The effect of special educational assistance in early childhood education and care on psycho-social difficulties in elementary school children

Guido Biele, PhD[1,3], Ratib Lekhal, PhD[2], Kristin R. Overgaard, PhD[3], Mari Vaage Wang, PhD[1], Ragnhild Eek Brandlistuen, PhD[1], Svein Friis, PhD[3], & Paal Zeiner, PhD[3]

[1] Norwegian Institute of Public Health

[2] Norwegian Business School

[3] Oslo University Hospital

Author Note

The authors made the following contributions. Guido Biele, PhD: Conceptualization, Formal analysis, Funding acquisition, Writing - Original Draft Preparation, Writing - Review & Editing; Ratib Lekhal, PhD: Conceptualization, Writing - Original Draft Preparation, Writing - Review & Editing; Kristin R. Overgaard, PhD: Conceptualization, Funding acquisition, Writing - Review & Editing; Mari Vaage Wang, PhD: Conceptualization, Writing - Review & Editing; Ragnhild Eek Brandlistuen, PhD: Conceptualization, Writing - Review & Editing; Svein Friis, PhD: Conceptualization, Writing - Review & Editing; Paal Zeiner, PhD: Conceptualization, Funding acquisition, Writing - Review & Editing.

Correspondence concerning this article should be addressed to Guido Biele, PhD, Marcus Thranes gate 6, 0473 Oslo, Norway. E-mail: guido.biele@fhi.no

Abstract

Three to seven percent of pre-schoolers have developmental problems or child psychiatric disorders. Randomized controlled trials (RCTs) indicate that interventions in early childhood education and care improve long-term outcomes of children from disadvantaged backgrounds. It is unknown if effects generalize beyond the well-structured context of RCTs and to children who may not have a disadvantaged background but have developmental problems or psychiatric disorders. We use data from the population-based Norwegian Mother, Father and Child Cohort Study, recruiting pregnant women from 1999 to 2008, with child follow-up from ages 6, 18, and 36 months to ages 5, 7, and 8 years. This sub-study included 2499 children with developmental problems or psychiatric disorders at age five. We investigate the effects of special educational assistance at age five on mother-reported internalizing, externalizing, and communication problems at age eight. We analyse bias due to treatment by indication with directed acyclic graphs, adjust for treatment predictors to reduce bias, and estimate effects in different patient groups and outcome domains with a hierarchical Bayesian model. In the adjusted analysis, pre-schoolers with special educational assistance had on average by 0.1 (0.03-0.16) standardised mean deviation weaker psycho-social difficulties in elementary school. Mean effect sizes varied between groups and outcomes. We estimate positive effects of educational assistance during the transition from preschool to the school years. It should therefore be considered as an intervention for pre-schoolers with developmental or behaviour problems. More research with improved measurements of treatment and outcomes is needed to identify success factors for their implementation.

Keywords: ADHD, ASD, Language difficulties, Behaviour problems, early childhood education and care, psycho-social intervention, special education, inattention, hyperactivity/impulsivity, oppositional behaviour, mood, anxiety, and communication, directed acyclic graph, hierarchical Bayesian modelling.

The effect of special educational assistance in early childhood education and care on psycho-social difficulties in elementary school children

Between three and seven percent of pre-schoolers have developmental problems or child psychiatric disorders [1, 2], which are an important risk factor for mental disorders in adulthood [3]. Efforts to promote healthy growth and development in children who struggle in the early years can accordingly improve children's long-term life opportunities [4]. This effect seems to decrease as children grow older. Therefore, investing resources later, at the age of school entry or beyond, may show less of an effect [5].

Interventions in early childhood are often described as an effective method to improve the long-term outcomes of children from disadvantaged backgrounds [6] or those with specific developmental or behavioural problems like attention deficit hyperactivity disorder, autism, or behaviour or language problems [7]. Interventions in early childhood education and care (ECEC) can be especially effective because in contrast to parental training programs, their implementation relies less on parents' abilities or motivation, and on average 93% of three to five year old children in Organisation for Economic Co-operation and Development (OECD) countries are enrolled in ECEC [8]. Randomized controlled trials (RCTs) reported clear effects of early interventions in ECEC for a horizon of up to nine months, for instance for language problems [9], children with ADHD or autism [10], and for teacher classroom management programs [11].

However, the effect sizes of such interventions are not generally large, and less is known about their effect when interventions are provided outside the well-structured context of RCTs. Even though RCTs are, due to their interval validity, the gold standard for estimating treatment effects, differences between study sample and target population and differences in treatment-implementation between study and regular care contexts, make a generalization of findings from RCT samples to populations of interest difficult [12–15]. Since RCTs often take place in a controlled setting, it may be difficult to replicate the results in other, less rigid settings. For instance, field professionals in ECEC institutions will draw on a much wider range of sources than formal experimental evidence in order inform their actions. Thus, while evidence from RCTs is encouraging, it remains unclear how it generalizes to interventions in ECEC provided in regular care.

Only a handful of studies examined the effects of special educational assistance (SEA) interventions in ECEC when they are implemented outside of RCTs. These studies used propensity scores to deal with the problematic internal validity in observational studies—due to treatment by indication—and found that children who received SEA in ECEC showed the same or worse outcomes compared children who did not receive SEA [16, 17].

The Norwegian ECEC-system facilitates the investigation of SEA, because children who cannot fully benefit from standard education and care have the right to receive free SEA. Similar to other OECD countries [2], around 4.5% of pre-schoolers in Norwegian ECEC have impaired functioning, the most common impairment being language and communication difficulties, followed by psycho-social difficulties and behaviour problems [18]. Around 2.6% of pre-schoolers receive SEA, which is provided for several hours per week for individual children [18]. After stimulation of language development, social- and behaviour-training are the most frequent types of SEA provided. To date, no study has–to the best of our knowledge–examined the effect of SEA in ECEC on children's psycho-social difficulties. Related studies on SEA in Norwegian schools report that students who received it have similar or slightly worse scholastic outcomes compared to those who did not receive it [19, 20, see also 21].

In sum, the few studies examining effects of SEA in ECEC outside the context of RCTs reported small negative, to no effects of SEA. Moreover, most studies focused on educational outcomes, such that the effect of SEA on the development of psycho-social difficulties remains largely unclear. Hence, this large-scale prospective cohort study adds to the existing literature by investigating how SEA in ECEC provided outside RCTs affects the psycho-social development of children with developmental or behavioural problems.

Methods

Participants

The sample is a sub-sample of the Norwegian Mother, Father and Child Cohort Study (MoBa), a prospective population-based pregnancy cohort study conducted by the Norwegian Institute of Public Health [22, 23]. Participating mothers from all over Norway were recruited during routine ultrasound assessment in week 17 or 18 of their pregnancy in the period from 1999 to 2009. 41% of the invited women consented to participate. MoBa participants received questionnaires in gestational week 17 or 18, week 22 and week 30, at child's age 6 and 18 months, 3, 5, and 8 years and onward. The study is still on-going. The reported analyses also use information from the Medical Birth Registry of Norway [24]. Figure 1 shows the inclusion-flowchart.

The study sample is comprised of children whose mothers indicated developmental or behaviour problems in MoBa's age five years questionnaire, and for whom information about outcomes in the age eight years questionnaire are available. This study focuses on children with one or more of the following developmental or behavioural problems: Attention deficit hyperactivity disorder, language development, oppositional defiant or conduct disorder, autism spectrum disorder, and learning disabilities.

Materials

The current study used rating scales from MoBa questionnaires sent out at child ages five, and eight years. Exposure and inclusion criteria were based on responses in the five year questionnaire, whereas outcome measures were taken from the eight year questionnaire. The first, 1.5 and three year MoBa questionnaires and the Medical Birth Registry of Norway provided covariates.

Exposure. To measure the provision of SEA, we relied on following question: "Does your child receive, or has received any extra resources in the kindergarten?" If mothers responded "Yes" to this question, they were additionally asked about the number of hours per week. SEA is provided to individual children, both inside and outside the context of regular preschool activities.

Outcome variables. Outcome variables (PSD_8 in Figure 2) were *sum scores* from different scales about psycho-social difficulties. Outcome dimensions were attentional, hyperactivity/impulsivity, and behavioural (ODD or CD) problems measured with the Parent Rating Scale for Disruptive Behaviour Disorders (RS-DBD, [25]), emotional problems measured with the Short Mood and Feelings Questionnaire (SMFQ, [26]) and the Screen for Child Anxiety Related Disorders (SCARED, [27]), and communication problems measured with the Children's Communication Checklist-2 (CCC-2, [28]).

Adjustment variables. Adjustment variables and those to control for loss to follow up were chosen based on the directed acyclic graph (DAG) shown in figure Figure 2. One important set of confounders includes children's psycho-social difficulties at baseline, because these can be seen as causes of treatment and are related to later psycho-social difficulties. A number of scales in MoBa assessed psycho-social difficulties at age five and served as baseline measures (PSD_5 in Figure 2). These included the Conners' Parent Rating Scale-Revised, Short Form (CPRS-R (S), [29]), Child Behaviour Checklist (CBCL, [30]), the Ages and Stages Questionnaire (ASQ, [31]), and the Children's Communication Checklist-2.

Additional variables used for adjustment or prediction of loss to follow-up included maternal age, education, and mental health (ADHD symptoms measured with the Adult ADHD Self-Report Scale [32] at child age three and depressive symptoms measured with the SCL-5 [33] at child age five, parity, preterm birth, birth-month, hours special education per week, number of developmental of behaviour problems, and contact with rehabilitation services, Child and Adolescent Psychiatric Units, or Educational and Psychological Counseling Service at child age five years.

Classification into groups with different developmental or behavioural problems

To classify if and in which area a child had developmental or behavioural problem (DBP), we used MoBa questions about mental health problems at age five. Mothers were asked if their child "suffered, or is currently suffering from any of the following long-term illnesses or health problems." In addition, mothers were asked if they had been in contact with a Child and Adolescent Psychiatric Unit or the Educational Psychology Counseling Services and if the health problem was confirmed by a professional. Only children for whom mothers reported a health problem *and* who indicated that the problem was evaluated by a mental health professional were included in the sample.

Disorders or health problems for which MoBa's age 5 questionnaire has questions included Epilepsy, Cerebral Palsy, impaired hearing, which were excluded from the current analysis, together with children for whom mothers indicated a chromosomal defect. MoBa also asked mothers about autism spectrum disorders (ASD), hyperactivity and attention problems (ADHD), language difficulties (Lang), and behavioural problems (Beh). Additional questions about learning disabilities (LD) were also used to identify cases of interest for this study. Each child was classified in one of the following DBP groups: 1. ASD, 2. LD, 3. ADHD & Beh & Lang, 4. ADHD & Beh, 5. ADHD & Lang, 6. ADHD, 7. Lang, 8. Beh. For some children, mothers indicated multiple DBP, in which case the child was assigned to the first group it fell into. If, for example, a mother indicated ASD, ADHD, and language problems, the child was assigned to the ASD group (details in supplementary materials and Table S1). The rational underlying this classification scheme was to use existing psychiatric diagnoses, and to classify children according to their most impairing problem because these have typically more severe and persistent effects psycho-social on psycho-social development.

Data analysis

All analyses were performed using R [34]. The Bayesian hierarchical regression model was implemented with the brms package [35]. The analyses are described in more detail in the supplementary material, and analysis scripts are available at https://github.com/gbiele/SPS358.

Bias from treatment by indication and loss to follow up. Estimation of treatment effects from observational data is difficult because treatment is not assigned randomly. Instead, individuals with more psycho-social difficulties at age five, who are also more likely to have psycho-social difficulties in the future, more likely receive treatment (treatment by indication). In addition, loss to follow up makes estimation of treatment effects difficult. Therefore, we used a directed acyclic graph [DAG, 36, see Figure 2] to explicate the assumed causal structure and to determine with which approach to deal with potential biases. Given this structural model, inverse probability of continued participation weighting was needed to reduce bias from loss to follow up [37], whereas adjustment for predictors of SEA was sufficient to control bias from treatment by indication. This means that we effectively estimated the effect of SEA on the change of psycho-social difficulties from preschool to elementary school.

Estimation of the treatment effects. We used a Bayesian adjusted and weighted hierarchical ordinal regression to estimate effects of SEA [35, 38, 39]. A hierarchical regression induces partial pooling (shrinkage) of estimates, which reduces the variance of estimates [40] and controls the multiple comparison problem [41]. Importantly, when analysing related patient groups hierarchical regression results in more accurate association estimates then independent analysis of these groups [40]. We used an ordinal regression model, because the estimation of latent, normally distributed traits that underlie the rating-scale responses facilitates the presentation of results in terms of standardised mean differences (SMD). The reported results were obtained by pooling over the independent analyses of the 50 imputed data sets [42]. Consistent with recent recommendations to focus on estimation of effect sizes instead of significance testing [43, 44] we generally report means and the 90% credible intervals.

Results

The study sample includes 2499 participants (c.f., Figure 1). Thirty-three percent of the children in the sample received SEA. Table 1 describes the study sample. Figures S4 and S5 show that children with more severe problems (e.g. ASD) were more likely to receive SEA and also received SEA from better educated personnel.

Inverse probability weights reduced the differences in mean values for covariates between participants followed up and those lost to follow up to less than 0.1 SMD (c.f. Figure S1; [45]). Cumulative distribution plots showed that weighting balanced the entire distributions of covariates (Figures S7 and S8).

Effects of special educational assistance

Consistent with the structural model shown in Figure 2, the analysis without adjustment showed that SEA at age five was associated with more psycho-social difficulties at age eight (c.f. Table S3 and Figure S7). Table S4 and Figures S9 and S10, S11, and S12 show coefficients of the adjusted regression model, which indicates that after adjustment for confounders SEA was associated with less psycho-social difficulties at age eight.

Over all psycho-social outcomes and groups of developmental or behaviour problems the estimated average treatment effect (ATE) was a symptom reduction by 0.10 standardised mean deviations (SMD) (Credible Interval CI: 0.04, 0.16). Figure 3 shows that the 90% credible interval is for all groups above 0. The pairwise comparisons of all groups did not show clear differences in the estimated treatment effects between groups (c.f. Table S5 and Figure S14)

Figure 3 and Table 2 also show estimated effect sizes stratified by outcomes and indicate that SEA had a positive effect on all measured psycho-social outcomes. While there were some differences in the effect size estimates for different outcomes, in particular smaller effects for anxiety and communication problems, pairwise comparisons did not show reliable differences between them (c.f. Table S6 and Figure S15). Effect size estimates did not vary substantially by the child sex (c.f. Figure S18).

Discussion

This research used observational data from a longitudinal population based cohort study to investigate the effect of special educational assistance (SEA) in ECEC on psycho-social difficulties of children with developmental or behaviour problems. We found that, after adjustment for treatment indicators, mothers of children who received SEA in kindergarten reported fewer psycho-social difficulties three years later, compared to mothers whose children did not receive SEA.

While there was some variation in the extent of the positive effect of SEA between groups and different psycho-social difficulties, these differences were not reliably different from zero (c.f. Figures S14 and S15). Because the credible intervals for these differences are large compared to the magnitude of the estimated overall effect and the random effects standard deviations are clearly non-zero (S4), these results do not exclude the possibility of group differences. Instead, they might reflect difficulties in reliably measuring exposure, covariates, and outcomes based on parent reports only. Still, the available data were sufficient to reveal an overall positive effect of SEA.

While the positive effect reported in this study is consistent with the results of randomized controlled trials [10, 11] and with reports of the positive effects of preschool child care quality [46], it also stands in contrast to previous observational studies, which estimated no or a small negative "effects" of special education. This apparent contradiction can be due to a number of differences between the current and previous studies. We had estimates of pre-treatment difficulties, and could estimate effects of special education on the change of psycho-social difficulties. Moreover, we used adjustment for treatment predictors instead of propensity score weighting. Adjustment is the preferable approach if treatment-predictors are not colliders on a backdoor path from outcome to treatment and if the sample size is large enough to allow for inclusion of many of adjustment variables. Another important difference is that whereas previous studies focused on scholastic outcomes, we focused on the effect on psycho-social difficulties. This is a to date little examined but important outcome of SEA, because early psycho-social difficulties are associate with impaired functioning in adulthood [3]. Interestingly, the clear results of SEA on externalizing behaviour suggests that in addition to helping children with DBP, it can also benefit their families by reducing disruptive behaviour.

The estimated effect size for the reduction of psycho-social difficulties is with on average 0.10 standardised mean difference small. In comparison, previous meta analysis about school– or ECEC–based interventions found effect sizes of between -.3 and 1.3 SMD for children with or at risk for ADHD [47, 48] or SMD between 0.3 and 1.1 for children with autism [49]. Randomized trials of classroom management training for kindergarten teachers showed effect sizes similar to our results [Cohen's d around 0.3 for high risk children at the nine-months follow up, 50]. It is possible that the small effect sizes we estimated are, in addition to above mentioned measurement problems, due to the fact the SEA was often provided by personnel with limited training, especially for children with typically less severe problems (c.f. Figure S5). More generally, the decentralized organization of the Educational and Psychological Counselling Service is likely to lead to a large variation in the implementation of SEA [51]. MoBa did not collect more detailed data about SEA, which could help to elucidate when it is most effective. Another possible explanation is that the composition of the study sample, which over-represents well-educated families compared to the population [37], leads to an underestimation of the true effect size, because well-educated parents could reduce children's psycho-social difficulties even without SEA [52].

While the current study showed that mothers report fewer psycho-social difficulties in elementary school when their children received SEA in ECEC, a causal interpretation of this result as reflecting an effect of SEA rests on a number of assumptions encoded in Figure 2. One un-testable assumption is that there are no unmeasured confounders that predict both which children receive SEA and their developmental pathway. Even though the reported analysis includes obvious confounders, other unobserved confounder could still account for some of the positive association of SEA and psycho-social development. However, because RCTs of SEA and similar interventions typically report positive effects, and thus confirm a causal role of SEA, it appears unlikely that the effects estimated in this study are primarily due to confounding.

Given that the current study does not have more precise and detailed measures, future studies that assess outcomes through blinded raters or objective instruments and measure quality and quantity of the treatment more thoroughly are needed. Studies that better assess variations in treatment quality and content and use more representative samples will be useful to investigate reasons for the relatively small effects observed in the current study, and to identify criteria for effective interventions in ECEC.

Conclusion

Previous RCTs about special educational assistance and teacher management programs showed that interventions in ECEC have a positive immediate impact for children with developmental or behavioural problems, but provide little guidance on long-term effects. The current study has due to its observational character a lower internal validity than RCTs, but complements them in terms of external validity and by examining long-term effects. It thus strengthens the view that interventions in ECEC are a useful approach to support pre-schoolers with developmental or behavioural problems.

In sum, the current study suggests that the psycho-social development of children with developmental or behaviour problems can be modified in a positive way through interventions in ECEC, also when provided outside the structured context of randomized controlled trials. Future research with better measurements and more representative samples should investigate under which conditions such interventions are most effective.

Key points and relevance

- Parent training programs are considered a key component of early interventions against the development of developmental or mental health problems
- Special educational assistances in early childhood education and care (SEA in ECEC) showed positive effects on later outcomes in RCTs, but population based cohort studies reported no or even negative associations
- Results from our large, population based cohort study indicate that SEA in ECEC is associated with reduced psycho-social difficulties in elementary school
- SEA in ECEC may be effective also when implemented outside the structured context of RCTs and for children who do not come from disadvantaged backgrounds
- Easy access to SEA in ECEC may be a component of early intervention strategies to prevent or mitigate development of psycho-social difficulties in pre-schoolers at risk.

References

1.

Wichstrøm L, Berg-Nielsen TS, Angold A, et al (2012) Prevalence of psychiatric disorders in preschoolers. Journal of child psychology and psychiatry, and allied disciplines 53:695–705

2.

Global Research on Developmental Disabilities Collaborators (2018) Developmental disabilities among children younger than 5 years in 195 countries and territories, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. The Lancet Global health 6:e1100–e1121

3.

Copeland WE, Wolke D, Shanahan L, Costello EJ (2015) Adult Functional Outcomes of Common Childhood Psychiatric Problems: A Prospective, Longitudinal Study. JAMA psychiatry 72:892–899

4.

Knudsen EI, Heckman JJ, Cameron JL, Shonkoff JP (2006) Economic, neurobiological, and behavioral perspectives on building America's future workforce. Proceedings of the National Academy of Sciences of the United States of America 103:10155–10162

5.

Heckman JJ (2006) Skill formation and the economics of investing in disadvantaged children. Science 312:1900–1902

6.

Reynolds AJ, Temple JA, Robertson DL, Mann EA (2001) Long-term effects of an early childhood intervention on educational achievement and juvenile arrest: A 15-year follow-up of low-income children in public schools. JAMA: the journal of the American Medical Association 285:2339–2346

7.

Perrin EC, Sheldrick RC, McMenamy JM, et al (2014) Improving parenting skills for families of young children in pediatric settings: a randomized clinical trial. JAMA pediatrics 168:16–24

8.

OECD (2019) Education at a Glance 2019: OECD Indicators. OECD Publishing, Paris

9.

Hagen ÅM, Melby-Lervåg M, Lervåg A (2017) Improving language comprehension in preschool children with language difficulties: a cluster randomized trial. Journal of child psychology and psychiatry, and allied disciplines 58:1132–1140

10.

Zwaigenbaum L, Bauman ML, Choueiri R, et al (2015) Early Intervention for Children With Autism Spectrum Disorder Under 3 Years of Age: Recommendations for Practice and Research. Pediatrics 136 Suppl 1:S60–81

11.

Reinke WM, Herman KC, Dong N (2018) The Incredible Years Teacher Classroom Management Program: Outcomes from a Group Randomized Trial. Prevention science: the official journal of the Society for Prevention Research 19:1043–1054

12.

Balzer LB (2017) "All Generalizations Are Dangerous, Even This One."-Alexandre Dumas. Epidemiology 28:562–566

13.

Cole SR, Stuart EA (2010) Generalizing evidence from randomized clinical trials to target populations: The ACTG 320 trial. American journal of epidemiology 172:107–115

14.

Huitfeldt A, Stensrud MJ (2018) Re: Generalizing Study Results: A Potential Outcomes Perspective. Epidemiology 29:e13–e14

15.

Lesko CR, Buchanan AL, Westreich D, et al (2017) Generalizing Study Results: A Potential Outcomes Perspective. Epidemiology 28:553–561

16.

Sullivan AL, Field S (2013) Do preschool special education services make a difference in kindergarten reading and mathematics skills?: A propensity score weighting analysis. Journal of school psychology 51:243–260

17.

Dempsey I, Valentine M, Colyvas K (2016) The Effects of Special Education Support on Young Australian School Students. International Journal of Disability, Development and Education 63:271–292

18.

Wendelborg C, Caspersen J, Kittelsaa AM, et al (2015) Barnehagetilbudet til barn med særlige behov. NTNU Samfunnsforskning

19.

Lekhal R (2018) Does special education predict students' math and language skills? European journal of special needs education 33:525–540

20.

Kvande MN, Bjørklund O, Lydersen S, et al (2018) Effects of special education on academic achievement and task motivation: a propensity-score and fixed-effects approach. European journal of special needs education 1–15

21.

Morgan PL, Frisco M, Farkas G, Hibel J (2010) A Propensity Score Matching Analysis of the Effects of Special Education Services. The Journal of special education 43:236–254

22.

Magnus P, Birke C, Vejrup K, et al (2016) Cohort profile update: The norwegian mother and child cohort study (MoBa). International journal of epidemiology

23.

Magnus P, Irgens LM, Haug K, et al (2006) Cohort profile: The norwegian mother and child cohort study (MoBa). International journal of epidemiology 35:1146–1150

24.

Irgens LM (2000) The medical birth registry of norway. Epidemiological research and surveillance throughout 30 years. Acta obstetricia et gynecologica Scandinavica 79:435–439

25.

Silva RR, Alpert M, Pouget E, et al (2005) A rating scale for disruptive behavior disorders, based on the DSM-IV item pool. The Psychiatric quarterly 76:327–339

26.

Sharp C, Goodyer IM, Croudace TJ (2006) The Short Mood and Feelings Questionnaire (SMFQ): A Unidimensional Item Response Theory and Categorical Data Factor Analysis of Self-Report Ratings from a Community Sample of 7-through 11-Year-Old Children. Journal of abnormal child psychology 34:365–377

27.

Birmaher B, Brent DA, Chiappetta L, et al (1999) Psychometric properties of the Screen for Child Anxiety Related Emotional Disorders (SCARED): a replication study. Journal of the American Academy of Child and Adolescent Psychiatry 38:1230–1236

28.

Norbury CF, Nash M, Baird G, Bishop D (2004) Using a parental checklist to identify diagnostic groups in children with communication impairment: a validation of the Children's Communication Checklist–2. International journal of language & communication disorders / Royal College of Speech & Language Therapists 39:345–364

29.

Conners CK, Sitarenios G, Parker JDA, Epstein JN (1998) The Revised Conners' Parent Rating Scale (CPRS-R): Factor Structure, Reliability, and Criterion Validity. Journal of abnormal child psychology 26:257–268

30.

Nøvik TS (1999) Validity of the child behaviour checklist in a norwegian sample. European child & adolescent psychiatry 8:247–254

31.

Achenbach TM, Ruffle TM (2000) The Child Behavior Checklist and related forms for assessing behavioral/emotional problems and competencies. Pediatrics in review / American Academy of Pediatrics 21:265–271

32.

Kessler RC, Adler L, Ames M, et al (2005) The world health organization adult ADHD Self-Report scale (ASRS): A short screening scale for use in the general population. Psychological medicine 35:245–256

33.

Strand BH, Dalgard OS, Tambs K, Rognerud M (2003) Measuring the mental health status of the norwegian population: A comparison of the instruments SCL-25, SCL-10, SCL-5 and MHI-5 (SF-36). Nordic journal of psychiatry 57:113–118

34.

R Core Team (2017) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria

35.

Bürkner P-C (2017) brms: An R package for Bayesian multilevel models using Stan. Journal of Statistical Software 80:1–28. https://doi.org/10.18637/jss.v080.i01

36.

Greenland S, Pearl J, Robins JM (1999) Causal diagrams for epidemiologic research. Epidemiology 10:37–48

37.

Biele G, Gustavson K, Czajkowski NO, et al (2019) Bias from self selection and loss to follow-up in prospective cohort studies. European journal of epidemiology 34:927–938

38.

Stan Development Team (2018) RStan: The R interface to Stan

39.

Carpenter B, Gelman A, Hoffman M, et al (2017) Stan: A probabilistic programming language. Journal of statistical software 76:1–32

40.

Greenland S (2000) Principles of multilevel modelling. International journal of epidemiology 29:158–167

41.

Gelman A, Hill J, Yajima M (2012) Why We (Usually) Don't Have to Worry About Multiple Comparisons. Journal of research on educational effectiveness 5:189–211

42.

Buuren S van, Groothuis-Oudshoorn K (2011) Mice: Multivariate imputation by chained equations in r. Journal of Statistical Software 45:1–67

43.

Sullivan GM, Feinn R (2012) Using Effect Size-or Why the P Value Is Not Enough. Journal of graduate medical education 4:279–282

44.

Wasserstein RL, Schirm AL, Lazar NA (2019) Moving to a World Beyond "$p < 0.05$". The American statistician 73:1–19

45.

Austin PC, Stuart EA (2015) Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. Statistics in medicine 34:3661–3679

46.

Vandell DL, Belsky J, Burchinal M, et al (2010) Do effects of early child care extend to age 15 years? Results from the NICHD study of early child care and youth development. Child development 81:737–756

47.

Richardson M, Moore DA, Gwernan-Jones R, et al (2015) Non-pharmacological interventions for attention-deficit/hyperactivity disorder (ADHD) delivered in school settings: Systematic reviews of quantitative and qualitative research. Health technology assessment 19:1–470

48.

Gaastra GF, Groen Y, Tucha L, Tucha O (2016) The Effects of Classroom Interventions on Off-Task and Disruptive Classroom Behavior in Children with Symptoms of Attention-Deficit/Hyperactivity Disorder: A Meta-Analytic Review. PloS one 11:e0148841

49.

Reichow B (2012) Overview of meta-analyses on early intensive behavioral intervention for young children with autism spectrum disorders. Journal of autism and developmental disorders 42:512–520

50.

Fossum S, Handegård BH, Britt Drugli M (2017) The Incredible Years Teacher Classroom Management Programme in Kindergartens: Effects of a Universal Preventive Effort. Journal of child and family studies 26:2215–2223

51.

Nilsen S, Herlofsen C (2012) National Regulations and Guidelines and the Local Follow-up in the Chain of Actions in Special Education. International journal of special education 27:136–147

52.

Russell AE, Ford T, Russell G (2015) Socioeconomic Associations with ADHD: Findings from a Mediation Analysis. PloS one 10:e0128248

Table 1

Study sample

Variable	w/o SEA	with SEA	Total
Special educational assistance (SEA)			
boy	1063 (63.8%)	586 (70.3%)	1649 (66%)
girl	602 (36.2%)	248 (29.7%)	850 (34%)
Hours	0 (0, 0)	4.76 (1, 6)	1.59 (0, 1)
Developmental or behaviour problem (DBP) group			
ASD	11 (0.7%)	32 (3.8%)	43 (1.7%)
LD	19 (1.1%)	63 (7.6%)	82 (3.3%)
ADHD & Beh & Lang	12 (0.7%)	19 (2.3%)	31 (1.2%)
ADHD & Lang	58 (3.5%)	85 (10.2%)	143 (5.7%)
ADHD & Beh	108 (6.5%)	38 (4.6%)	146 (5.8%)
ADHD	330 (19.8%)	71 (8.5%)	401 (16%)
Lang	847 (50.9%)	486 (58.3%)	1333 (53.3%)
Beh	280 (16.8%)	40 (4.8%)	320 (12.8%)
Psycho-social difficulties (PSD) at child age five			
Attention	6.03 (2, 9)	6.98 (2, 10)	6.34 (2, 9)
Hyperactivity	4.67 (3, 6)	4.68 (3, 6)	4.67 (3, 6)
Externalizing (CBCL)	3.98 (2, 6)	3.73 (1, 6)	3.9 (2, 6)
Internalizing (CBCL)	2.01 (0, 3)	2.16 (0, 3)	2.06 (0, 3)
Communication (CCC)	3.93 (2, 6)	4.76 (3, 7)	4.21 (2, 6)
Development (ASQ)	1.34 (0, 2)	2.31 (1, 3)	1.67 (0, 2)
Psycho-social difficulties (PSD) at child age eight			
Attention (ATT, RS-DBD)	7.51 (4, 10)	8.3 (4, 12)	7.77 (4, 11)
Hyperactivity (HYP, RS-DBD)	6.07 (2, 9)	5.77 (1, 8)	5.97 (2, 9)
Oppositional (OPP, RS-DBD)	5.18 (2, 7)	4.44 (1, 6)	4.93 (2, 7)
Mood (MOOD, SMFQ)	3.06 (1, 4)	2.96 (1, 4.75)	3.03 (1, 4)
Anxiety (ANX, SCARED)	1.21 (0, 2)	1.22 (0, 2)	1.21 (0, 2)
Communication (COMM, CCC)	7.75 (4, 11)	10.29 (5, 14)	8.6 (4, 12)
Education (years)	14.01 (12, 15)	13.98 (12, 16)	14 (12, 15)
Maternal characteristics			
Age (yars)	30.52 (27, 34)	30.82 (28, 34)	30.62 (28, 34)
ADHD (ADHD-RS)	7.38 (5, 10)	7.15 (5, 9)	7.3 (5, 10)
Depression (SCL-5)	2.53 (0, 4)	2.43 (0, 3)	2.5 (0, 3)

ASD = Autisum spectrum disorder, LD = Learning difficulties, Lang = Language problems, Beh = behaviour problems, SEA = special educational assistance. Abbreviations and original scales for PSD are given in parentheses (see methods section for full names). Numbers in parentheses are percent or first and third quartiles.

Table 2

Estimated average treatment effects (ATE) stratified by groups with different developmental and behavioural problems (rows) and psycho-social difficulties (columns).

	ATT	HYP	OPP	MOOD	ANX	COMM	Average
ASD	0.08 (-0.06,0.21)	0.11 (-0.02,0.25)	0.11 (-0.02,0.25)	0.1 (-0.03,0.24)	0.08 (-0.06,0.21)	0.07 (-0.08,0.2)	0.09 (0, 0.18)
LD	0.09 (-0.02,0.21)	0.11 (0,0.23)	0.11 (0,0.23)	0.11 (0,0.23)	0.1 (-0.01,0.22)	0.08 (-0.05,0.19)	0.1 (0.03, 0.18)
ADHD & Beh & Lang	0.1 (-0.04,0.24)	0.11 (-0.03,0.26)	0.09 (-0.05,0.23)	0.11 (-0.03,0.25)	0.1 (-0.04,0.25)	0.08 (-0.07,0.22)	0.1 (0, 0.19)
ADHD & Lang	0.07 (-0.05,0.18)	0.11 (-0.01,0.22)	0.13 (0.02,0.26)	0.1 (-0.01,0.21)	0.13 (0.02,0.26)	0.1 (-0.02,0.21)	0.11 (0.03, 0.18)
ADHD & Beh	0.15 (0.03,0.29)	0.11 (-0.01,0.23)	0.11 (-0.01,0.24)	0.1 (-0.02,0.22)	0.11 (-0.01,0.23)	0.09 (-0.03,0.22)	0.11 (0.04, 0.19)
ADHD	0.09 (-0.02,0.2)	0.12 (0.02,0.23)	0.14 (0.03,0.26)	0.09 (-0.03,0.19)	0.11 (0,0.23)	0.08 (-0.03,0.19)	0.11 (0.03, 0.18)
Lang	0.1 (0.01,0.18)	0.15 (0.06,0.24)	0.15 (0.07,0.24)	0.1 (0.01,0.18)	0.06 (-0.03,0.15)	0.06 (-0.03,0.14)	0.1 (0.04, 0.16)
Beh	0.09 (-0.03,0.2)	0.07 (-0.06,0.18)	0.11 (0,0.24)	0.06 (-0.07,0.17)	0.11 (0,0.23)	0.1 (-0.02,0.21)	0.09 (0.02, 0.16)
Average	0.1 (0.01, 0.18)	0.15 (0.06, 0.24)	0.15 (0.07, 0.24)	0.1 (0.01, 0.18)	0.06 (-0.03, 0.15)	0.06 (-0.03, 0.14)	

Note. ATEs are reported as standardised mean differences (SMD). Numbers are means (90% credible intervals).

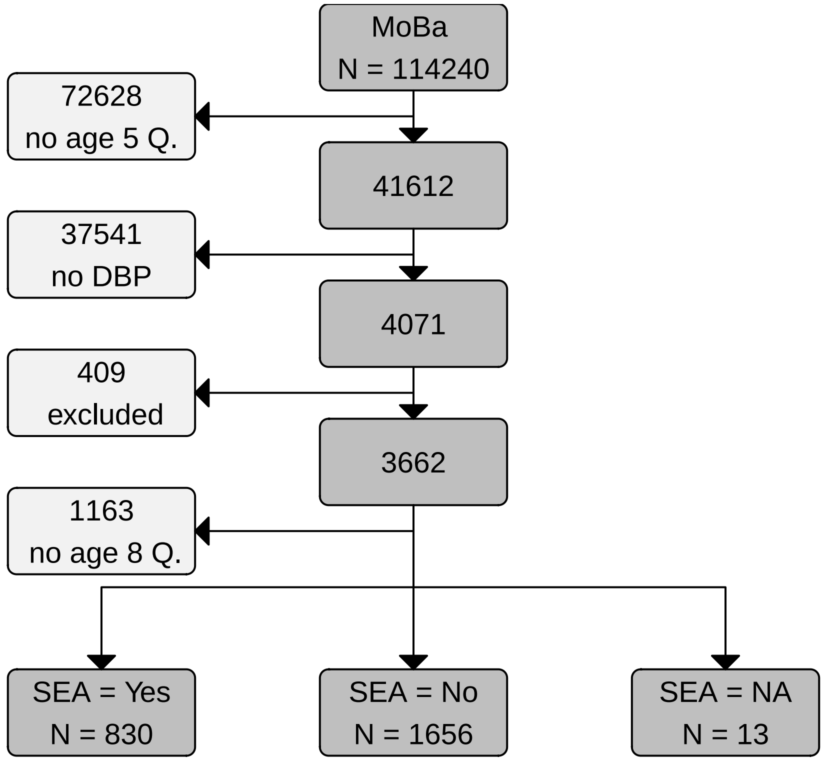

Figure 1. Inclusion flow chart. Age 5 and 8 Q are MoBa questionnaires sent out at child age five and eight years. Children in the study sample have a parent-reported developmental or behaviour problem (DBP) at age five. Children with epilepsy, cerebral palsy, chromosomal defects, severe developmental delay, or hearing loss were excluded from this study. SEA = special educational assistance in early childhood education and care (ECEC).

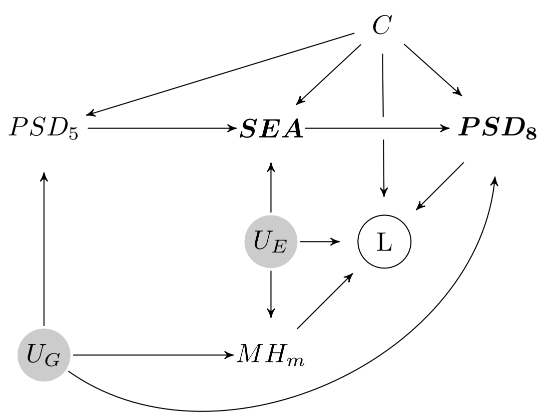

Figure 2. Directed Acyclic Graph of the hypothesized causal relationships between special educational assistance (*SEA*) psycho-social difficulties (PSD_5,PSD_8), loss to follow up (*L*), maternal mental health (MH_m), unobserved environmental and genetic causes (U_E,U_G) and additional confounders (*C*, i.e. contact with mental health services, maternal education, birth order, birth month, preterm birth). Current and prior psycho-social difficulties PSD_5 are confounders causing bias due to treatment by indication and can be controlled through adjustment. Because maternal mental health (MH_m) predicts loss to follow up (*L*), which is a collider on a backdoor path between *SEA* and PSD_8, loss to follow up has to be controlled through inverse probability weighting.

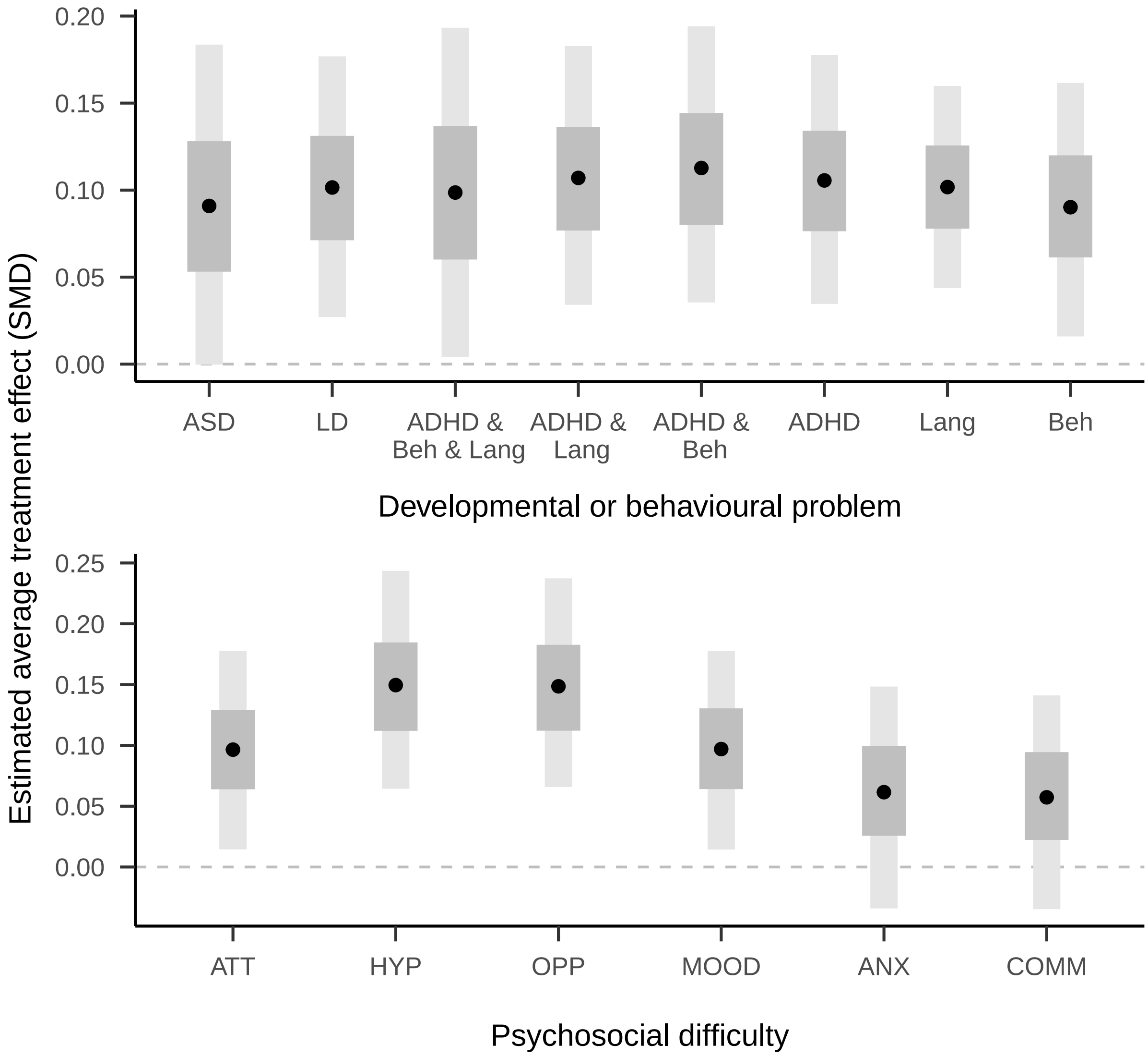

Figure 3. Estimated average treatment effects stratified by group (top) and outcome (bottom). Lines indicate means, grey and dark-grey bands indicate 50 and 95% credible intervals. SMD = standardised mean deviation. Abbreviations as in Table 1.

Supplementary Material

Supplementary Methods

All analysis were performed with R [Version 4.0.2; [1]][1]. Scripts for all analysis steps are available at https://github.com/gbiele/SPS358.

Classification into groups with different developmental or behavioral problems (DBP). Many children had combinations of developmental or behavioral problems. Altogether, we found 25 patterns of DBP, which we reduced to 8 non-overlapping groups. The guiding principle was to categorize each child according to it most severe problem, where the order of severity was ASD, LD, ADHD, language, and behavior problems. More specifically, we used following classification rules:

1. ASD = children with ASD (could have learning disabilities)
2. LD = children with learning disabilities, but no ASD
3. ADHDBehLang = children with ADHD and behavior and language problems but no ASD or LD,
4. ADHDLang = children with ADHD and language problems but no ASD, LD, or behavior problems,
5. ADHDBeh = children with ADHD and behavior problems but no ASD, LD, or

[1] We, furthermore, used the R-packages *apaTables* [Version 2.0.8; [2]], *arsenal* [Version 3.5.0; [3]], *bayesplot* [Version 1.8.0; [4]], *brms* [Version 2.14.4; [5]; [6]], *colorspace* [Version 2.0.0; [7]; [8]], *data.table* [Version 1.13.6; [9]], *DescTools* [Version 0.99.34; [10]], *english* [Version 1.2.5; [11]], *flextable* [Version 0.6.1; [12]], *Formula* [Version 1.2.3; [13]], *ggplot2* [Version 3.3.3; [14]], *ggpubr* [Version 0.4.0; [15]], *Gmisc* [Version 1.9.2; [16]], *Hmisc* [Version 4.4.0; [17]], *htmlTable* [Version 2.1.0; [18]], *jtools* [Version 2.1.0; [19]], *kableExtra* [Version 1.3.1; [20]], *knitr* [Version 1.30; [21]], *lattice* [Version 0.20.41; [22]], *magrittr* [Version 2.0.1; [23]], *officer* [Version 0.3.16; [24]], *papaja* [Version 0.1.0.9997; [25]], *pdftools* [Version 2.3.1; [**R-pdftools?**]], *Rcpp* [Version 1.0.6; [26]; [27]], *rstan* [Version 2.21.2; [28]; [29]], *rstanarm* [Version 2.21.1; [29]], *spelling* [30], *StanHeaders* [Version 2.21.0.6; [31]], and *survival* [Version 3.2.7; [32]].

language problems,

6. ADHD = children with ADHD problems but no ASD, LD, behavior or language problems,
7. Lang = children with language problems but no ASD, LD, ADHD or language problems,
8. Beh = children with oppositional defiant and/or conduct disorder problems (but no ASD, LD, ADHD or language problems)

Table S1 shows to which categories children with different developmental difficulties were assigned.

Table S1

Patterns of developmental and behavioral problems and their classification into groups

	ASD	LD	ALB	AL	AB	ADHD	Lang	Beh
ADHD	0	0	0	0	0	401	0	0
ADHD + ASD	2	0	0	0	0	0	0	0
ADHD + Beh	0	0	0	0	146	0	0	0
ADHD + Beh + ASD	1	0	0	0	0	0	0	0
ADHD + Beh + ASD + LD	1	0	0	0	0	0	0	0
ADHD + Beh + Lang	0	0	31	0	0	0	0	0
ADHD + Beh + Lang + ASD	3	0	0	0	0	0	0	0
ADHD + Beh + Lang + ASD + LD	2	0	0	0	0	0	0	0
ADHD + Beh + Lang + LD	0	9	0	0	0	0	0	0
ADHD + Beh + LD	0	2	0	0	0	0	0	0
ADHD + Lang	0	0	0	143	0	0	0	0
ADHD + Lang + ASD	8	0	0	0	0	0	0	0
ADHD + Lang + ASD + LD	3	0	0	0	0	0	0	0
ADHD + Lang + LD	0	24	0	0	0	0	0	0
ADHD + LD	0	4	0	0	0	0	0	0
ASD	8	0	0	0	0	0	0	0
ASD + LD	1	0	0	0	0	0	0	0
Beh	0	0	0	0	0	0	0	320
Beh + ASD	5	0	0	0	0	0	0	0
Beh + Lang	0	0	0	0	0	0	26	0
Lang	0	0	0	0	0	0	1307	0

Table S1 continued

	ASD	LD	ALB	AL	AB	ADHD	Lang	Beh
Lang + ASD	7	0	0	0	0	0	0	0
Lang + ASD + LD	2	0	0	0	0	0	0	0
Lang + LD	0	28	0	0	0	0	0	0
LD	0	15	0	0	0	0	0	0

Note. Cells show number of cases. ALB = ADHD & Lang & Beh, AL = ADHD & Lang, AB = ADHD & Beh. All other appreviations as in Table 1 of the main article

Multiple chained imputation of missing values. Table S2 shows the percentage of missing data. The analysis used altogether 214 variables, of which 172 were items assessing psycho-social difficulties.

To deal with missing data among co-variates, we generated 50 data sets with imputed missing data using the multiple chained imputation approach as implemented in the R package mice [33]. The imputation data set included 3662 participants for which the five or eight year questionnaire was available. The rating data was imputed on the level of individual rating items, which lead to a large number of 214 variables for the imputation process. Because the data set included 187 ordinal variables (mostly rating scale items) and the mice package does currently not offer facilities for fast imputation of ordinal data, we wrote a mice extension (https://github.com/gbiele/spolr) for efficient calculation of penalized ordered, logistic, and Gaussian regression using the rstan package. Here, penalization is implemented by putting weakly informative priors on regression weights for z-standardized predictors and estimating maximum a-posteriori parameter estimates.

We ran mice for four chains with 100 iterations each and verified through visual inspection that all chains had converged.

Inverse probability of continued participation weights. To predict loss to follow up from the MoBa age five to eight year questionnaire we used following groups of variables :

- child psycho-social difficulties at age 5
- maternal mental health
- type of the developmental or behavioral problem
- maternal age and education
- parity, birth month, child gender
- contact with health services
- special educational assistance (see also Figure S1)

Consistent with recommendations for calculating weights for hierarchical analysis, we used a hierarchical regression analysis [implemented in rstanarm, 34] as the basis for calculating stabilized inverse probability of continued participation weights (IPPW) for groups with different DBPs. In particular, we estimated a hierarchical logistic regression model with random intercepts and slopes, with type of developmental or behavioral problem, and child gender as grouping variables for random/group-specific effects. We fit a selection model for each of the 50 imputed data in R, using rstanarm [35], and obtained inverse probability of continued participation weights as the continuation probability for all participants in the 5 year questionnaire divided by each individuals continuation probability predicted by the selection model.

When evaluating inverse probability of continued participation weights, the focus should be on the successful balancing of the weighted follow up sample with the original inclusion sample, with respect to potential confounders [36]. Figure S1 shows standardized mean deviations (SMD) for all imputed samples, once non-weighted and once weighted by inverse probability of participation weights. A conventional threshold is that the magnitude of the SMD between participants who were and were not lost to follow up should be below

0.1. In the Figure, dots to the right (left) of the vertical zero line indicate that participants with high values on the variable were more (less) likely to also participate in the 8 year questionnaire.

In addition to means, the entire distributions of potential confounders should be matched. This is typically investigate through visual inspection of cumulative density plots, which are shown in Figure S2. Figure S3 directly shows the difference of the empirical cumulative distribution function and reinforces that weighting reduces the differences in the distribution of values for all predictor variables.

Table S2

Percent missing data

	Participants with MoBa age five, eight data			Participants with at least MoBa age eight data		
	all variables	age 5 scales	age 8 scales	all variables	age 5 scales	age 8 scales
Minimum	0.0	0.2	0.3	0.0	0.5	31.9
1st Quartile	0.6	0.7	0.5	0.9	0.7	32.1
Median	0.8	0.8	0.6	8.9	0.9	32.2
Mean	3.1	1.9	0.7	15.0	2.0	32.2
3rd. Quartile	1.8	1.0	0.8	32.1	1.3	32.3
Maximum	65.7	65.7	1.4	67.2	67.2	32.7

Note. Age 5/8 scales refers to items to assess psycho-social difficulties. MoBa = Norwegian Mother, Father and Child Cohort Study

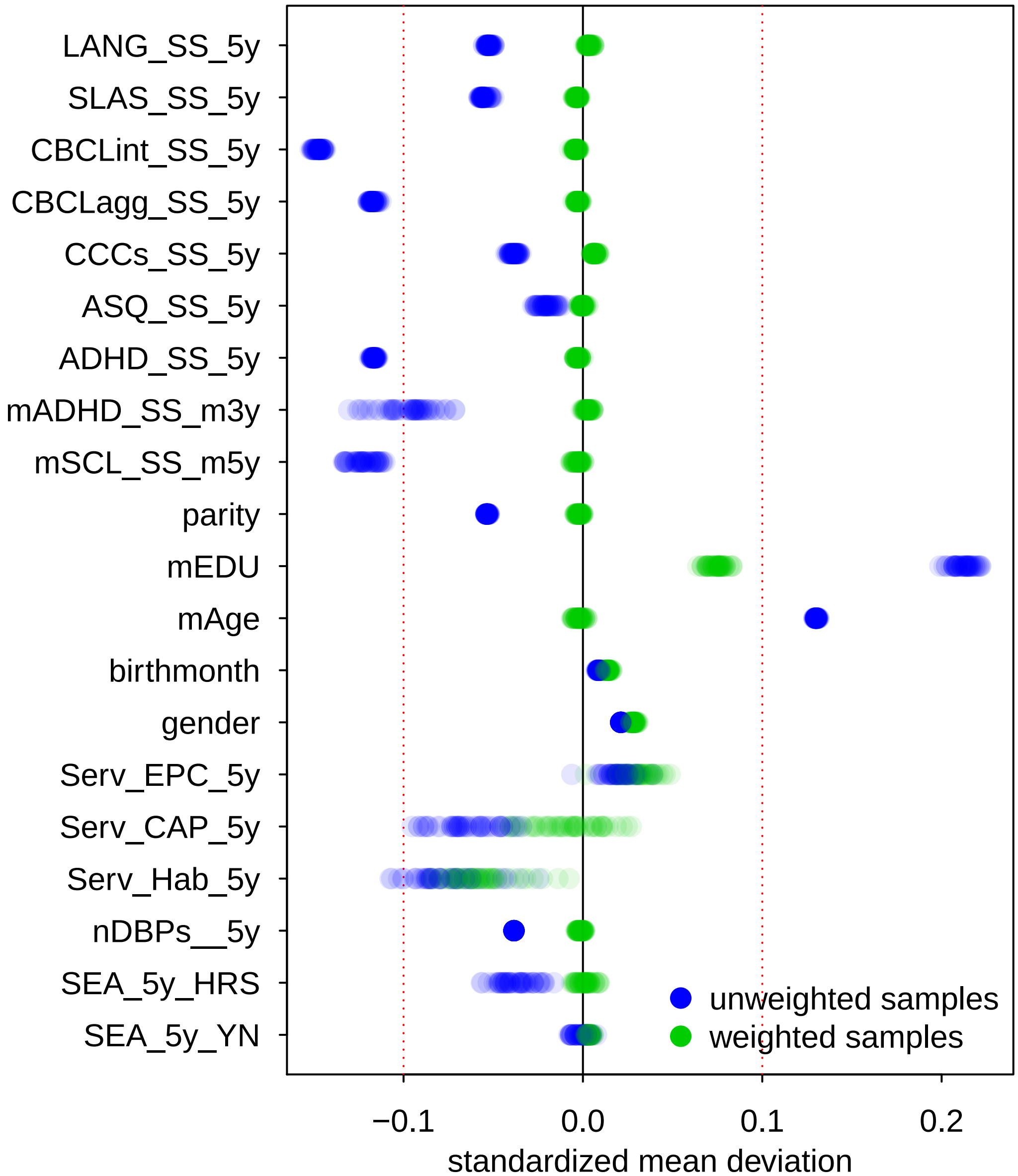

Figure S1. Standardized mean difference between respondents to the age 8 year questionnaire and those lost to follow up. For each variable that influences participation, each sample is represented with 50 transparent, overlapping dots (one for each imputed data set) for the standardized mean difference (SMD) between participants who were and were not lost to follow up. SS = sum-score, 5y = 5 year questionnaire, nDBP = number of developmental of behavioral problems, EPC = educational and psychological counseling service, CAP = Child and adolescent psychiatric units.

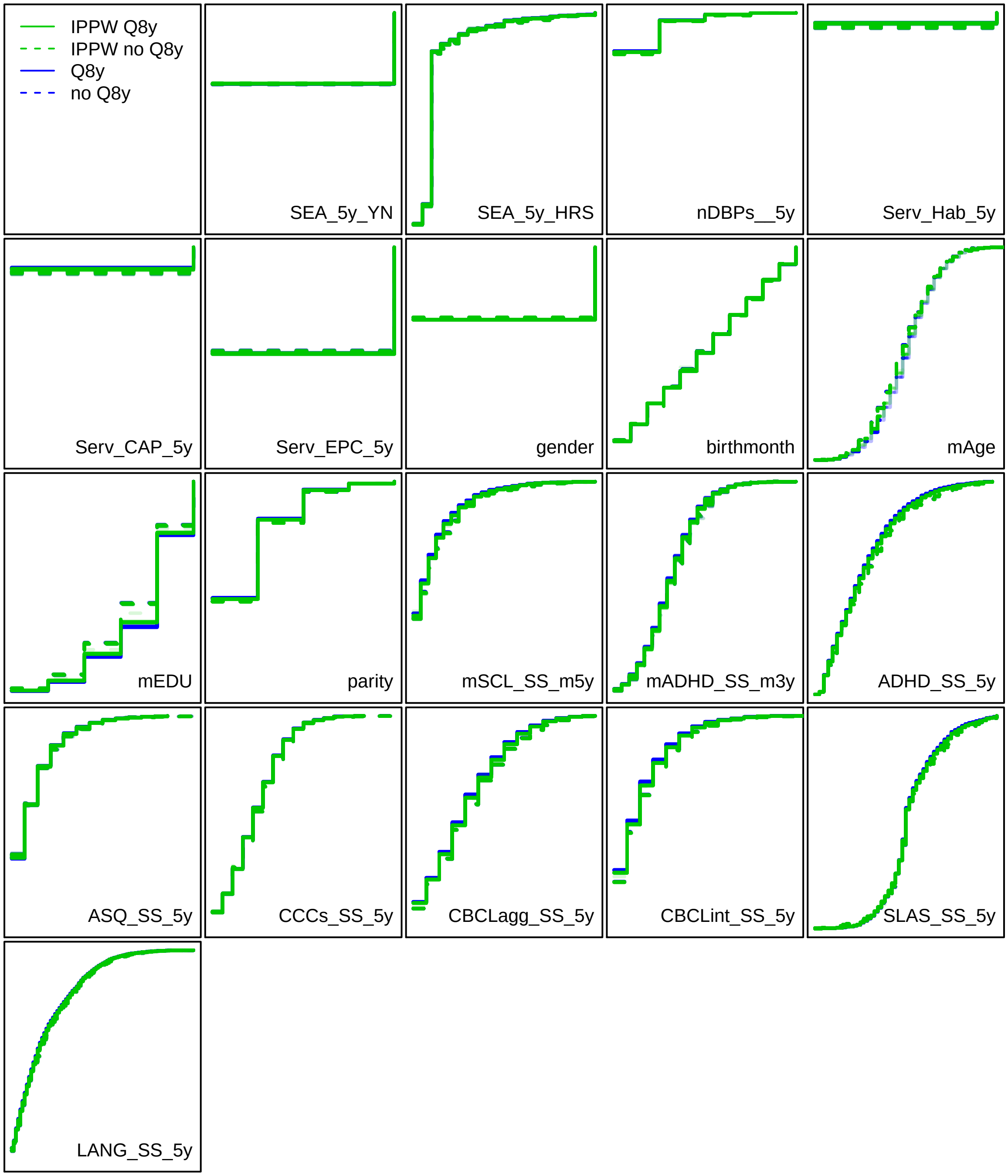

Figure S2. Cumulative distributions for respondents to the 8 year questionnaire (Q8y) and those lost to follow up (no Q8y) before weighting (blue) and after weighting (green). Increasing values of participation predictors are on the x-axis, the cumulative proportion of participants up to a variable value are on the y-axis. The sample is properly balanced, if the cumulative distribution function for participants with and without 8 year questionnaire are identical (i.e. the slid and dashed lines overlap).

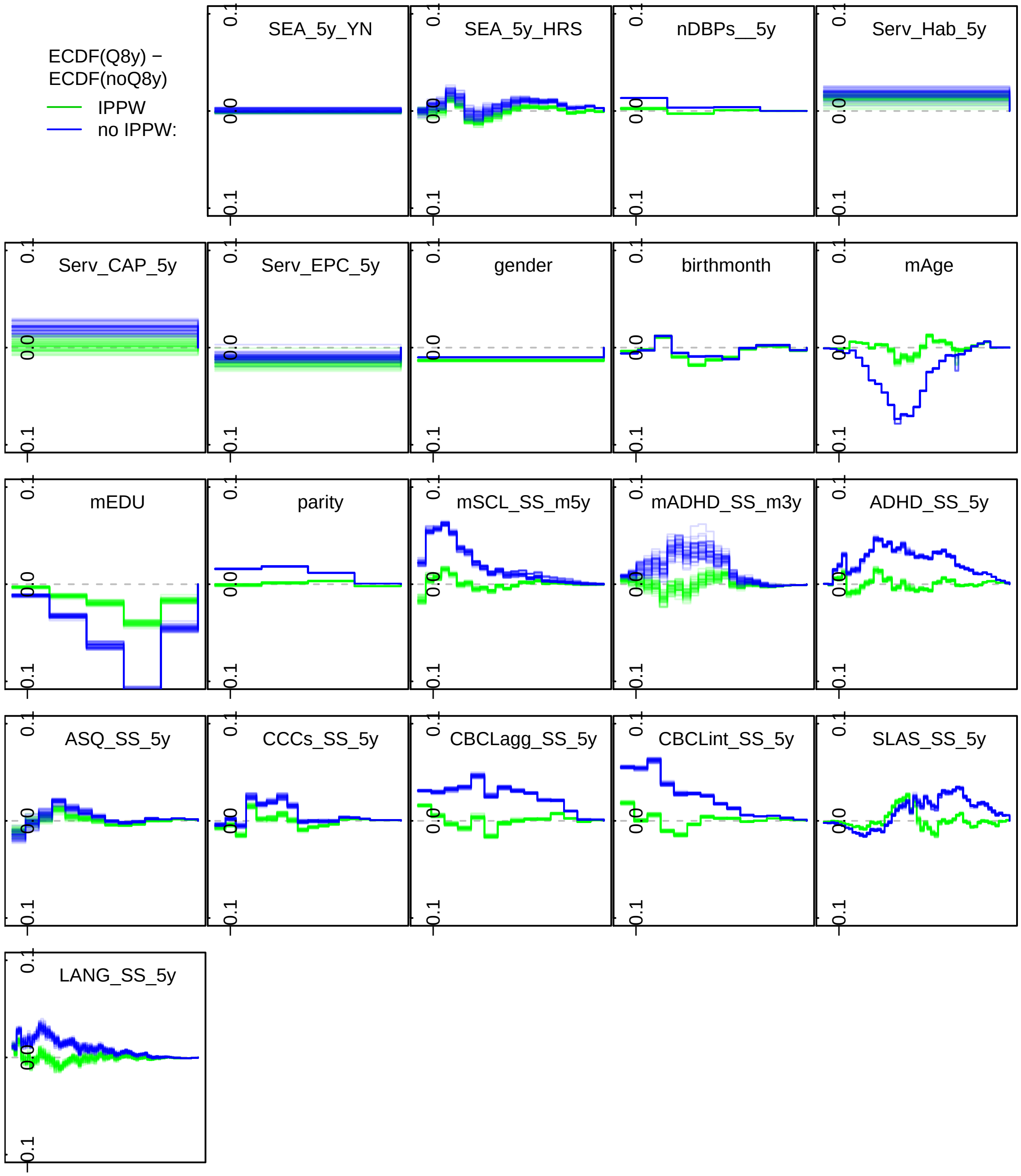

Figure S3. Difference in cumulative distributions for respondents to the 8 year questionnaire (Q8y) and those lost to follow up (no Q8y) before weighting (blue) and after weighting (green). Increasing values of participation predictors are on the Z axis, the differences of the cumulative proportion of participants up to a variable value are on the y axis. The sample is properly balanced, if the differences between participants with and without 8 year questionnaire is close to zero.

Regression models and estimation of average treatment effects. Outcome variables are sum-scores, which we model as ordinal variables. That is, the analysis model assumes latent traits for psycho-social difficulties, which in combination with cut K-1 cut points result in one of K possible sum-scores. This approach captures the intuition that sum-scores from questionnaires do not measure psycho-social difficulties on an interval scale and it also facilitates calculation of average treatment effects on the scale of standardized mean differences.

In order to obtain reliable estimates also for smaller groups, and in order to deal with the multiple comparison problem, we estimated the effects for all groups and psycho-social difficulties jointly in a hierarchical regression.

In particular, we estimate a random-intercept random-slope model for strata determined by (a) developmental or behavioral problems, (b) psycho-social difficulties, and (c) sex Further, we account for repeated measures within individuals by estimating individual level intercepts/random effects. We estimated the parameters for this model with a version of the (cumulative) ordinal regression implemented in the brms package[37, 38], which we modified to account for the fact that different outcome measures had different numbers of categories.

Un-adjusted model.

Even though the unadjusted model cannot be used to obtained unbiased effect estimate, we estimated such a model for reasons of completeness and to conform with STROBE reporting guidelines.

The unadjusted model estimated the effect of SEA as

$$\begin{aligned} sumscore \sim & 1 + SEA + h_{SEA} + h^2_{SEA} + \\ & (1 + SEA + h_{SEA} + h^2_{SEA} | DBP : PSD : sex) + \\ & (1|ID) \end{aligned}$$

where $SEA + h_{SEA} + h^2_{SEA}$ are the fixed effects for SEA and the linear and quadratic effects of the number of hours SEA, $(SEA + h_{SEA} + h^2_{SEA} | DBP : PSD : sex)$ are random effects for strata define by the type of developmental or behavior problem, outcome, and gender. $(1|ID)$ are participant level random effects.

The estimation of the model used inverse probability of continued participation weights to account for loss to follow up.

All non-binary predictors (including linear and quadratic effects) were scaled to mean of zero and a standard deviation of 1. We employed weakly informative shrinkage priors (normal distribution with a mean of zero and a standard deviation of 2) for the estimation of the fixed effects and for the estimation of the random effects standard deviation (half-normal distribution with a mean of zero and a standard deviation of 3). To verify convergence of all models we checked that all $\hat{R}$ values were below 1.1[39], and that the model estimation was completed without divergent iterations.

For each of the 50 imputed data sets, we estimated one model with four chains, each with 1000 warmup iteration and 500 post-warmup iterations. The reported results are based on the pooled results over all these analyses, i.e. based on 500 * 4 * 50 = 100,000 post warmup samples.

Adjusted model.

The basic structure of the adjusted model is the same as described for the un-adjusted model, except that we used additional covariates.

Figure 2 shows the causal assumptions underlying our analysis. Based on these causal assumptions we control bias due to treatment by indication by adjusting for maternal education and the degree of psycho-social difficulties at age five. Regarding the latter, we adjust for linear and quadratic effects of sum-scores for ADHD symptoms, externalizing and internalizing behavior (CBCL) and communication (CCC), developmental difficulties (ASQ) as well as interactions between these variables. We also included the number of developmental or behavioral problems at age five and contact to different types of mental health services as indicators of psycho-social difficulties at age five. We further adjusted for pregnancy duration, maternal education and mental health (ADHD and depression symptoms) as potential confounders, as well as for birth month and birth order, which are not of interest for the current study, but were shown to be associated with child mental health problems [40]. Lastly, we also adjusted for the number of children in the kindergarten group and the number of children per adult as indicators for variations in the quality of ECEC care.

Calculation of average treatment effects.

We calculated average treatment effects as

$$
\begin{aligned}
ATE_{g,o} =& \frac{\sum_{i=1}^{n_g} \omega_i TE_i}{\sum_{i=1}^{n_g} \omega_i} \\
TE =& \hat{Y}_1 - \hat{Y}_0 \\
\hat{Y}_1 =& \beta_F X_1 + \beta_R z_1 \\
\hat{Y}_0 =& \beta_F X_0 + \beta_R z_0
\end{aligned}
$$

where $ATE_{g,o}$ is the average treatment effect for a particular group and outcome, ω is a vector or IPPW weights, and TE is a vector of treatment effects and n_g is the number of individuals in a group. X is a matrix with fixed effects predictors and β_F are the associated regression weights. z are random effects predictors and β_R are the associated regression weights. In X_1 and z_1 the value for SEA is set to 1, and the value for the hours SEA is set to the true value for children who received SEA and to the imputed values for children who did

Table S3

Coefficients of the hierarchical regression model without adjustment

predictor	FE coefficient	RE stand. dev.
supp_5y_YN	1.053 (0.607, 1.513)	2.111 (1.829, 2.456)
supp_5y_HRS.L	-0.142 (-0.287, 0)	0.592 (0.503, 0.696)
supp_5y_HRS.Q	0.174 (0.055, 0.297)	0.52 (0.437, 0.619)

Note. supp_5y_YN = Special educational assistance (SEA) received, supp_5y_HRS = hours SEA per week, FE = fixed effects, RE = random effects. L = linear effects, Q = quadratic effects. Numbers are means and 90% credible intervals.

not receive SEA. In X_0 and z_0 the values for SEA and hours SEA is set to 0. For reasons of transparency this description omits indices for the 50 imputations and 500 post warm-up MCMC samples per imputations, but all ATEs were calculated by averaging over all imputations and MCMC samples.

Supplementary Results

Unadjusted model: The association between SEA and psycho-social and developmental problems. The unadjusted model estimated a hierarchical regression model with only the presence of SEA and linear and quadratic effects of the number of hours SEA per week as predictors. Given the causal relationships displayed in Figure 2 the positive association between SEA and psycho-social and developmental problems at age 8 was expected. This should however not be taken as evidence a negative effect of SEA.

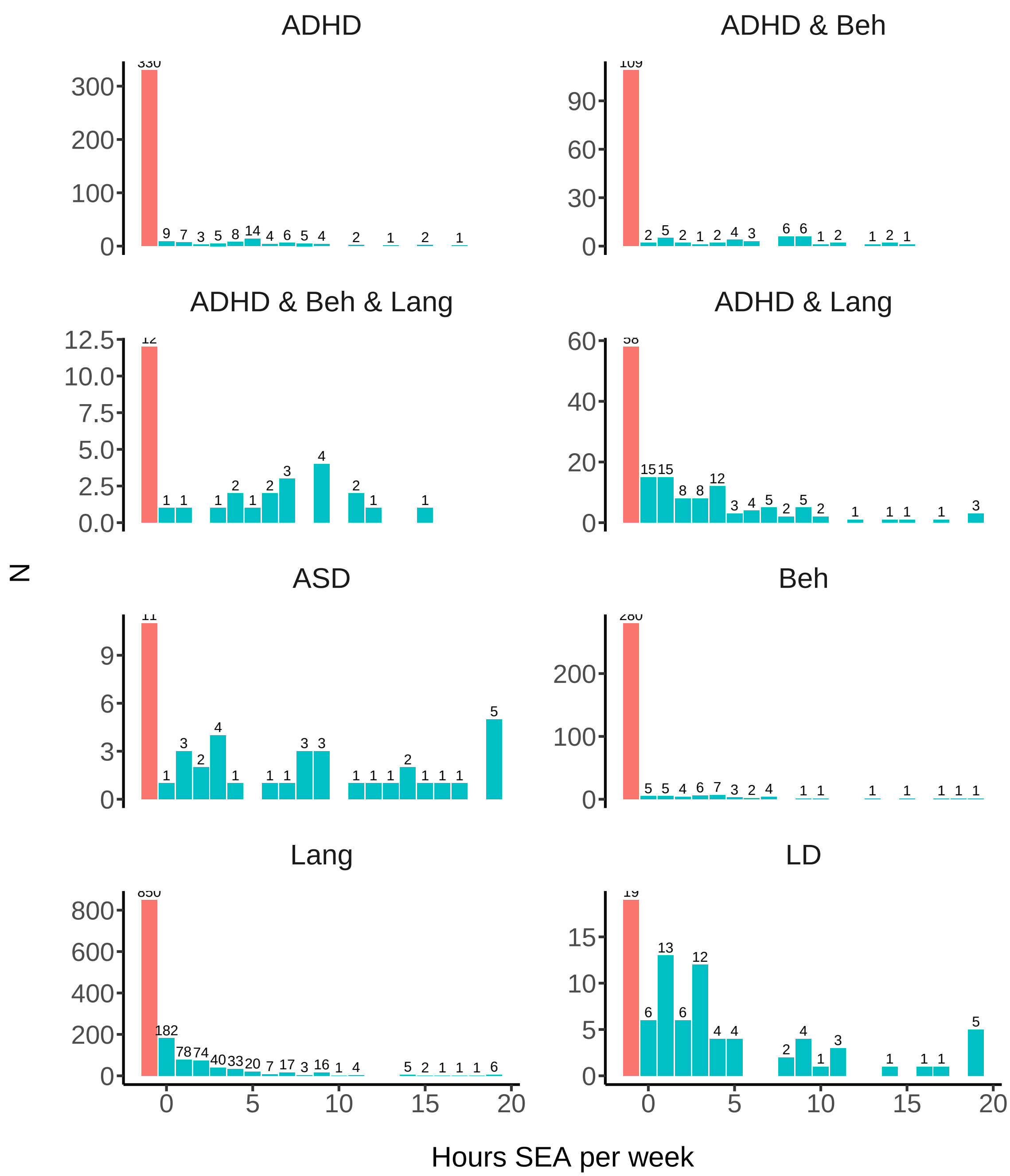

Figure S4. Distribution of weakly hours special educational assistance (SEA), stratified by type of developmental or behavioral problem.

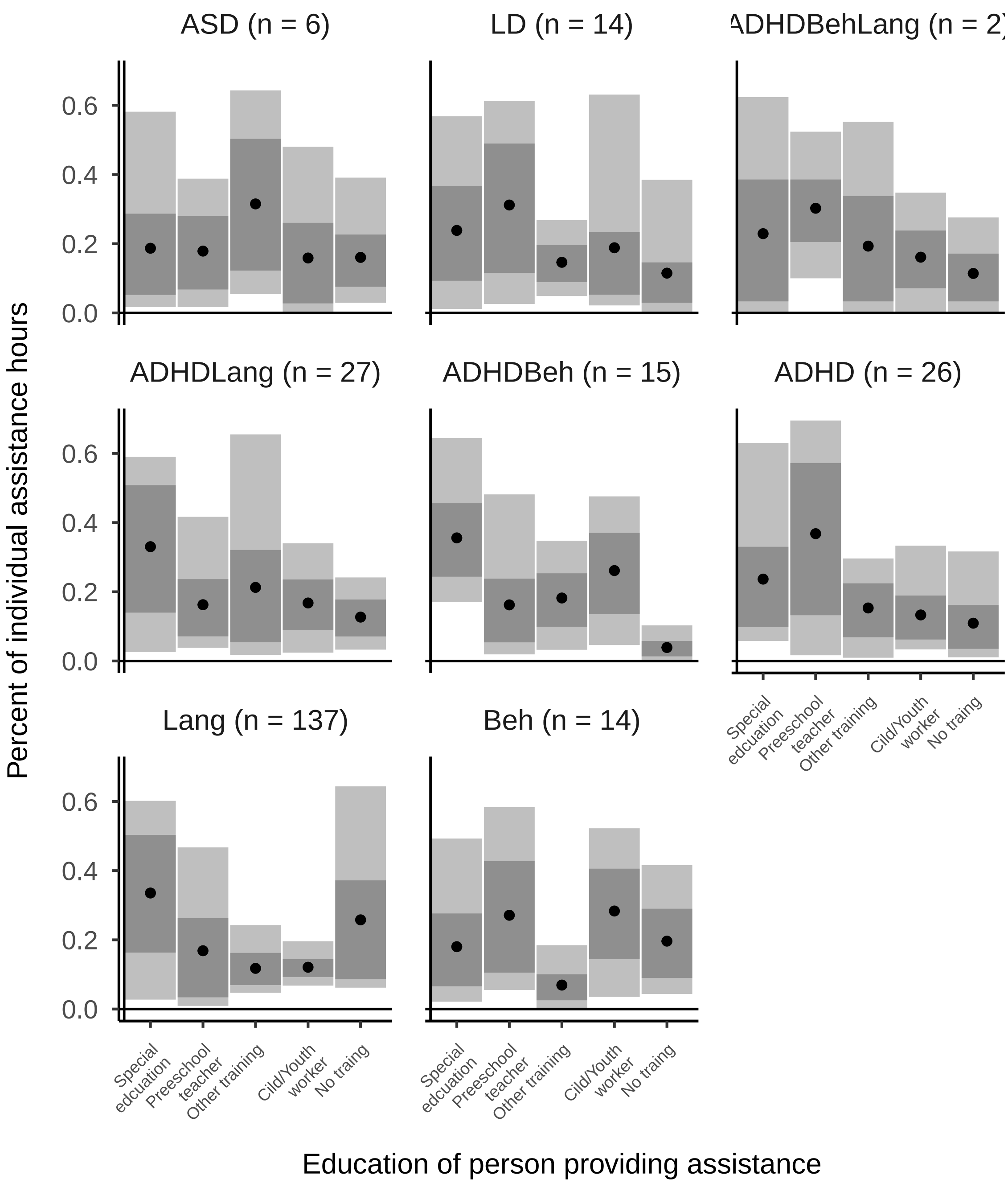

Figure S5. Education of employees providing individual assistance in ECEC to children with DBP. Only employees in the categories "Special education" and "Preschool teacher" have a longer pedagogical education. Only children with ASD, delayed development or combined ADHD and language problems receive the majority of assistance from persons with training in special education. All other children receive assistance mostly from regular preschool teachers or adults without dedicated education. Data are from the MoBa Kindergarten questionnaire, which was sent to and returned from only a subsample of the MoBa population.

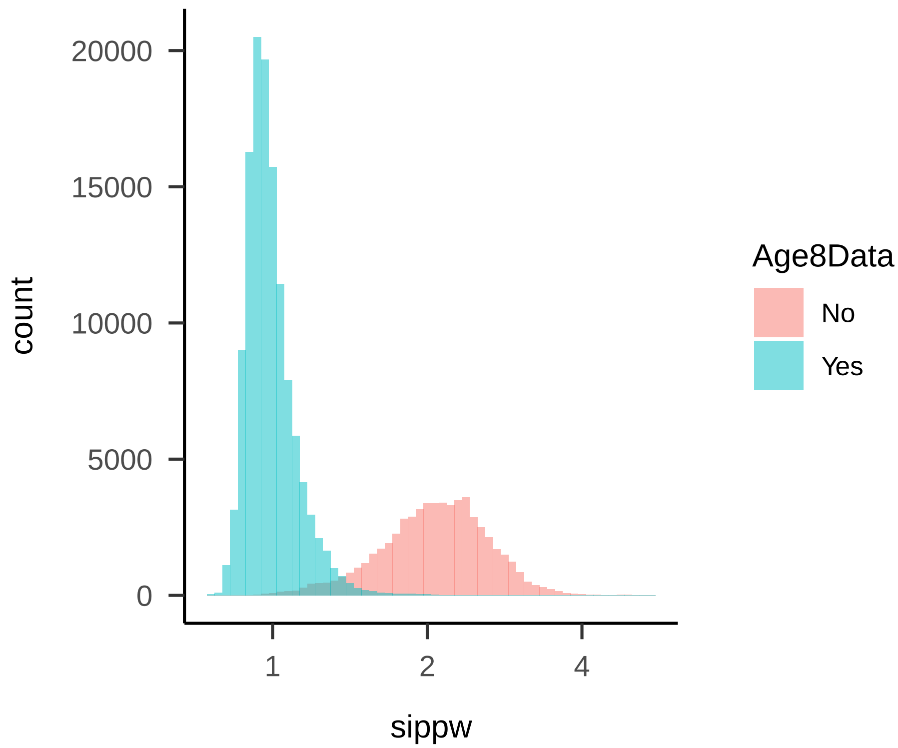

Figure S6. Distribution of stabilized inverse probability of continued participation weights (IPPW). The histogram shows weights for all 50 imputed data sets. Weights for participants with data from the age eight questionnaire range between 0.76 and 2.27.

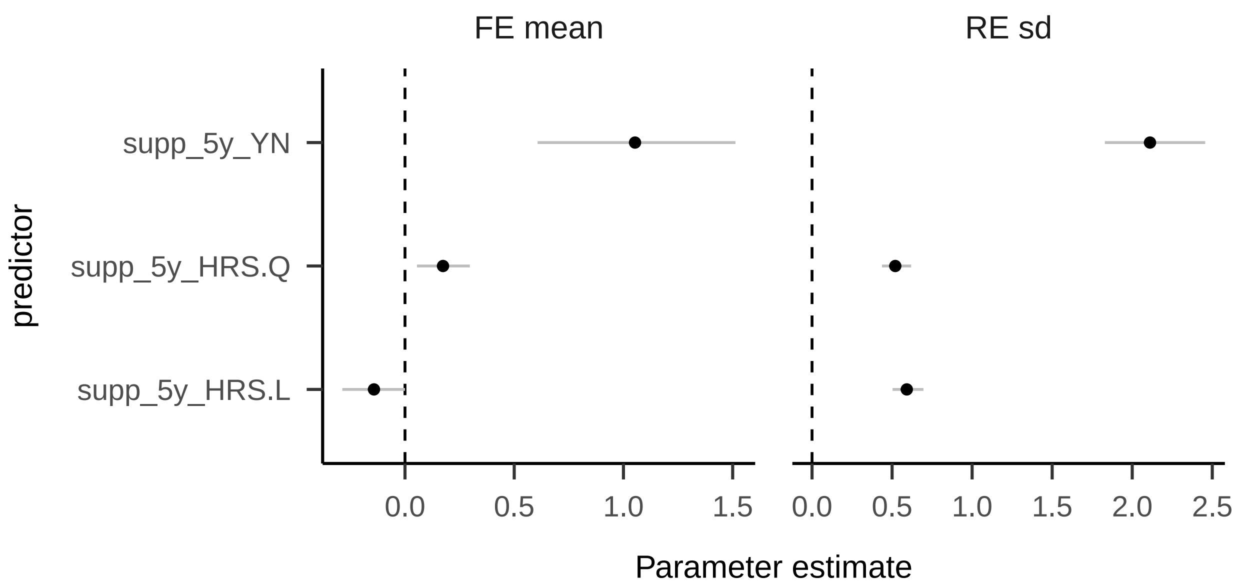

Figure S7. Coefficients of the hierarchical regression model without adjustment. FE mean = fixed effects, RE sd = standard deviation of random effects. L = linear effects, Q = quadratic effects. Colons between variables indicate interactions. The plot shows means as dots and 90% credible intervals as vertical lines.

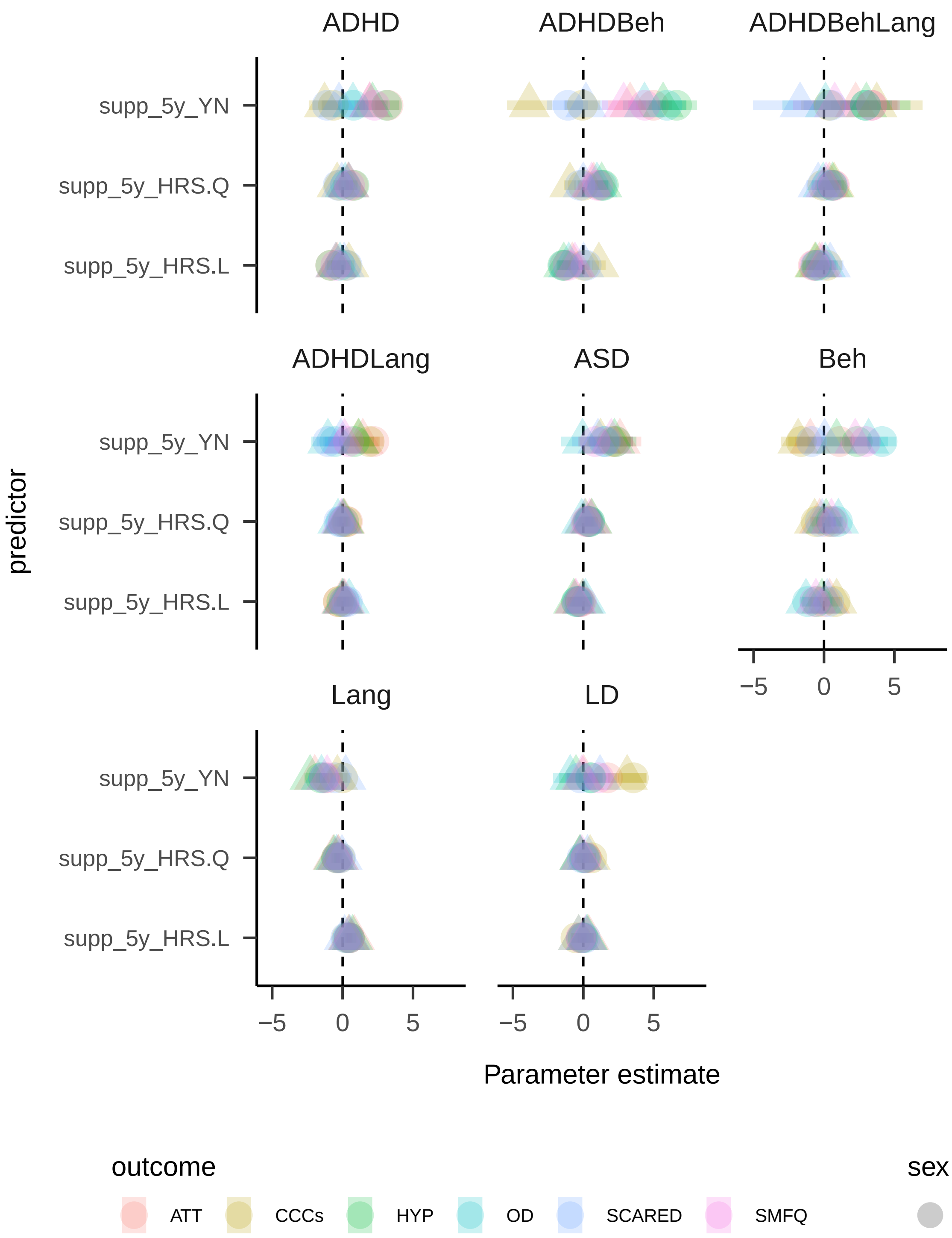

Figure S8. Group specific regression coefficients for the model without adjustment (fixed effects + random effects)

Table S4

Coefficients of the hierarchical regression model with adjustment

predictor	FE coefficient	RE stand. dev.
supp_5y_YN	-0.213 (-0.391, -0.037)	0.126 (0.018, 0.261)
supp_5y_HRS.L	0.014 (-0.063, 0.09)	0.036 (0.003, 0.093)
supp_5y_HRS.Q	0.005 (-0.056, 0.067)	0.027 (0.003, 0.076)
mEDU.L	-0.056 (-0.108, -0.004)	0.1 (0.06, 0.145)
mEDU.Q	0.02 (-0.027, 0.067)	0.057 (0.01, 0.105)
nMHPs_5y.L	0.055 (-0.029, 0.138)	0.086 (0.009, 0.193)
nMHPs_5y.Q	-0.081 (-0.145, -0.018)	0.059 (0.007, 0.131)
parity.L	-0.095 (-0.152, -0.037)	0.138 (0.086, 0.196)
parity.Q	0.056 (0.01, 0.102)	0.024 (0.002, 0.069)
preg_week.L	0.017 (-0.028, 0.063)	0.03 (0.003, 0.079)
preg_week.Q	-0.007 (-0.051, 0.038)	0.025 (0.002, 0.068)
birthmonth.L	0.05 (-0.003, 0.102)	0.113 (0.07, 0.162)
birthmonth.Q	-0.002 (-0.048, 0.044)	0.055 (0.009, 0.107)
LNX_SS_5y.L	0.371 (0.243, 0.499)	0.573 (0.496, 0.665)
LNX_SS_5y.Q	-0.033 (-0.096, 0.031)	0.099 (0.028, 0.166)
CBCLagg_SS_5y.L	0.437 (0.337, 0.539)	0.353 (0.292, 0.425)
CBCLagg_SS_5y.Q	0.029 (-0.037, 0.095)	0.036 (0.003, 0.096)
CBCLint_SS_5y.L	0.248 (0.168, 0.329)	0.248 (0.202, 0.302)
CBCLint_SS_5y.Q	-0.016 (-0.073, 0.041)	0.102 (0.053, 0.149)
ATT_SS_5y.L	0.24 (0.112, 0.368)	0.388 (0.307, 0.479)
ATT_SS_5y.Q	-0.058 (-0.163, 0.048)	0.144 (0.042, 0.226)
HYP_SS_5y.L	0.261 (0.144, 0.378)	0.273 (0.209, 0.349)
HYP_SS_5y.Q	0.042 (-0.057, 0.142)	0.106 (0.023, 0.178)
mADHD_SS_m3y.L	0.109 (0.052, 0.166)	0.102 (0.045, 0.159)
mADHD_SS_m3y.Q	-0.066 (-0.115, -0.016)	0.063 (0.013, 0.116)
mSCL_SS_m5y.L	0.094 (0.042, 0.147)	0.068 (0.01, 0.129)
mSCL_SS_m5y.Q	-0.027 (-0.079, 0.024)	0.081 (0.011, 0.163)
Serv_BUP_5yJa	0.098 (-0.093, 0.293)	0.185 (0.02, 0.41)
Serv_Hab_5yJa	0.177 (-0.106, 0.452)	0.217 (0.022, 0.542)

Table S4 continued

predictor	FE coefficient	RE stand. dev.
Serv_PPT_5yJa	0.034 (-0.073, 0.14)	0.08 (0.008, 0.199)
lnum_childr_KgGr	-0.132 (-0.291, 0.026)	0.123 (0.064, 0.171)
num_childr_per_adult_KgGr	0.037 (-0.006, 0.081)	0.018 (0.002, 0.05)
LNX_SS_5y.L:CBCLagg_SS_5y.L	-0.051 (-0.116, 0.015)	0.105 (0.035, 0.164)
LNX_SS_5y.L:CBCLint_SS_5y.L	-0.024 (-0.075, 0.027)	0.024 (0.002, 0.066)
LNX_SS_5y.L:ATT_SS_5y.L	0.029 (-0.052, 0.109)	0.034 (0.003, 0.094)
LNX_SS_5y.L:HYP_SS_5y.L	0.007 (-0.069, 0.083)	0.061 (0.007, 0.131)
CBCLagg_SS_5y.L:CBCLint_SS_5y.L	-0.079 (-0.141, -0.018)	0.058 (0.008, 0.111)
CBCLagg_SS_5y.L:ATT_SS_5y.L	-0.022 (-0.112, 0.066)	0.034 (0.003, 0.096)
CBCLagg_SS_5y.L:HYP_SS_5y.L	-0.071 (-0.157, 0.016)	0.028 (0.003, 0.082)
CBCLint_SS_5y.L:ATT_SS_5y.L	0.056 (-0.02, 0.131)	0.039 (0.004, 0.1)
CBCLint_SS_5y.L:HYP_SS_5y.L	0.052 (-0.02, 0.125)	0.028 (0.002, 0.079)
ATT_SS_5y.L:HYP_SS_5y.L	-0.046 (-0.162, 0.07)	0.077 (0.009, 0.154)

Note. supp_5y_YN = Special educational assistance (SEA) received, supp_5y_HRS = hours SEA per week, nMHPs = number of developmental or behavior problems, LNX = composite sum score for language difficulties, mADHD / mSCL = maternal ADHD and depressive symptoms, Serv_BUP/PPT/Hab = contact with child and adolescent psychiatric unit/educational and psychological counselling service, habilitation service, num_childr_KgGr = number of children in kindergten group, FE = fixed effects, RE = random effects. L = linear effects, Q = quadratic effects. 3y/5y = measured with MoBa age three/five years questionnaires. Colons between variables indicate interactions. Numbers are means and 90% credible intervals.

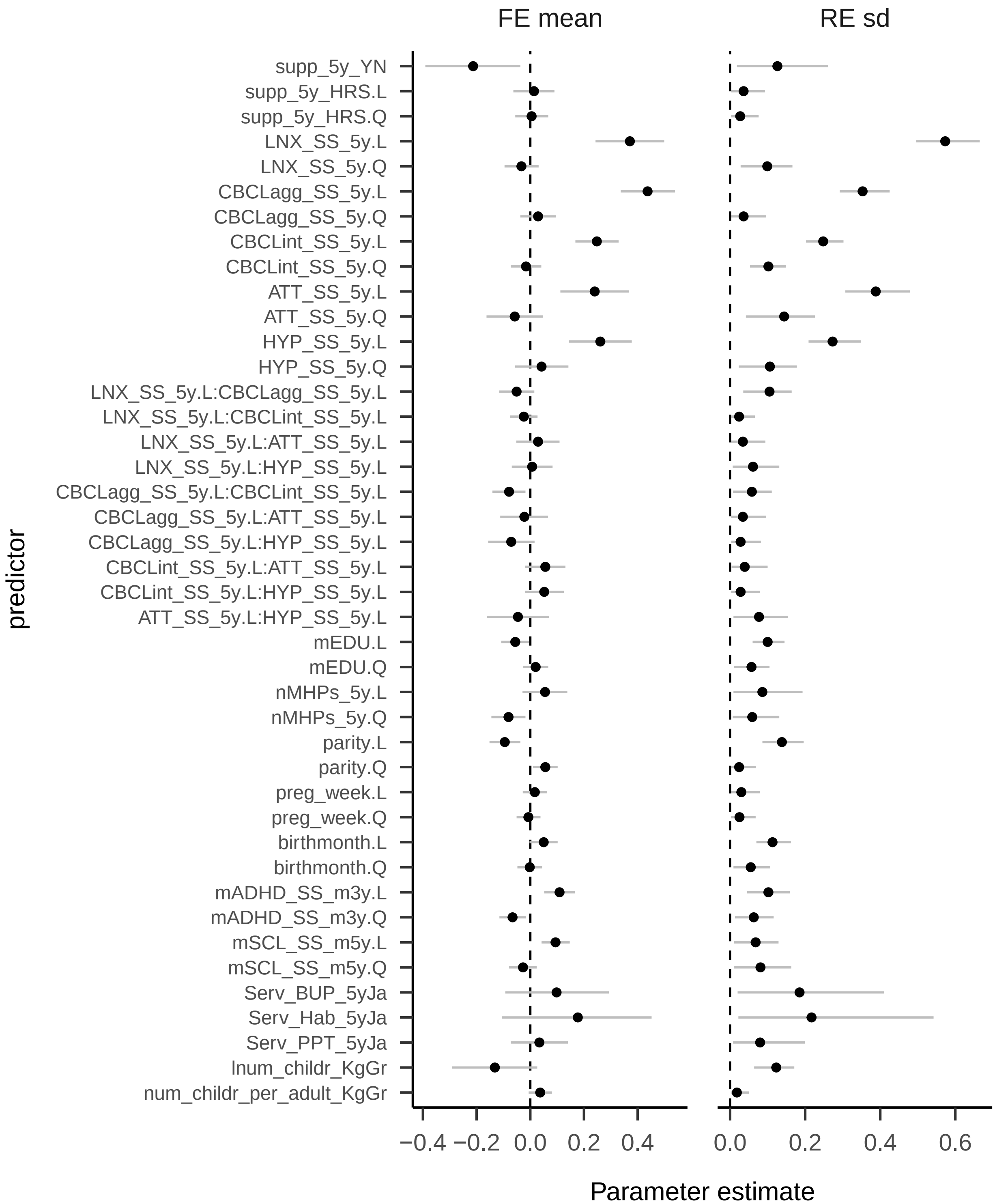

Figure S9. Coefficients of the hierarchical regression model with adjustment. FE mean = fixed effects, RE sd = standard deviation of random effects. L = linear effects, Q = quadratic effects. Colons between variables indicate interactions. The plot shows means as dots and 90% credible intervals as vertical lines.

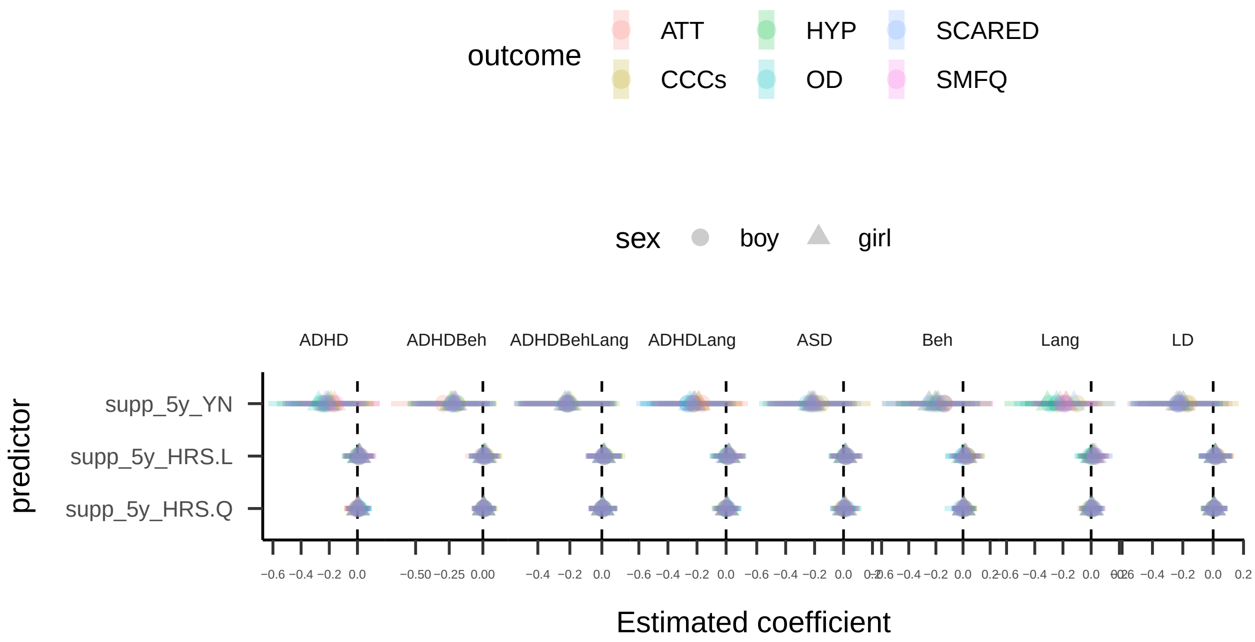

Figure S10. Group specific regression coefficients for the model with adjustment (fixed effects + random effects): Special educational assistance.

Estimated average treatment effects. Average treatment effects are only reported for the model with adjustment, because the model without adjustment produces obviously biased results (due to treatment by indication).

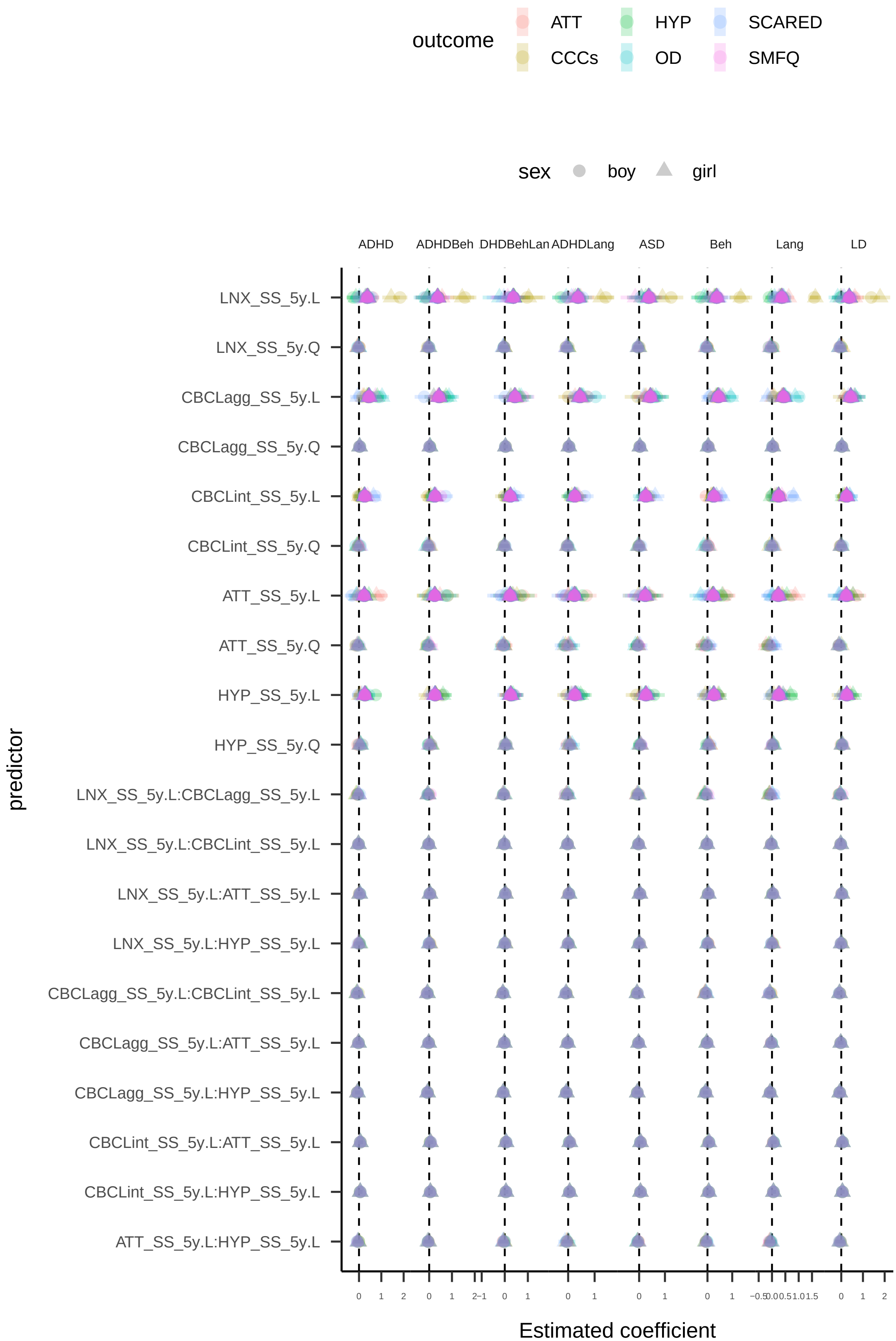

Figure S11. Group specific regression coefficients for the model with adjustment (fixed effects + random effects): Difficulties at baseline.

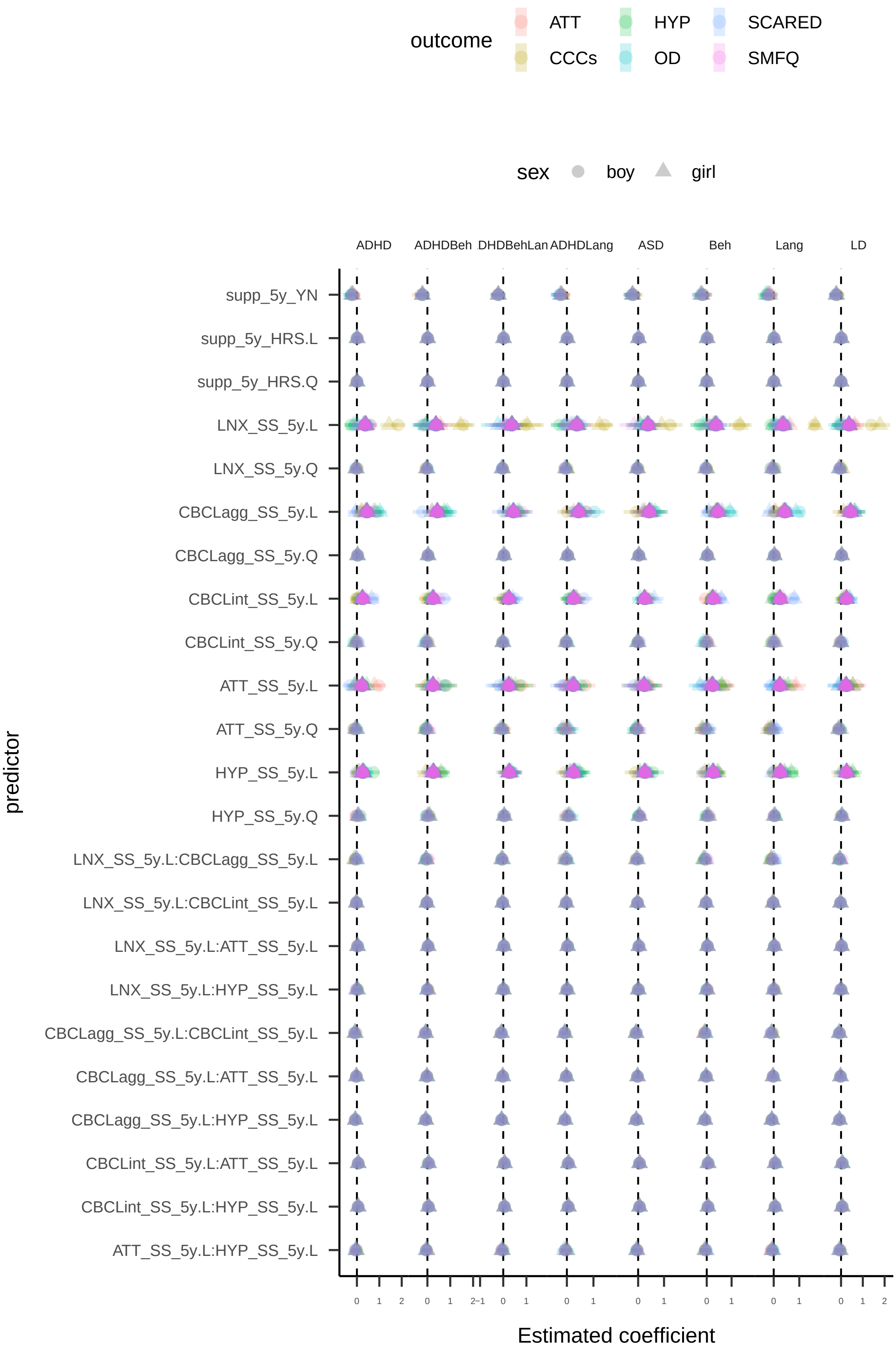

Figure S12. Group specific regression coefficients for the model with adjustment (fixed effects + random effects): Maternal characteristics and ECEC.

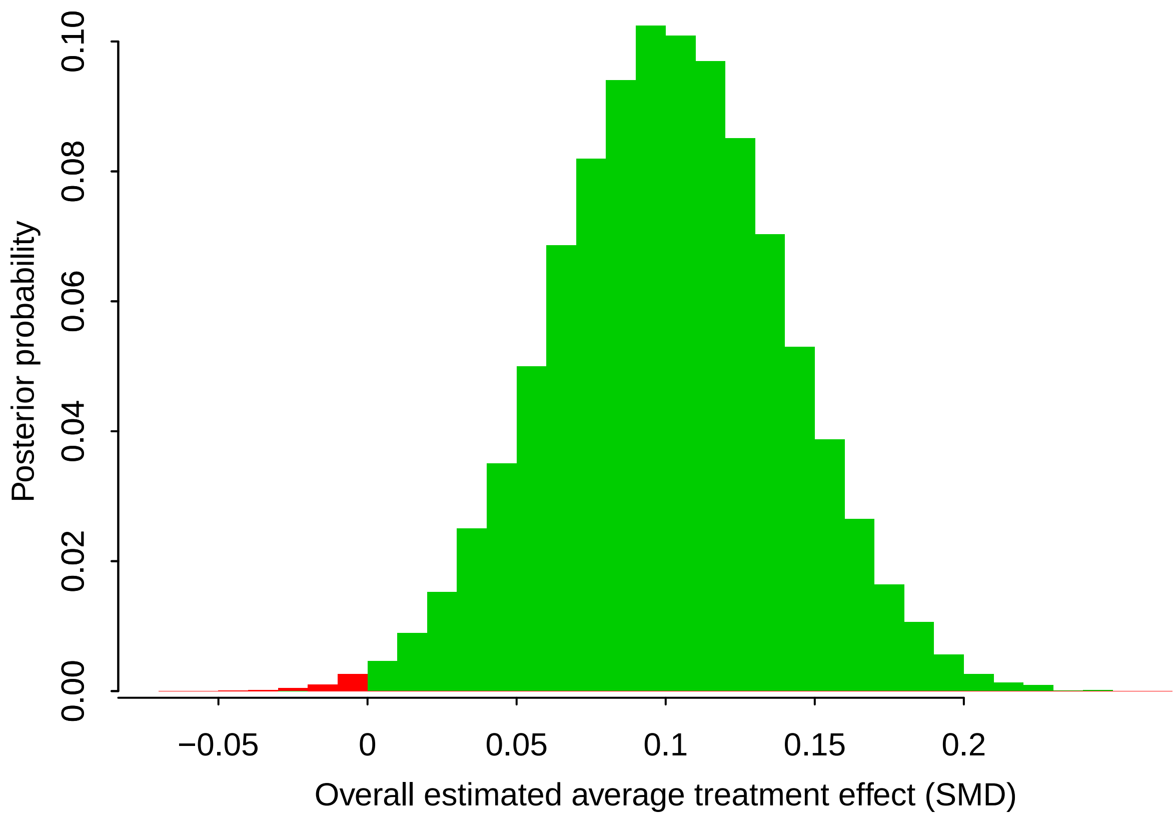

Figure S13. Posterior distribution of the average effect of special pedagogical assistance across all psycho-social difficulties in the study sample. The effect size is on the x-axis, the posterior probability of an effect size is on the y axis. Positive effects are displayed in green, negative effects in red. The posterior probability of a positive effect of SEA is 99.99%.

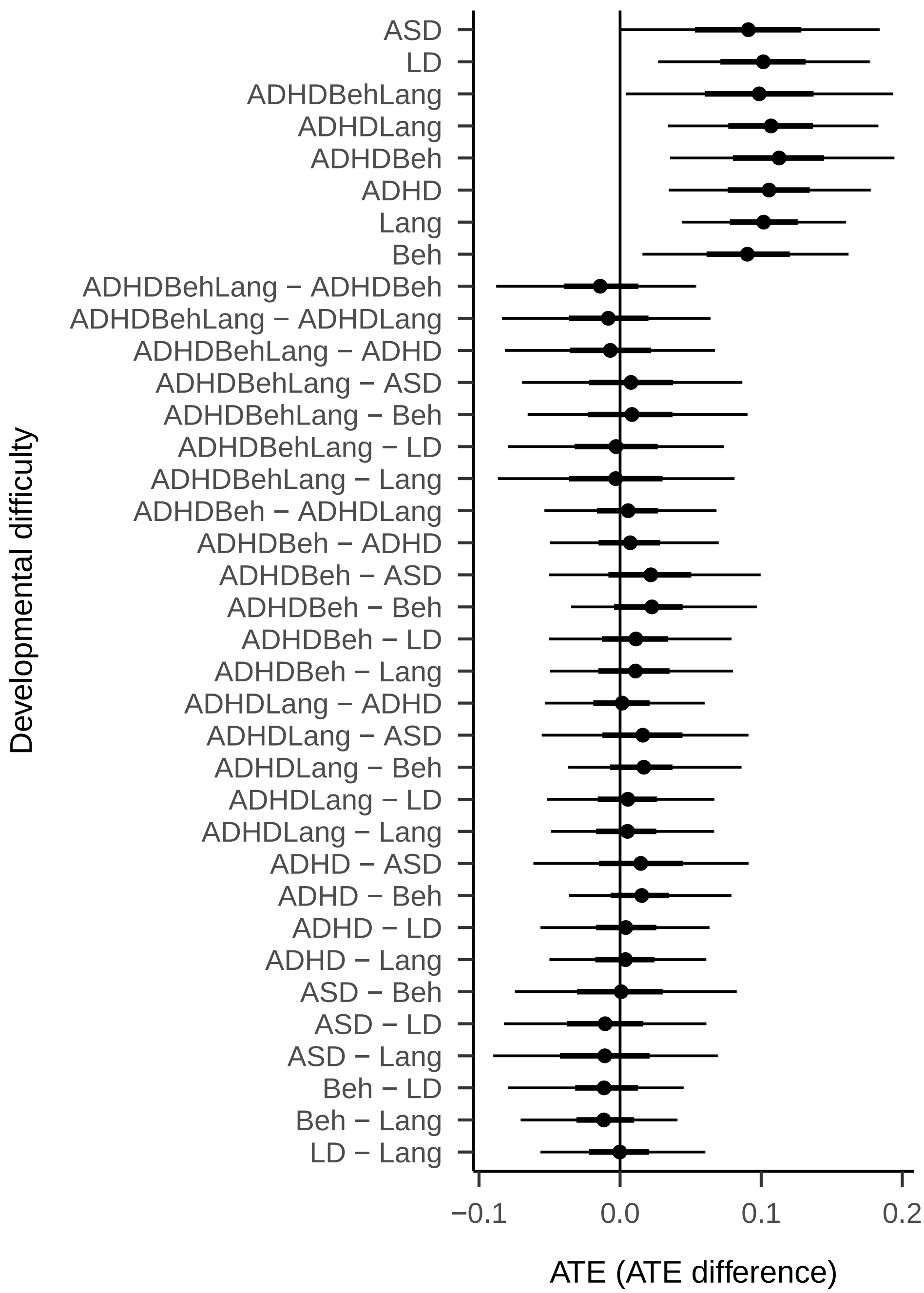

Figure S14. Estimated average treatment effects (ATE) and effect differences for children with different behavioral or developmental Problems. The top 6 rows show effects for the different groups (for each group averaged overall all psycho-social difficulties). The remaining rows show all pair-wise comparisons between groups. Dots are means, thick and thin lines are 50% and 90% credible intervals. Lines show 50% and 90% credible intervals.

Table S5

SEA effects and effect differences for children with different developmental difficulties.

	ASD	LD	ADHDBehLang	ADHDLang	ADHDBeh	ADHD	Lang	Beh
ASD	.091 (0,.184)	-.011 (-.082,.061)	.008 (-.069,.086)	.016 (-.055,.091)	.022 (-.051,.099)	.015 (-.061,.091)	-.011 (-.09,.069)	.001 (-.075,.083)
LD	-.011 (-.082,.061)	.102 (.027,.177)	-.003 (-.08,.073)	.006 (-.052,.067)	.011 (-.05,.079)	.004 (-.056,.063)	0 (-.056,.06)	-.011 (-.079,.045)
ADHDBehLang	.008 (-.069,.086)	-.003 (-.08,.073)	.099 (.004,.193)	-.008 (-.084,.064)	-.014 (-.088,.054)	-.007 (-.082,.067)	-.003 (-.087,.081)	.008 (-.065,.09)
ADHDLang	.016 (-.055,.091)	.006 (-.052,.067)	-.008 (-.084,.064)	.107 (.034,.183)	.006 (-.054,.068)	.001 (-.053,.06)	.005 (-.049,.066)	.017 (-.037,.086)
ADHDBeh	.022 (-.051,.099)	.011 (-.05,.079)	-.014 (-.088,.054)	.006 (-.054,.068)	.113 (.035,.194)	.007 (-.05,.07)	.011 (-.05,.08)	.023 (-.035,.097)
ADHD	.015 (-.061,.091)	.004 (-.056,.063)	-.007 (-.082,.067)	.001 (-.053,.06)	.007 (-.05,.07)	.106 (.035,.178)	.004 (-.05,.061)	.015 (-.036,.079)
Lang	-.011 (-.09,.069)	0 (-.056,.06)	-.003 (-.087,.081)	.005 (-.049,.066)	.011 (-.05,.08)	.004 (-.05,.061)	.102 (.044,.16)	-.012 (-.07,.04)
Beh	.001 (-.075,.083)	-.011 (-.079,.045)	.008 (-.065,.09)	.017 (-.037,.086)	.023 (-.035,.097)	.015 (-.036,.079)	-.012 (-.07,.04)	.09 (.016,.162)

Note. Numbers are mean ATEs (90% credible intervals). Group wise effects are on the main diagonal and pairwise comparisons on the off diagonals.

Table S6

SEA effects and effect differences for different outcomes.

	ATT	HYP	OPP	MOOD	ANX	COMM
ATT	.096 (.024,.168)	-.015 (-.069,.033)	-.024 (-.091,.03)	0 (-.053,.054)	-.005 (-.068,.054)	.015 (-.041,.078)
HYP	-.015 (-.069,.033)	.111 (.039,.184)	-.01 (-.066,.039)	.015 (-.034,.071)	.01 (-.051,.072)	.03 (-.029,.101)
OPP	-.024 (-.091,.03)	-.01 (-.066,.039)	.121 (.047,.198)	.025 (-.021,.083)	.02 (-.035,.082)	.04 (-.021,.117)
MOOD	0 (-.053,.054)	.015 (-.034,.071)	.025 (-.021,.083)	.096 (.025,.167)	-.005 (-.061,.045)	.015 (-.041,.078)
ANX	-.005 (-.068,.054)	.01 (-.051,.072)	.02 (-.035,.082)	-.005 (-.061,.045)	.101 (.026,.178)	.02 (-.038,.09)
COMM	.015 (-.041,.078)	.03 (-.029,.101)	.04 (-.021,.117)	.015 (-.041,.078)	.02 (-.038,.09)	.081 (.004,.156)

Note. Numbers are means ATEs (90% credible intervals). Outcome wise effects are on the main diagonal and pairwise comparisons on the off diagonals.

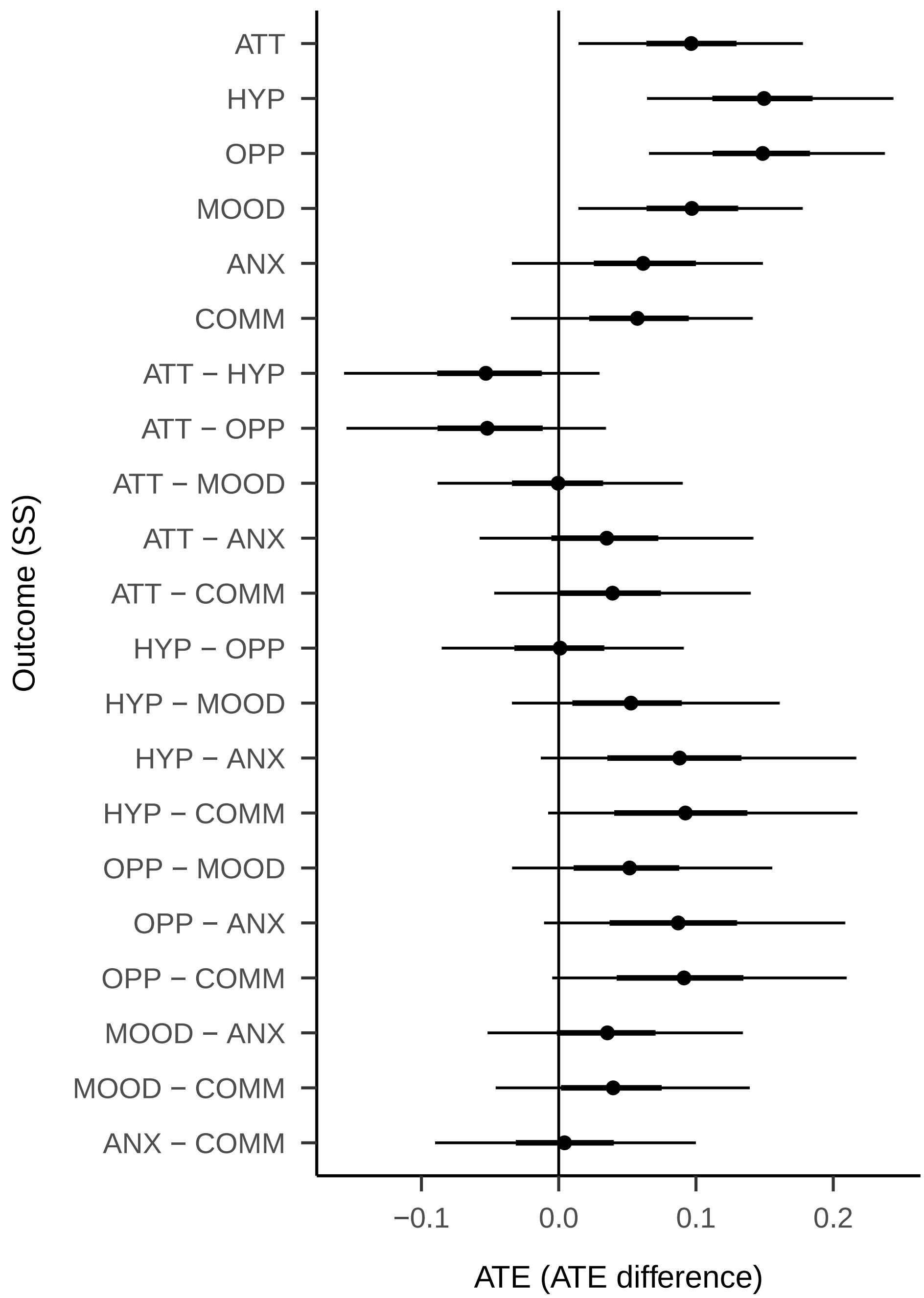

Figure S15. Estimated average treatment effects (ATE) and effect differences for different psycho-social difficulties. The top 6 rows show effects for the different psycho-social difficulties (for each difficulty averaged overall all DBPs). The remaining rows show all pair-wise comparisons between difficulties. Dots are means, thick and thin lines are 50% and 90% credible intervals.

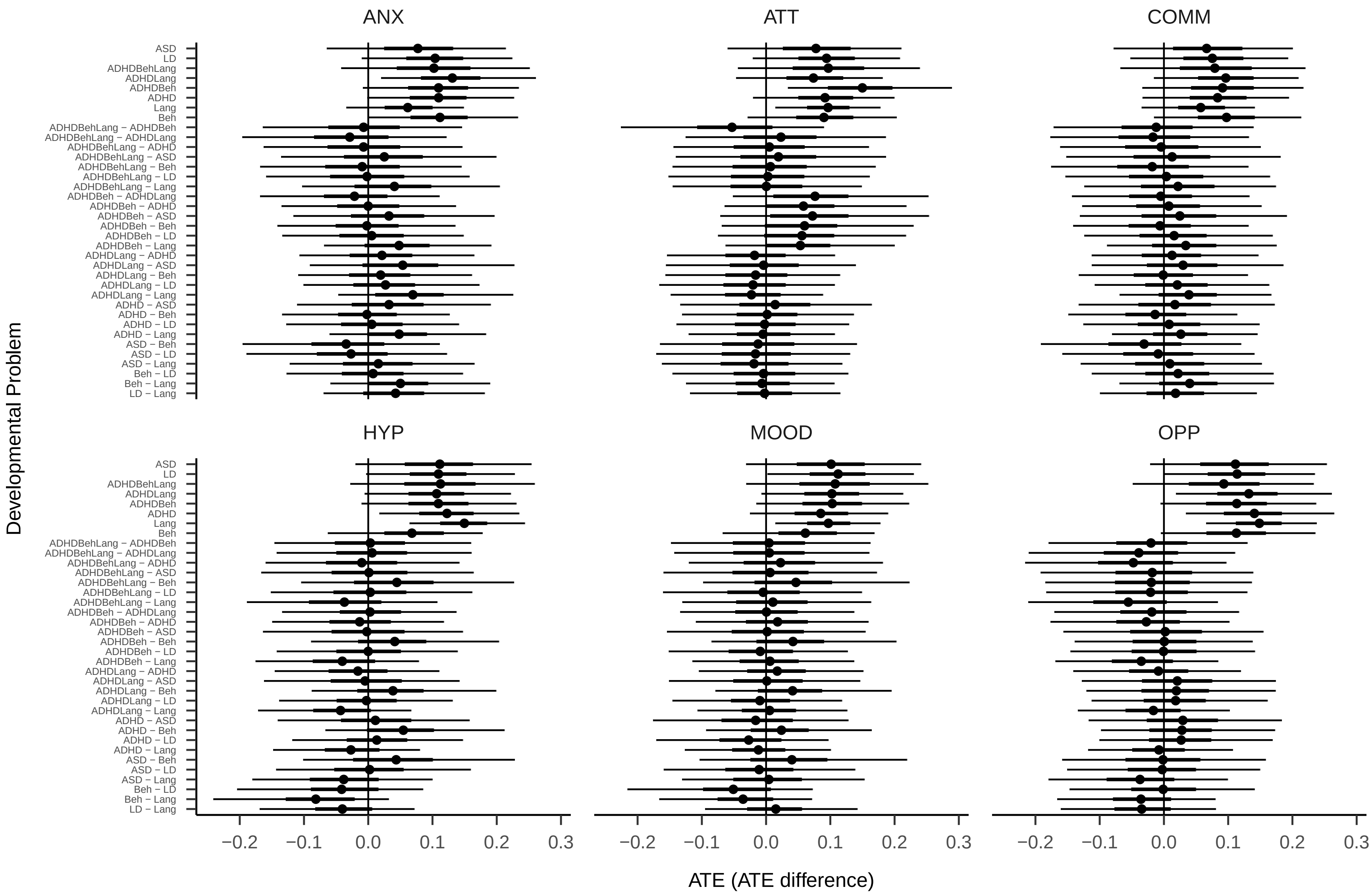

Figure S16. Estimated average treatment effects (ATE) and effect differences by group and outcome.

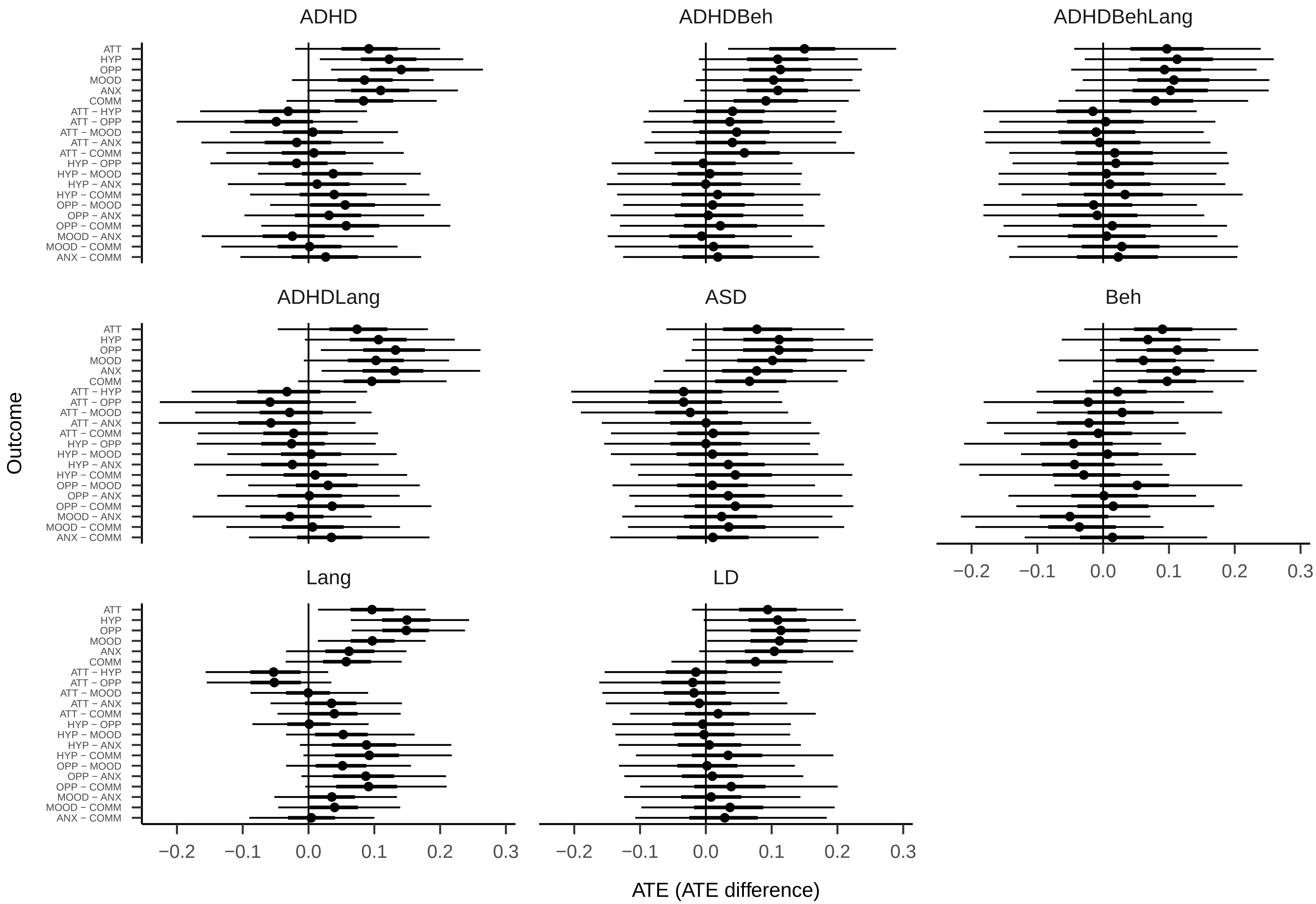

Figure S17. Estimated average treatment effects (ATE) and effect differences by outcome and group.

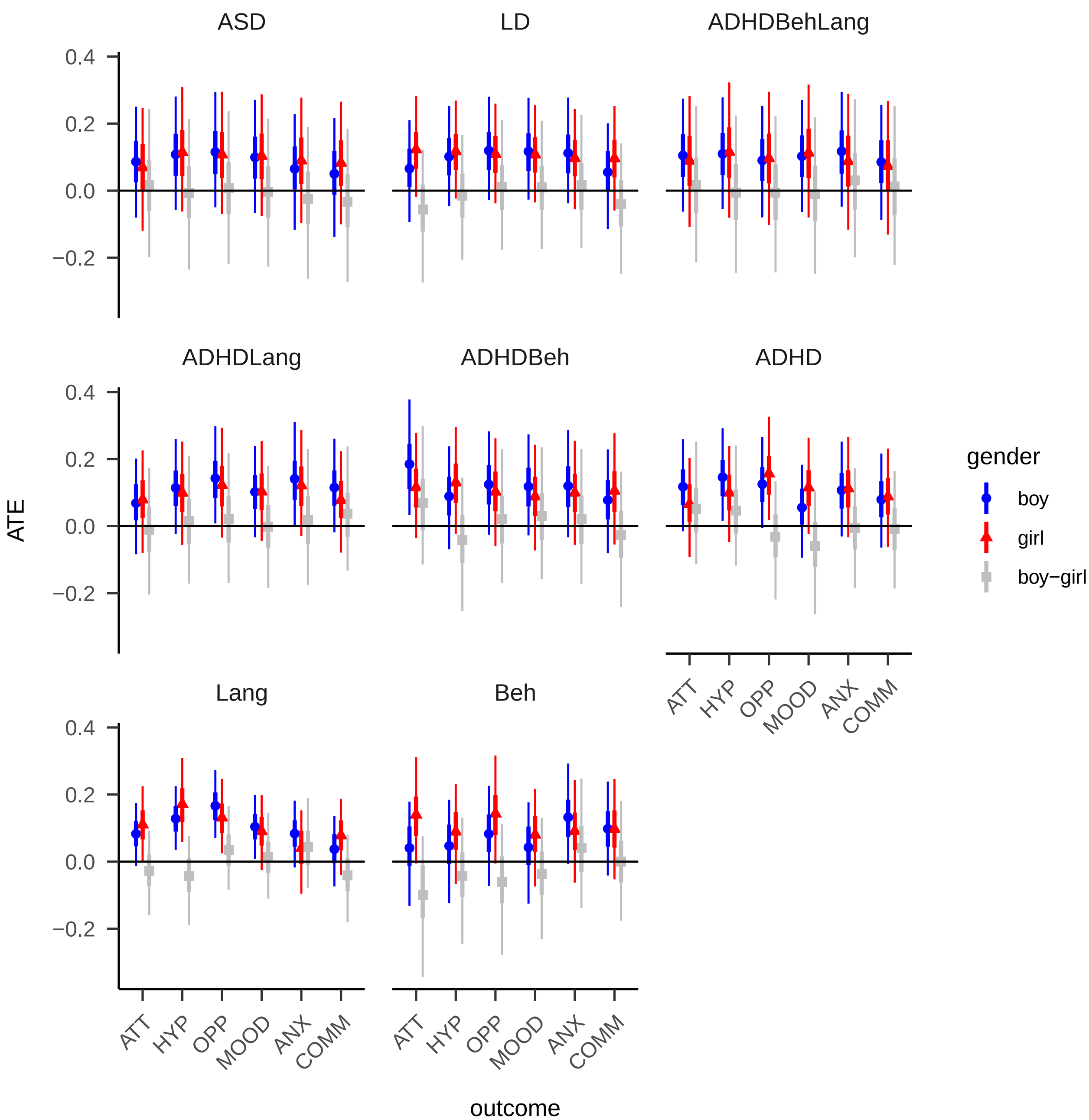

Figure S18. Estimated average treatment effects, split by developmental or behavioral problem, psycho-social difficulty, and gender. Dots are means, thick and thin lines are 50% and 90% credible intervals.

Supplementary References

1.

R Core Team (2017) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria

2.

Stanley D (2018) apaTables: Create american psychological association (APA) style tables

3.

Heinzen E, Sinnwell J, Atkinson E, et al (2018) Arsenal: An arsenal of ’r’ functions for large-scale statistical summaries

4.

Gabry J, Mahr T (2018) Bayesplot: Plotting for bayesian models

5.

Bürkner P-C (2017) brms: An R package for Bayesian multilevel models using Stan. Journal of Statistical Software 80:1–28. https://doi.org/10.18637/jss.v080.i01

6.

Bürkner P-C (2018) Advanced Bayesian multilevel modeling with the R package brms. The R Journal 10:395–411. https://doi.org/10.32614/RJ-2018-017

7.

Zeileis A, Hornik K, Murrell P (2009) Escaping RGBland: Selecting colors for statistical graphics. Computational Statistics & Data Analysis 53:3259–3270. https://doi.org/10.1016/j.csda.2008.11.033

8.

Stauffer R, Mayr GJ, Dabernig M, Zeileis A (2009) Somewhere over the rainbow: How to make effective use of colors in meteorological visualizations. Bulletin of the American Meteorological Society 96:203–216. https://doi.org/10.1175/BAMS-D-13-00155.1

9.

Dowle M, Srinivasan A (2018) Data.table: Extension of ‘data.frame‘

10.

al. AS et mult. (2019) DescTools: Tools for descriptive statistics

11.

Fox J, Venables B, Damico A, Salverda AP (2019) English: Translate integers into english

12.

Gohel D (2020) Flextable: Functions for tabular reporting

13.

Zeileis A, Croissant Y (2010) Extended model formulas in R: Multiple parts and multiple responses. Journal of Statistical Software 34:1–13. https://doi.org/10.18637/jss.v034.i01

14.

Wickham H (2016) ggplot2: Elegant graphics for data analysis. Springer-Verlag New York

15.

Kassambara A (2019) Ggpubr: ’ggplot2’ based publication ready plots

16.

Gordon M (2018) Gmisc: Descriptive statistics, transition plots, and more

17.

Harrell Jr FE, Charles Dupont with contributions from, others. many (2018) Hmisc: Harrell miscellaneous

18.

Gordon M, Gragg S, Konings P (2018) htmlTable: Advanced tables for markdown/HTML

19.

Long JA (2019) Jtools: Analysis and presentation of social scientific data

20.

Zhu H (2019) kableExtra: Construct complex table with ’kable’ and pipe syntax

21.

Xie Y (2015) Dynamic documents with R and knitr, 2nd ed. Chapman; Hall/CRC, Boca Raton, Florida

22.

Sarkar D (2008) Lattice: Multivariate data visualization with r. Springer, New York

23.

Bache SM, Wickham H (2014) Magrittr: A forward-pipe operator for r

24.

Gohel D (2021) Officer: Manipulation of microsoft word and PowerPoint documents

25.

Aust F, Barth M (2018) papaja: Create APA manuscripts with R Markdown

26.

Eddelbuettel D, François R (2011) Rcpp: Seamless R and C++ integration. Journal of Statistical Software 40:1–18. https://doi.org/10.18637/jss.v040.i08

27.

Eddelbuettel D, Balamuta JJ (2017) Extending extitR with extitC++: A Brief Introduction to extitRcpp. PeerJ Preprints 5:e3188v1. https://doi.org/10.7287/peerj.preprints.3188v1

28.

Stan Development Team (2018) RStan: The R interface to Stan

29.

Goodrich B, Gabry J, Ali I, Brilleman S (2018) Rstanarm: Bayesian applied regression modeling via Stan.

30.

Ooms J, Hester J (2019) Spelling: Tools for spell checking in r

31.

Stan Development Team (2018) StanHeaders: Headers for the R interface to Stan

32.

Terry M. Therneau, Patricia M. Grambsch (2000) Modeling survival data: Extending the Cox model. Springer, New York

33.

Buuren S van, Groothuis-Oudshoorn K (2011) Mice: Multivariate imputation by chained equations in r. Journal of Statistical Software 45:1–67

34.

Goodrich B, Gabry J, Ali I, Brilleman S (2019) rstanarm: Bayesian applied regression modeling via Stan

35.

Goodrich B, Gabry J, Ali I, Brilleman S (2018) Rstanarm: Bayesian applied regression modeling via Stan.

36.

Austin PC, Stuart EA (2015) Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. Statistics in medicine 34:3661–3679

37.

Bürkner P-C (2017) brms: An R package for Bayesian multilevel models using Stan. Journal of Statistical Software 80:1–28. https://doi.org/10.18637/jss.v080.i01

38.

Bürkner P-C (2018) Advanced Bayesian multilevel modeling with the R package brms. The R Journal 10:395–411. https://doi.org/10.32614/RJ-2018-017

39.

Gelman A, Rubin DB (1992) Inference from iterative simulation using multiple sequences. Statistical science: a review journal of the Institute of Mathematical Statistics 7:457–472

40.

Root A, Brown JP, Forbes HJ, et al (2019) Association of Relative Age in the School Year With Diagnosis of Intellectual Disability, Attention-Deficit/Hyperactivity Disorder, and Depression. JAMA pediatrics