined at age 14–15 and 2,900 at age 16–17. At both ages, 38%–39% of children reported their health to be better than what was indicated by their parent, while the parent reported better child health than the child in 16% of cases. Research questions in this field could, for example, examine the extent of disagreement between parents and children, whether the discrepancy occurs more often among families with certain characteristics, and the consequences of regarding health differently on service access and use.

PedsQL Physical Health Scale

The physical health subscale of the PedsQL measures overall levels of physical health and was designed to collect the same information from both parents and children. This paper analysed scale scores derived using parent and child responses from the K cohort collected when children were aged 14–15 and 16–17. Results showed scores differed between parents and children, in line with prior research (e.g. Cremeens, Eiser & Blades, 2006). There was agreement in around 47% of dyads studied at age 14–15 (*N* = 3,255) and 45% at age 16–17 (*N* = 2,846). Thus, it was slightly lower when children were older; results also showed it tended to be lower among single-parent families. This may reflect constrained resources of single parents in cases where communication or time spent with the child was limited, be related to single parents' mental health and wellbeing, or indeed be reflective of older children independently managing their health, thus limiting involvement of the single parent.

Generally, where there was disagreement, scores calculated on parent responses tended to indicate lower physical health compared to what was suggested by scores calculated from child responses, consistent with previous studies (e.g. Sanker et al., 2015). Such parent under-reporting occurred in 38% of dyads at age 14–15 and 42% at 16–17. Differences between scores were sufficiently high as to be meaningful and could lead to different conclusions depending on which set of responses was used. The probability of parent under-reporting varied according to parent and child characteristics. A sample that had a high proportion of children with a medical condition or disability, and with children that reported more frequent arguments with their parents, would likely contain a relatively high proportion of dyads where parent scores indicated lower physical health than what would be indicated from child scores. This could also happen if the sample contained a high proportion of parents who did not have a university level education or who were born in non-English speaking countries.

Researchers are advised to carefully examine the composition of their analytic sample and be mindful of the above when deciding which of parent and child responses to use from the PedsQL physical health scale, as well as when interpreting results and conclusions from their analyses.

Strengths and Difficulties Questionnaire

The SDQ was intended to be a multi-informant scale and designed to collect data from both parents and children. In LSAC, it was collected when study children were aged 10–11, 12–13, 14–15 and 16–17. Analysis in this paper was conducted on pooled data from both the B and K cohorts. Concordance was assessed on scores for each of the five subscales of hyperactivity, emotional symptoms, conduct problems, peer problems and prosocial behaviour; classifications of the subscales scores into average or raised/high symptoms (low for prosocial), and on the total SDQ score. Factors associated with disagreement on the total score were also examined. Results are summarised below. When taken together, they suggest that parents' perspectives taken in isolation could under-estimate the social-emotional difficulties reported by children themselves (e.g. Michels et al., 2013; Van Roy, Groholt, Heyerdahl, & Clench-Aas, 2010). Consistent with findings from Gray and colleagues (2021), this paper shows LSAC participants endorse a level of difficulties not identified by their parents. Analysts are recommended to use both parent and child responses in a multi-informant approach to mitigate any effect of underestimating social-emotional difficulties.

Subscale scores

Results showed that, for each of the SDQ subscale scores, different conclusions could be drawn depending on whether parent or child responses were used. Concordance was relatively low across all five domains. However, while differences between scores were assessed as substantively meaningful, they would nevertheless be small in the population. On average, scores calculated from child responses indicated slightly higher levels of hyperactivity, emotional symptoms and conduct problems and lower prosociality at each age compared to what was indicated by parent scores. There was a clear trend on the hyperactivity subscale in that the difference between scores increased as children got older. In contrast, on the prosociality scale, the discrepancy between parent and child scores decreased as children got older.

Concordance was highest on the peer problems scale. A meaningful, albeit small, effect was observed only when children were aged 16–17. At that age , scores calculated from child responses were higher, on average, than those derived from parent responses, indicating a higher level of problems than what would be suggested from parent scores.

Subscale classifications

Prevalence of disagreement on the categorisation of subscale scores into average versus raised/high (low for prosocial) ranged from 12% to 27% across all domains and ages. While concordance was assessed as substantial or higher by Gwet's AC, the prevalence of disagreement was sufficiently high and showed trends with age that could impact on analysis. The hyperactivity categorised score had a clear trend; 22% of dyads disagreed at age 10–11 and this increased to 27% at 16–17. The increase was explained by more instances of child over-reporting within dyads, where the child score indicated a raised/high level of symptoms but the parent score indicated average.

Disagreement on the peer problems and prosociality subscales was explained by parent rather than child over-reporting. Parent scores tended to indicate a raised or high level of peer problems compared to average as suggested by child scores. In the case of the prosociality subscale, disagreement was mostly attributable to parent scores indicating average prosociality while child scores suggested low.

Total score

Consistent with the findings by subscale, concordance on total scores of the SDQ was such that different results and conclusions could be obtained depending on whether parent or child data was used. Total scores tended to indicate higher levels of difficulty when calculated on child responses compared to parent responses on average; that is, difficulties tended to be under-reported when assessed using parent rather than child scores. However, scores tended to be closer to each other when the parent and child scores indicated a lower level of difficulties on average. The difference between parent versus child scores was inconsistent across age groups – it was slightly larger when children were 16–17 than when they were younger. Parent scores were also likely to be lower, rather than agree with, child scores where the child had no ongoing health condition or disability; where the child reported lower levels of trust and communication, higher levels of child depressive symptoms and negative social behaviours.

In light of these findings, analysts are advised to carefully examine the composition of the sample they are using. If it comprised a relatively high proportion of children with the above characteristics, then the total SDQ score calculated on parent responses would likely underestimate their child's social emotional difficulties compared to what would be obtained from the child's own reported data.

Conclusion

This paper gave a detailed examination of parent-child agreement on selected items and measures in LSAC. It showed that concordance was generally high on items relating to child participation in and reasons for paid work, general health, and use of prescribed medicines but lower on reasons for changing schools, and family cohesion. The prevalence of disagreement on the PedsQL Physical Health scale score and the SDQ subscale and total scores was such that different results and conclusions could be obtained from parent compared to child responses. Scores calculated from parent responses were more likely to under-estimate children's physical health quality of life and over-estimate their social-emotional wellbeing compared to what would be indicated from children's responses.

Concordance on health-related measures matters for access and use of children's health care services, with possible consequences for treatment and outcomes. This paper highlights the significance of the parent-child relationship on parents' and children's ability to assess wellbeing in a similar way. The frequency with which children argued with their parent/s affected the likelihood of over-reporting on total SDQ scores derived from parent rather than child responses and under-reporting of PedsQL Physical Health scores calculated from the same. Where arguing occurred sometimes or often, parent scores were more likely to indicate higher overall levels of psychosocial problems and lower physical health quality of life than what was suggested by child responses. Where there was no arguing, scores were more likely to agree. Levels of trust and communication between children and parents were also relevant. If trust and communication was low, scores derived from parent responses were more likely to under-estimate psychosocial difficulties and over-estimate physical health quality of life rather than agree with what was indicated by child responses. Analysts are advised to take a multi-informant approach, and consider both parent and child reported data, to obtain more accurate estimates of prevalence than what would be obtained from a single source.

References

Australian Mental Health Outcomes and Classification Network (AMHOCN). (2005). *Strengths and Difficulties Questionnaire: Training manual.* Parramatta, NSW: Australian Mental Health Outcomes and Classification Network.

Bland, J. M., & Altman, D. G. (2010). Statistical methods for assessing agreement between two methods of clinical measurement. *International Journal of Nursing Studies*, *47*(8), 931–936.

Christian, H., Mitrou, F., Cunneen, R., & Zubrick, S. R. (2020). Pets are associated with fewer peer problems and emotional symptoms, and better prosocial behavior: Findings from the longitudinal study of Australian children. *The Journal of Pediatrics*, *220*, 200–206.

Cleridou, K., Patalay, P., & Martin, P. (2017). Does parent-child agreement vary based on presenting problems? Results from a UK clinical sample. *Child and Adolescent Psychiatry and Mental Health*, *11*(1), 1–11.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cremeens, J., Eiser, C., & Blades, M. (2006). Factors influencing agreement between child self-report and parent proxy-reports on the Pediatric Quality of Life Inventory™ 4.0 (PedsQL™) Generic Core Scales. *Health and Quality of Life Outcomes*, *4*, 1–8.

Eiser, C., & Morse, R. (2001). Can parents rate their child's health-related quality of life? Results of a systematic review. *Quality of Life Research*, *10*, 347–357.

Goodman, R. (2001). Psychometric properties of the strengths and difficulties questionnaire. *Journal of the American Academy of Child & Adolescent Psychiatry*, *40*(11), 1337–1345.

Goodman, R., Ford, T., Simmons, H., Gatward, R., & Meltzer, H. (2003). Using the Strengths and Difficulties Questionnaire (SDQ) to screen for child psychiatric disorders in a community sample. *International Review of Psychiatry*, *15*, 166–172.

Gisev, N., Bell, J. S., & Chen, T. F. (2013). Interrater agreement and interrater reliability: key concepts, approaches, and applications. *Research in Social and Administrative Pharmacy*, *9*(3), 330–338.

Goodman, A., & Goodman, R. (2009). Strengths and Difficulties Questionnaire as a dimensional measure of child mental health. *Journal of the American Academy of Child & Adolescent Psychiatry*, *48*(4), 400–403. doi.org/10.1097/CHI.0b013e3181985068

Goolsby, J., Rich, B. A., Hinnant, B., Habayeb, S., Berghorst, L., de Los Reyes, A., Alvord, M. K. et al. (2018). Parent-child informant discrepancy is associated with poorer treatment outcome. *Journal of Child and Family Studies*, *27*, 1228–1241. doi.org/10.1007/s10826-017-0946-7

Gray, M., & Smart, D. (2009). *Growing Up in Australia*: The Longitudinal Study of Australian Children. A valuable new data source for economists. *Australian Economic Review*, *42*(3), 367–376.

Gray, E. J., Scott, J. G., Lawrence, D. M., & Thomas, H. J. (2021). Concordance between adolescents and parents on the Strengths and Difficulties Questionnaire: Analysis of an Australian nationally representative sample. *Australian & New Zealand Journal of Psychiatry*, *55*(11), 1058–1070.

Gwet, K. L. (2008). Computing inter-rater reliability and its variance in the presence of high agreement. *British Journal of Mathematical and Statistical Psychology*, *61*, 29-48.

Gwet, K. L. (2014). *Handbook of inter-rater reliability: The definitive guide to measuring the extent of agreement among raters (4th ed.)*. Gaithersburg, MD: Advanced Analytics.

Hogan, J. N., Garcia, A. M., Tomko, R. L., Squeglia, L. M., & Flanagan, J. C. (2023). Parent-child concordance and discordance in family violence reporting: a descriptive analysis from the Adolescent Brain Cognitive Development Study®. *Journal of Interpersonal Violence*, *38*(1–2), NP646-NP669.

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155–163.

May, T., & Williams, K. (2022). Mother and child mental health over time in children with Autism and/or ADHD in the Longitudinal Study of Australian Children. *Development and Psychopathology*, 1–11.

Michels, N., Vanaelst, B., Stoppie, E., Huybrechts, I., Bammann, K., De Henauw, S. et al. (2013). Parental and children's report of emotional problems: Agreement, explanatory factors and event-emotion correlation. *Child and Adolescent Mental Health*, 18(3), 180–186. doi: 10.1111/j.1475-3588.2012.00672.x

Mohal, J., Lansangan, C., Gasser, C., Taylor, T., Renda, J., Jessup, K., et al. (2021). Growing Up in Australia: *The Longitudinal Study of Australian Children – Data User Guide*, Release 9C1, June 2021. Melbourne: Australian Institute of Family Studies. doi:10.26193/BAA3N6

Rebholz, C. E., Chinapaw, M. J., van Stralen, M. M., Bere, E., Bringolf, B., De Bourdeaudhuij, I.et al. (2014). Agreement between parent and child report on parental practices regarding dietary, physical activity and sedentary behaviours: The ENERGY cross-sectional survey. *BMC Public Health*, *14*(1), 1–15.

Sanders, T., Parker, P. D., del Pozo-Cruz, B., Noetel, M., & Lonsdale, C. (2019). Type of screen time moderates effects on outcomes in 4013 children: Evidence from the Longitudinal Study of Australian Children. *International Journal of Behavioral Nutrition and Physical Activity*, *16*, 1–10.

Sanker, K., Yit Siew, C., Rezali, F. W., Fui Chee, W., Chu Chien, H., Mohd Shariff, Z. et al. (2015). Adolescent self-report and parent proxy-report of health-related quality of life: An analysis of validity and reliability of PedsQL[TM] 4.0 among a sample of Malaysian adolescents and their parents. *Health and Quality of Life Outcomes*, *13*(1).

Soloff, C., Lawrence, D., & Johnstone, R. (2005). Growing Up in Australia: *The Longitudinal Study of Australian Children – Technical Paper No. 1 Sample Design, Release 9C1*, May 2005. Melbourne: Australian Institute of Family Studies.

Steichen, T., & Cox, N. (1998). sg84: Concordance correlation coefficient. *Stata Technical Bulletin*, *43*, 35–39.

Van Roy, B., Groholt, B., Heyerdahl, S., & Clench-Aas, J. (2010). Understanding the discrepancies in parent-child reporting of emotional and behavioural problems: Effects of emotional and socio-demographic factors. *BMC Psychiatry*, *10*, 56–68.

Varni, J. W., & Limbers, C. A. (2009). The PedsQL 4.0 generic core scales young adult version: Feasibility, reliability and validity in a university student population. *Journal of Health Psychology*, *14*(4), 611–622.

Varni, J. W., Seid, M., & Kurtin, P. S. (2001). PedsQL 4.0: Reliability and validity of the Pediatric Quality of Life Inventory version 4.0 generic core scales in healthy and patient populations. *Medical Care*, *39*(8), 800–812.

Ware Jr, J. E., Nelson, E. C., Sherbourne, C. D., & Stewart, A. L. (1992). Preliminary tests of a 6-item general health survey: A patient application. In A. L. Stewart & J. E. Ware Jr (Eds.), *Measuring functioning and well-being: The medical outcomes study approach*. Durham NC: Duke University Press, pp. 291–303.

White-Koning, M., Arnaud, C., Dickinson, H. O., Thyen, U., Beckung, E., Fauconnier, J. et al. (2007). Determinants of child-parent agreement in quality-of-life reports: A European study of children with cerebral palsy. *Pediatrics*, *120*(4), e804-814. doi: 10.1542/peds.2006-3272

Xie, Q. (2013). *Agree or disagree? A demonstration of an alternative statistic to Cohen's kappa for measuring the extent and reliability of agreement between observers*. Unpublished manuscript. Retrieved from nces.ed.gov/FCSM/pdf/J4_Xie_2013FCSM.pdf

Yeh, M., & Weisz, J. R. (2001). Why are we here at the clinic? Parent-child disagreement on referral problems at outpatient treatment entry. *Journal of Consulting and Clinical Psychology*, 69, 1018–1025.

Youthinmind. (n.d.). *Frequently asked questions*. Retrieved from youthinmind.com/faq

Appendix

Sample distribution

Table 16: Sample distribution variables in models of agreement for SDQ and PedsQL at age 12–13

Variable	*N*/mean	%/*sd*
Socio-demographic characteristics		
Sex		
Female	3,546	48.8
Male	3,724	51.2
Total	7,270	100.0
Family structure		
Two-parent family	6,071	83.5
Single-parent family	1,199	16.5
Total	7,270	100.0
Parent country of birth		
At least one parent born Australia	5,071	69.8
At least one parent born English speaking country (not Australia)	844	11.6
Parents born in another country	1,355	18.6
Total	7,270	100.0
Child disability or medical condition (lasted or likely to last 6 months or more)		
Yes	332	4.6
No	6,878	95.4
Total	7,210	100.0
Highest level of parental education		
No university degree	3,835	52.8
Has university degree	3,430	47.2
Total	7,265	100.0
Number of siblings in the household		
None	686	9.4
One	3,224	44.3
Two or more	3,360	46.2
Total	7,270	100.0
Parent-child relationship		
Child–parent trust and communication (child report)		
Very high trust/communication	3,983	56.8
Mid/high trust/communication	1,952	27.9
Lower trust/communication	1,072	15.3
Total	7,007	100.0
Frequency parent-child disagree and fight (child report)		
Not at all	1,713	24
A little	3,480	48.8
Sometimes or more often	1,931	27.1
Total	7,124	100.0

Table continued over page →

Variable	*N*/mean	%/*sd*
Mental health and behaviour		
K6 psychological distress scale for main caregiver		
Mean score	9.14	3.71
n	7,109	
Child depressive symptoms (Short Mood & Feelings Questionnaire)		
Below cut-off	5,713	81.4
Above cut-off	1,304	18.6
Total	7,017	100.0
Sum of delinquent behaviour (any frequency)		
0	5,149	72.0
1	1,094	15.3
2	368	5.1
3	169	2.4
4	85	1.2
5	47	0.7
6	39	0.5
7	22	0.3
8	17	0.2
9	17	0.2
10+	145	2.1
Total	7,152	100.0

Notes: Sample descriptives are presented at age 12–13 as this is the baseline for the regression models of concordance. SDQ: Strengths and Difficulties questionnaire. PedsQL: Pediatric Quality of Life inventory. *sd*: standard deviation.

Source: LSAC, K cohort, Wave 5 and B cohort, Wave 7 (age 12–13)

Sample selection and missing data for education and work items

Sample for analysis of 'Any difficulties in changing schools?'

The eligible sample for analysing concordance on items capturing difficulty in changing schools was children aged 12–13 from B cohort, Wave 7 and K cohort, Wave 5 (combined *N* = 7,337). Parent/child pairs were retained in the analytic sample if:

- the parent responded to the question 'Since we last interviewed you has study child changed schools?' **and**
- the child responded to the question 'Have you changed schools since we last talked to you?' **and**
- both parent and child responded 'Yes' to their respective question.

Of the possible 7,337 parent/child pairs, 57 (0.8%) were removed because both parent and child were missing a response to the above questions (Table 17). A further 403 (5.5%) were removed due to missing one of the two responses. This left 6,877 pairs in which both parent and child had responses. Of these, 4,953 parents and children (72%) both responded 'yes' and were retained in the analytic sample, and 1,745 (25.4%) were removed because both parent and child reported that no change in school had occurred. One hundred and seventy-nine (2.6%) were removed because parent and child responses were inconsistent.

Sample for items on type of difficulty changing schools

The analytic sample for the seven individual difficulty questions was selected from the 4,953 parent and child pairs identified above. Within each pair, the parent and child were asked to indicate whether there had been 'any difficulties changing school'. Three hundred and eighty pairs (7.7%) were retained in the sample because both parent and child replied 'Yes'. No difficulties were reported by 3,829 pairs (77.3%). The remaining 744 pairs (15%) were removed due to inconsistent responses between parents and children. Of the 380 pairs in the analytic sample, all parents and children responded to each of the seven questions on type of difficulty experienced.

Sample for analysis of 'Has child worked?'

The eligible sample for analysing concordance on reasons for paid work was K cohort children aged 14–15 from Wave 6 (*N* = 3,537). Parent/child pairs were retained in the analytic sample if:

- the parent responded to the question 'In the last 12 months has child worked? (Excludes unpaid work experience, voluntary work, household chores for pocket money)' **and**
- the child responded to the question 'In the last 12 months have you worked?' **and**
- both parent and child responded 'Yes' to their respective question.

Of the possible 3,537 parent/child pairs, 70 (2%) were removed because both parent and child were missing a response to the above questions (Table 17). Just under 5% (*N* = 164, 4.6%) were removed due to missing one of the parent or child responses. This left 3,303 pairs in which both parent and child had responses. Of those, 1,837 (55.6%) were removed because both parent and child reported no work, and a further 188 (5.7%) were removed because responses were inconsistent. This left 1,278 parent/child pairs (38.7%) where both responded 'yes' to the child having worked. Finally, one pair was removed from this group, where the child reported that their reason for working was to 'supplement family income'. That response option was not included in the equivalent question asked to parents. This gave a final analytic sample of *N* = 1,277.

Missing data for health and family cohesion items

For analysis of the general health status variable, 6.7% of pairings were removed from respondents at the age 14–15 survey, and 17% from the 16–17 survey, due to missingness in either or both of the parent and child items (Table 17). Just over 5% of respondents were removed for analysis of the item indicating use of prescription medicine.

Missing data for PedsQL Physical Health items

Patterns of missing data and non-response were consistent across items in the PedsQL Physical Health scale (Table 17). For each item surveyed at ages 14–15 and 16–17, approximately 8% of parent/child pairings were removed due to not having either or both of the parent or child response.

Missing data in SDQ

Table 18 contains details of missing data within parent/child pairs for each SDQ item, by age group. The majority of cases removed were deleted due to one of the parent or child not responding, rather than non-response of both; furthermore, child non-response accounted for a higher proportion of cases deleted than parent non-response. Just over 95% of the eligible sample of children aged 10–11 were analysed for each item. Approximately 3% were removed due to non-response from one of the parent or child, and 1% removed due to missing both parent and child responses. Of the eligible sample of children aged 12–13, just under 5% were removed from the analysis of each item because of missing one of the either parent or child response, and 0.7% were removed due to missing both. From the sample of children aged 14–15, 5% were missing a response from one of the pair across all items, and 2.5% were missing both responses. Finally, around 7% of the eligible sample of children aged 16–17 years were missing one of the parent or child response across all items, with less than half a per cent missing both.

Table 17: Missing data and non-response education, work, health and family cohesion

Variable	Age 12–13 (Combined B and K cohorts, *N* = 7,337)						Age 14–15 (K cohort, *N* = 3,537)						Age 16–17 (K cohort, *N* = 3,089)					
	Have response for SC and P1		Missing both SC and P1		Missing one of SC or P1		Have response for SC and P1		Missing both SC and P1		Missing one of SC or P1		Have response for SC and P1		Missing both SC and P1		Missing one of SC or P1	
	n	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%
Education and work																		
Changed schools	6,877	93.7	57	0.8	403	5.5												
Worked in last 12 months							3,303	93.4	70	2.0	164	4.6						
Health																		
General health status							3,303	93.4	69	2.0	165	4.7	2,561	82.9	184	6.0	340	11.0
Uses prescribed medicine													2,907	94.1	19	0.6	163	5.3
Family cohesion																		
Ability of family to get along													2,845	92.1	14	0.5	230	7.4
PedsQL Physical Health scale																		
Difficulty walking							3,254	92.0	92	2.6	191	5.4	2,847	92.2	27	0.8	215	7.0
Difficulty running							3,255	92.0	92	2.6	190	5.4	2,844	92.1	27	0.8	218	7.1
Difficulty playing sports or exercising							3,256	92.0	92	2.6	189	5.3	2,848	92.1	25	0.8	216	7.0
Difficulty lifting something heavy							3,253	92.0	92	2.6	192	5.4	2,845	92.1	27	0.8	217	7.0
Difficulty having a bath or shower							3,255	92.0	92	2.6	190	5.4	2,846	92.1	27	0.8	216	7.0
Difficulty helping around the house							3,253	92.0	92	2.6	192	5.4	2,844	92.1	27	0.8	218	7.1
Gets aches and pains							3,255	92.0	92	2.6	190	5.4	2,847	92.2	27	0.8	215	7.0
Low energy							3,255	92.0	92	2.6	190	5.4	2,846	92.1	27	0.8	216	7.0

Source: LSAC, K cohort, Waves 5, 6 and 7 (ages 12–13 to 16–17) and B cohort, Wave 7 (age 12–13)

Table 18: Patterns of missing data for SDQ items by age group

Variable	Age 10–11 (Combined B and K cohorts, *N* = 7,933)						Age 12–13 (Combined B, K cohorts, *N* = 7,337)						Age 14–15 (K cohort, *N* = 3,537)						Age 16–17 (K cohort, *N* = 3,089)					
	Have response for SC and P1		Missing both SC and P1		Missing one of SC or P1		Have response for SC and P1		Missing both SC and P1		Missing one of SC or P1		Have response for SC and P1		Missing both SC and P1		Missing one of SC or P1		Have response for SC and P1		Missing both SC and P1		Missing one of SC or P1	
	n	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%
Hyperactivity																								
Not been able to stay still	7,601	95.8	79	1.0	253	3.2	6,938	94.6	49	0.7	350	4.8	3,268	92.4	90	2.5	179	5.1	2,865	92.7	10	0.3	214	6.9
Constantly fidgeting, etc.	7,598	95.8	79	1.0	256	3.2	6,926	94.4	49	0.7	362	4.9	3,266	92.3	90	2.5	181	5.1	2,859	92.6	11	0.4	219	7.1
Easily distracted	7,599	95.8	79	1.0	255	3.2	6,936	94.5	49	0.7	352	4.8	3,266	92.3	90	2.5	181	5.1	2,865	92.7	10	0.3	214	6.9
Stopped to think before acting	7,602	95.8	79	1.0	252	3.2	6,933	94.5	50	0.7	354	4.8	3,266	92.3	90	2.5	181	5.1	2,868	92.8	10	0.3	211	6.8
Has a good attention span	7,593	95.7	79	1.0	261	3.3	6,927	94.4	50	0.7	360	4.9	3,268	92.4	90	2.5	179	5.1	2,866	92.8	11	0.4	212	6.9
Emotional symptoms																								
Complained of headaches, etc.	7,600	95.8	79	1.0	254	3.2	6,936	94.5	49	0.7	352	4.8	3,269	92.4	90	2.5	178	5.0	2,866	92.8	10	0.3	213	6.9
Often seems worried	7,602	95.8	79	1.0	252	3.2	6,934	94.5	49	0.7	354	4.8	3,269	92.4	90	2.5	178	5.0	2,866	92.8	11	0.4	212	6.9
Often been unhappy or tearful	7,598	95.8	79	1.0	256	3.2	6,936	94.5	49	0.7	352	4.8	3,264	92.3	90	2.5	183	5.2	2,864	92.7	10	0.3	215	7.0
Nervous or easily loses confidence	7597	95.8	80	1.0	256	3.2	6,934	94.5	50	0.7	353	4.8	3,266	92.3	90	2.5	181	5.1	2,865	92.7	11	0.4	213	6.9
Had many fears	7603	95.8	79	1.0	251	3.2	6,932	94.5	50	0.7	355	4.8	3,267	92.4	90	2.5	180	5.1	2,866	92.8	10	0.3	213	6.9
Conduct problems																								
Temper	7,599	95.8	79	1.0	255	3.2	6,937	94.5	49	0.7	351	4.8	3,268	92.4	90	2.5	179	5.1	2,866	92.8	10	0.3	213	6.9
Obeys requests	7,602	95.8	79	1.0	252	3.2	6,938	94.6	49	0.7	350	4.8	3,270	92.5	90	2.5	177	5.0	2,867	92.8	11	0.4	211	6.8

Table continued over page →

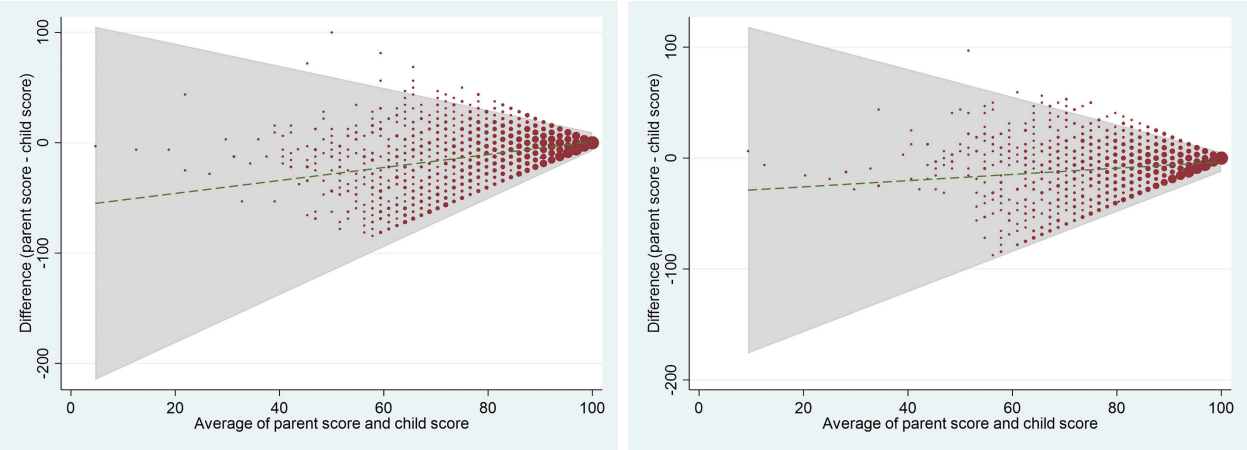
Variable	Age 10–11 (Combined B and K cohorts, *N* = 7,933)						Age 12–13 (Combined B, K cohorts, *N* = 7,337)						Age 14–15 (K cohort, *N* = 3,537)						Age 16–17 (K cohort, *N* = 3,089)					
	Have response for SC and P1		Missing both SC and P1		Missing one of SC or P1		Have response for SC and P1		Missing both SC and P1		Missing one of SC or P1		Have response for SC and P1		Missing both SC and P1		Missing one of SC or P1		Have response for SC and P1		Missing both SC and P1		Missing one of SC or P1	
	n	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%
Often fights/ bullies children	7,602	95.8	79	1.0	252	3.2	6,937	94.5	49	0.7	351	4.8	3,267	92.4	90	2.5	180	5.1	2,867	92.8	10	0.3	212	6.9
Often lies or cheats	7,599	95.8	79	1.0	255	3.2	6,934	94.5	50	0.7	353	4.8	3,267	92.4	90	2.5	180	5.1	2,867	92.8	10	0.3	212	6.9
Steals	7,600	95.8	79	1.0	254	3.2	6,932	94.5	50	0.7	355	4.8	3,266	92.3	90	2.5	181	5.1	2,865	92.7	10	0.3	214	6.9
Peer problems																								
Has been solitary	7,596	95.8	79	1.0	258	3.3	6,934	94.5	49	0.7	354	4.8	3,268	92.4	90	2.5	179	5.1	2,864	92.7	10	0.3	215	7.0
Has at least one good friend	7,604	95.9	79	1.0	250	3.2	6,937	94.5	49	0.7	351	4.8	3,268	92.4	90	2.5	179	5.1	2,868	92.8	10	0.3	211	6.8
Liked by other children	7,602	95.8	79	1.0	252	3.2	6,933	94.5	49	0.7	355	4.8	3,265	92.3	90	2.5	182	5.1	2,866	92.8	10	0.3	213	6.9
Picked on/ bullied by children	7,596	95.8	79	1.0	258	3.3	6,934	94.5	50	0.7	353	4.8	3,266	92.3	90	2.5	181	5.1	2,864	92.7	10	0.3	215	7.0
Gets on better with adults	7,600	95.8	79	1.0	254	3.2	6,933	94.5	50	0.7	354	4.8	3,267	92.4	90	2.5	180	5.1	2,865	92.7	11	0.4	213	6.9
Prosociality																								
Considerate of other's feelings	7,600	95.8	79	1.0	254	3.2	6,938	94.6	49	0.7	350	4.8	3,268	92.4	90	2.5	179	5.1	2,866	92.8	11	0.4	212	6.9
Readily shared with children	7,602	95.8	79	1.0	252	3.2	6,936	94.5	49	0.7	352	4.8	3,269	92.4	90	2.5	178	5.0	2,865	92.7	10	0.3	214	6.9
Helpful if someone is hurt, etc.	7,602	95.8	79	1.0	252	3.2	6,936	94.5	49	0.7	352	4.8	3,269	92.4	90	2.5	178	5.0	2,866	92.8	11	0.4	212	6.9
Kind to younger children	7,603	95.8	79	1.0	251	3.2	6,935	94.5	50	0.7	352	4.8	3,268	92.4	90	2.5	179	5.1	2,868	92.8	10	0.3	211	6.8
Often volunteered to help	7,601	95.8	79	1.0	253	3.2	6,934	94.5	49	0.7	353	4.8	3,267	92.4	90	2.5	180	5.1	2,866	92.8	10	0.3	213	6.9

Source: LSAC, K cohort, Waves 4, 5, 6 and 7 (ages 10–11 to 16-17) and B cohort, Waves 6 and 7 (ages 10–11 to 12–13).

PedsQL Physical Health scale score supporting figures

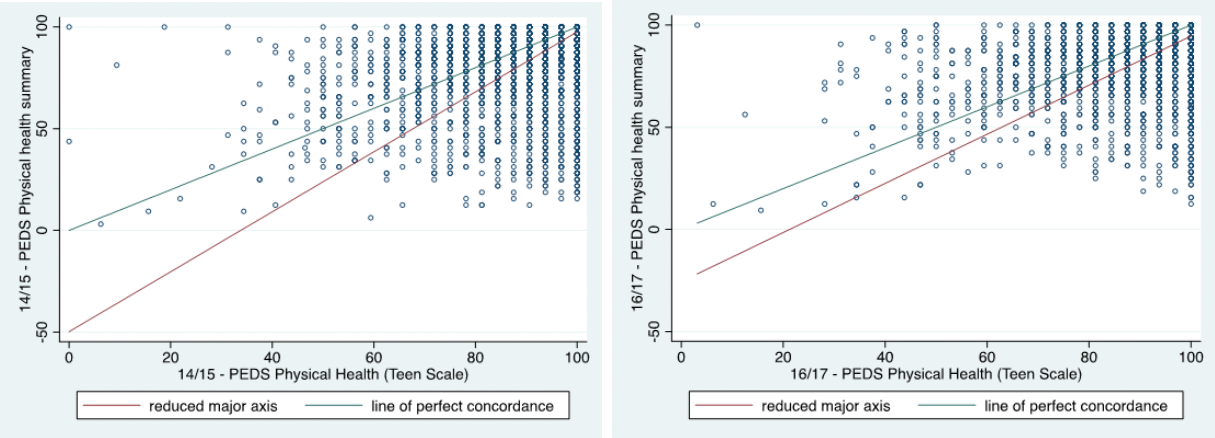
These figures plot the trends due to the non-zero correlation between the pairwise difference and mean scores:

Figure 1: Bland-Altman plots PedsQL Physical Health scale score at age 14–15 (left) and 16–17 (right)



Source: LSAC, K cohort, Waves 6–7

Figure 2: Reduced major axis plots PedsQL Physical Health scale score at age 14–15 (left) and 16–17 (right)

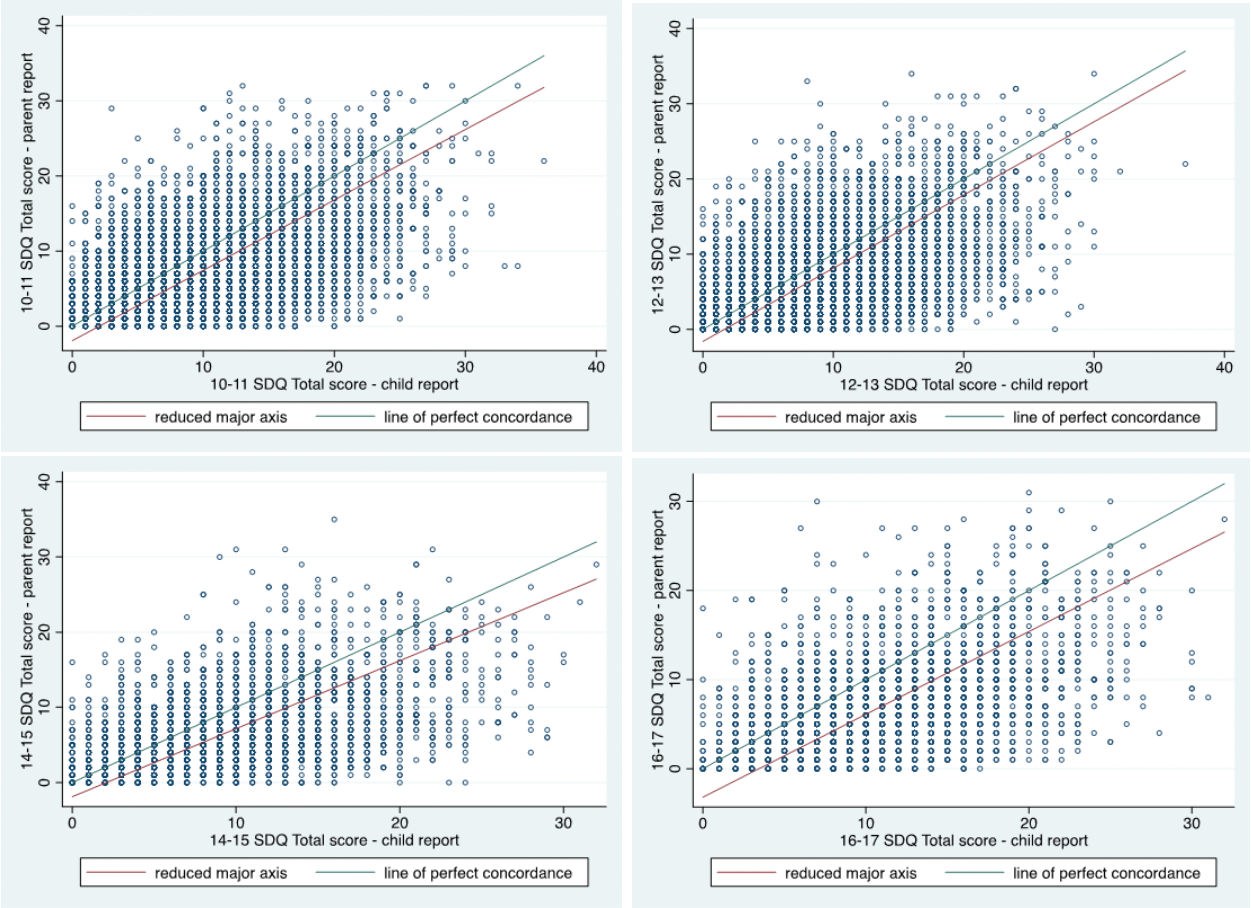


Source: LSAC, K cohort, Waves 6–7

Direction of disagreement SDQ total score – supporting figures

Figure 3 shows reduced major axis (RMA) plots for each year of age. They further illustrate the relatively low levels of agreement between total SDQ scores calculated on parent versus child responses. If there was perfect concordance, there would be no difference between parent scores and child scores, and all data points (represented by dots) would lie on the line of perfect concordance (in green). The RMA line (in red) and line of perfect concordance move further apart as the total difficulties scores increase, particularly at ages 10–11, 14–15 and 16–17. Thus, the figures suggest the disagreement between scores is larger where children have a higher level of difficulties, as previously suggested.

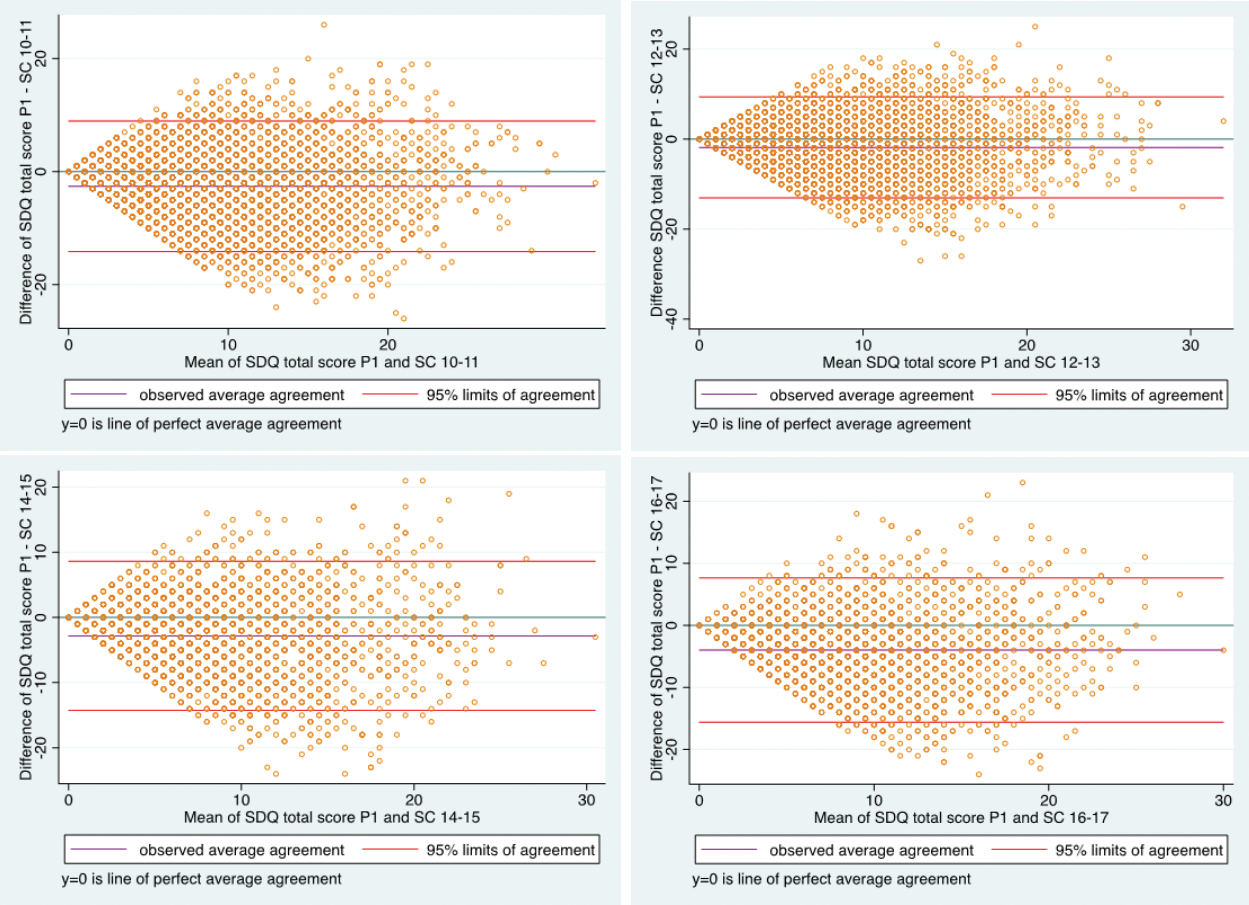
Figure 3: Reduced major axis SDQ total difficulties scores ages 10–11 to 16–17



Notes: SDQ: Strengths and difficulties score.

Source: LSAC, B cohort, Waves 6 and 7, K cohort, Waves 4–7

Figure 4: Bland-Altman plots SDQ total difficulties scores ages 10–11 to 16–17



Notes: SDQ: Strengths and difficulties score. P1: primary caregiver. SC: study child.

Source: LSAC, B cohort, Waves 6 and 7, K cohort Waves 4–7