USER GUIDE
SCALES SCREEN DATA

Courtenay Norbury, Sarah Griffiths, Debbie Gooch, Laura Lucas, and the SCALES Consortium

V3 05.01.22

Contents

Background

The SCALES screening data release contains data from the first phase of the Surrey Communication and Language in Education Study (SCALES). In this phase, Reception class teachers in state-maintained schools in the county of Surrey UK were invited to respond to an on-line questionnaire for each child in their class. The main focus of the study was to identify the prevalence of language difficulties at school entry and document the association between language difficulties, behavioural problems and educational attainment in the first year of school.

Documentation

1. The screening questionnaire instructions and items given to teachers.
2. Data dictionary that includes the list of variables and variable definitions.

The data dictionary includes a list of all variables and a description of how they were measured or derived. Once you have decided which variables you would like to include in your analysis, you can find details of each specific question in the screening questionnaire document (the variable list includes reference to the question number in the questionnaire).

Study Design

The population

The population of interest were children in reception class in state-maintained[1] schools in the County of Surrey in the UK in September 2011 (i.e. children that turned 5 between 01/09/11 and 31/08/12). The total population of children was 12,398, located in 263 schools (254 mainstream and 9 special schools) across Surrey.

The sample

The SCALES screening questionnaire was completed for 7,267 children (59% of the total population) by approximately 281 teachers in 161 state-funded schools (156 mainstream and 5 special schools). Participating schools (n = 161) did not differ from those that opted out (n = 102) on measures of socioeconomic disadvantage (percentage of children receiving free school meals, $t(261) = 1.38$, $p = .17$); children in receipt of a statement of special educational need, $t(261) = 0.19$, $p = .85$; or children speaking English as an additional language, $t(232) = 1.05$, $p = .29$. Children screened were all aged between 4 years 9 months (57 months) and 5 years 10 months (70 months). The screen was completed for 3,553 girls (49%) and 3,714 boys (51%). The sample came from a range of social economic status backgrounds as determined by Income Deprivation Affecting

[1] We initially planned to include children in the 69 independent schools in Surrey in the study. However, due to low buy in from the independent schools (16%) it was not possible to include this sub-population.

Children Index (McLennan, 2011) for children's home postcodes (mean IDACI rank = 21592, SD = 7830; 1 is the most deprived and 32482 the least deprived). 5,979 children (82%) were of white ethnic origin and 782 (11%) children had English as an Additional Language (EAL).

Ethics

Ethical approval was granted by the Ethics committee at Royal Holloway, University of London, where this study originated. All parents/caregivers of children attending schools that agreed to participate were sent information sheets about they study and given the opportunity to opt-out. Twenty parents returned opt-out forms to their child's school.

Measurements

The screening questionnaire included background questions about the child, a short version of the Children's Communication Checklist 2 (CCC-2, Bishop, 2003) to assess language, the Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997) to assess behaviour, and the Early Years Foundation Stage Profile (EYFSP) to assess educational attainment.

Background information

Teachers were asked to report each child's basic demographic information such as age, sex, ethnicity and home language. They were also asked whether the child had any developmental disabilities or sensory impairments. Finally they were asked whether they had any they had any additional learning needs and whether they were receiving any additional support (e.g. Statement of Special Educational Need).

Variables:

childid	unique SCALES ID
scales_schoolt1	unique SCALES school identifier
schooltype	type of state maintained school
idaci_quintilet1	post-code indicator of neighbourhood deprivation, reported as quintile score
q4ethnicity	child ethnicity
q5gender	child sex
age	child age at assessment, in months
q8timeknownchild	time respondent has known child
q9relationship	role of respondent
q10eal	is child learning English as an additional language
q11schoolaction	is the child receiving additional support at school
statement	does child have a statement of special educational need
q12sensoryimpairment	does child have sensory impairment
q13nophrasespeech	is child combining words into phrases or sentences
q14slt	has child ever been seen by a speech-language therapist

q15edpsych	has child ever been seen by an educational psychologist
anyconcerns	has teacher been concerned about child's development in any specified area
q17nodiagnoses	has child received clinical diagnosis from outside agency
q17asd	has child received clinical diagnosis of autism spectrum disorder
q17speechlanguagedisorder	has child received clinical diagnosis of speech or language disorder from outside agency

Children's Communication Checklist-Short (unpublished, based on Bishop, 2003).

Teacher completed a short, 13-item version of the Children's Communication Checklist (CCC-2; Bishop, 2003). The CCC-2 was designed to assess communication skills and screen for communication problems in children aged 4 to 16 years who speak in sentences, do not have permanent hearing loss, and who speak English as their first language. The short version of the CCC-2 (CCC-S) contains the 13 items that best discriminated typically developing children from peers with language impairment in a validation study (Norbury, Nash, Baird, & Bishop, 2004).

Teachers reported how often the child displayed 13 communication behaviours e.g. "When answering a question, provides just the right amount of information without being overly precise or too vague". Responses were made on a 4 point scale; less than once a week (0), at least once a week, but not every day (1), once or twice a day (2), several times (more than twice) a day (3). The data dictionary describes which type of language skill is measured by each item in CCC-S. It was not possible to include exact wording of each item in the dictionary due to test copyright.

The CCC-S provides a dimensional score of communication competence, with a maximum score of 39 indicative of greatest communication deficit. The CCC-S has high internal consistency (Cronbach's $\alpha = .87$) and correlates significantly with CCC-2 total scores (Pearson's $r(515) = .88$). If teachers endorsed 'not yet combining words into phrases or sentences', the child was recorded as having 'no phrase speech (NPS).' These children received a maximum score of 39 on the CCC-S, but no further questions were completed (as these require the child to be using sentences for communication.

For T1 we provide a z score for total ccc that has been age standardised using the LMS method

Variables:

ccc1	forgets words s/he knows, uses descriptives instead
ccc2	vague pronoun reference - uses 'he' or 'it' without making referent clear
ccc3	misses the point of jokes and puns

ccc4	leaves off past-tense/word endings
ccc5	takes in just one/two words of sentence, leading to confusion
ccc6	problems with sequencing events in a narrative
ccc7	talks about future events
ccc8	can have enjoyable conversation with child
ccc9	can produce long, complex sentences
ccc10	uses words to refer to whole category of objects
ccc11	produces all speech sounds in utterance clearly
ccc12	explains past events clearly
ccc13	provides just the right amount of information in response to a question
ccctotal	sum of all 13 items of the CCCS
z_cctotalt1	Children's Communication Checklist Short (CCCS)_age adjusted z score

Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997).

The SDQ is a short behavioural screening tool appropriate for use with children aged 3-16 years old. The SDQ consists of 5 subscales (emotional symptoms, conduct problems, hyperactivity/inattention, peer relationship problems and pro-social behaviour), each comprising 5 items. Each item is scored on a 3-point scale. Responses may be: 'somewhat true' (always scored as 1), 'not true', or 'certainly true'. Responses of 'not true' and 'certainly true' may be scored as either 0 or 2 depending on the item. In the case of missing responses, a score may be prorated providing at least 3 items have been completed. Scores on the first 4 subscales can be added together to generate a total difficulties score (max = 40).

Teachers also completed the impact supplement. The first impact question enquires if the respondent "considers the child to have a problem in any one the following areas; emotions, concentration, behaviour or being able to get along with other people". This was scored as: 0 = No difficulties, 1 = yes – minor difficulties, 2 = yes – definite difficulties, 3 = yes – severe difficulties. If the respondent reported any yes response, they are asked further questions about the chronicity (less than one month, one term, more than one term), distress, social impairment and burden to others that may result (Goodman, 1999). Items from the impact supplement are scored on a 4-point scale to assess the impact of any difficulties. The scores are as follows: 0 = not at all, 1 = only a little, 2 = quite a lot, 3 = a great deal. The items from overall distress and social impairment may be totalled to create a total impact score which may range from 0 to 10. Scores from the impact supplement may be used as continuous variables or classified as normal, borderline, or abnormal. A total impact a score of 2 or more is considered abnormal, a score of 1 is considered borderline and a score of 0 is normal for each item.

Variables:

sdq1	considerate of other people's feelings
sdq2	restless, overactive, cannot stay still for long
sdq3	often complains of headaches, stomach-aches or sickness
sdq4	shares readly with other children (treats, toys, pencils)
sdq5	often has temper tantrums or hot tempers
sdq6	rather solitary, tends to play alone
sdq7	generally obedient, unsually does what adults request
sdq8	many worries, often seems worried
sdq9	helpful if someone is hurt, upset or feeling ill
sdq10	constantly fidgeting or squirming
sdq11	has at least one good friend
sdq12	often fights with other children or bullies them
sdq13	often unhappy, down-hearted or tearful
sdq14	generally liked by other children
sdq15	easily distracted, concentration wanders
sdq16	nervous or clinicaly in new sitations, easily loses confidence
sdq17	kind to younger children
sdq18	often lies or cheats
sdq19	picked on or bullied by other children
sdq20	often volunteers to heal others (parents, teachers, other children)
sdq21	thinks things out before acting
sdq22	steals from home, school, or elsewhere
sdq23	gets on better with adults than with other children
sdq24	many fears, easily scared
sdq25	sees tasks through to the end, good attention span
sdq27	overall, do you think child has difficulties in one or more of these areas: emotions, concentration, behaviour, social
sdq28	roughly how long have difficulties been present
sdq29	do the difficulties upset the child
sdq30	do the difficulties interfere with child's everyday peer relations
sdq31	do the difficulties interfere with child's everyday classroom learning
sdq32	do the difficulties place burden on you or the class as a whole
sdqemotionalrs	summary of emotion items
sdqconductrs	summary of conduct items
sdqhyperactivityrs	summary of hyperactivity items
sdqpeerrs	summary of peer items
sdqprosocialrs	summary of prosocial items
sdqtotaldifficultiesrs	summary of emotion, conduct, hyperactivity and peer items

Early Years Foundation Stage Profile (EYFSP, www.gov.uk).

Since 2006, reception teacher have been required by the UK government to complete an EYFSP for each child in their class at the end of their reception year. At the time the SCALES cohort was in

reception (2011), the UK government was about to bring in a new early years curriculum and an amended EYFSP assessment. Teachers involved in SCALES were asked to complete the new EYFSP as part of the screening questionnaire so that we could assess children's attainment within the new framework, so that the findings were applicable to policy makers going forward. The closest government report on the version of the EYFSP used in SCALES can be found here: https://www.gov.uk/government/statistics/early-years-foundation-stage-profile-results-in-england-academic-year-2011-to-2012

The new EYFSP asked teachers to rate children's progress as Emerging, Expected or Exceeding expectation in 17 areas. The 17 areas included three areas related to *Personal, Social and Emotional Development*: (1) Self-confidence and self-awareness, (2) Managing feelings and behaviour (3) Making relationships; two areas related to *Physical Development*: (1) Moving and handling, (2) Health and Self-care; three areas of *Communication and Language Development:* (1) Listening and attention, (2) Understanding, (3) Speaking; two areas of *Literacy*: (1) Reading, (2) Writing; two areas of *Mathematics*: (1) Numbers, (2) Shape, space and measures; three areas related to *Understanding of the World*: (1) People and communities, (2) The world, (3) Technology; and two areas related to *Expressive Arts and Design*: (1) Exploring and using media and materials and (2) Being imaginative.

Children are awarded 1 point for emerging, 2 points for expected and 3 points for exceeding target expectations. Total scores range from 17-51. Children are additionally classified as having achieving a good level of development if they scored at least 2 (expected) in the 12 areas related to Personal, social and emotional development, Physical development, Communication and language development, Literacy and Mathematics.

Variables:

neweyfsp1	self-confidence/self-awareness
neweyfsp2	managing feelings and behaviour
neweyfsp3	making relationships
neweyfsp4	moving and handling
neweyfsp5	health and self-care
neweyfsp6	listening and attention
neweyfsp7	understanding
neweyfsp8	speaking
neweyfsp9	reading
neweyfsp10	writing
neweyfsp11	numbers
neweyfsp12	shape, space and measures
neweyfsp13	people and communities

neweyfsp14	the world
neweyfsp15	technology
neweyfsp16	exploring using media and material
neweyfsp17	being imaginative
neweyfsptotal	sum of all 17 early learning goals
neweyfsp_gld	sum of prime areas and literacy and maths (items included in 'good level of development' index)
neweyfsp_gldcat	categorical variable - children meet expected or exceeding on all 12 GLD items (as above)

IDACI rank (Noble et al. 2019)

Home postcodes provided by teachers were used obtain an Income Deprivation Affecting Children Index (IDACI) score to estimate children's socioeconomic status. IDACI 2010 combines a number of indicators, covering a range of economic, social and housing issues, into a single neighbourhood deprivation score for each small area in England. This allows each area to be ranked relative to one another according to their level of deprivation. As with the 2007 and 2004 Indices, the IDACI 2010 have been produced at Lower Super Output Area (SOA) level, of which there are 32,482 in the country. A relative ranking of areas, according to their level of deprivation is then provided. The SOA with a rank of 1 is the most deprived, and 32482 the least deprived on this overall measure. http://www.education.gov.uk/cgi-bin/inyourarea/idaci.pl. IDACI scores are reported in quintile bands to reduce identification risk.

Missing data

Household postcodes were unavailable for 148 children and have been replaced with the child's school postcode. One child was missing both SDQ and EYFSP scores and six were missing EYFSP scores due to teachers exiting the on-line screen before completion. The screen required a response to each individual item before teachers could progress to the next item, thus there were no further missing data. Missing data is coded as NA in the data file.

Funding

This phase of SCALES was funded primarily by Wellcome grant (WT094836AIA) to Courtenay Norbury (Principal Investigator) and Baird, Charman, Pickles and Simonoff (Co-Investigators)

SCALES consortium

The SCALES consortium includes co-investigators and post-doctoral research assistants on the SCALES project:

Professor Courtenay Frazier Norbury (UCL)

Dr Sarah Griffiths (UCL)

Dr Debbie Gooch (University of Surrey)

Professor Gillian Baird (Newcomen Centre, Guys & St Thomas's NHS Trust)

Professor Tony Charman (Kings College London)

Professor Andrew Pickles (Kings College London)

Professor Emily Simonoff (Kings College London)

We gratefully acknowledge the assistance of Surrey County Council, especially Jennifer Charters and Virginia Martin, and the staff and pupils of participating schools in contributing their time and expertise to this study.

References

Bishop, D.V.M. (2003). The Children's Communication Checklist version 2 (CCC-2) Psychological Corporation. *London, England*.

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: a research note. *Journal of Child Psychology and Psychiatry, 38*(5), 581-586.

Noble, S. et al. (2019). September 2019

Noble, S., McLennan, D., Noble, M., Plunkett, E., Gutacker, N., Silk, M. & Wright, G. (2019). The English Indices of Deprivation 2019: Research report. (www.gov.uk) Note: although does not give data for the year of SCALES screen (2011/2012), provides an overview of what IDACI measures)

Norbury, C. F., Gooch, D., Baird, G., Charman, T., Simonoff, E., & Pickles, A. (2016). Younger children experience lower levels of language competence and academic progress in the first year of school: evidence from a population study. *Journal of Child Psychology and Psychiatry*, *57*(1), 65-73.