Author's response to reviews

Title: Delayed school progression and mental health problems in adolescence: a population-based study in 10,803 adolescents.

Authors:

Wanda M. Tempelaar (w.tempelaar@umcutrecht.nl)
Christiaan P. Otjes (chrisotjes@gmail.com)
Clothilde J. Bun (CBun@ggdmn.nl)
Carolien M. Plevier (CPlievier@ggdmn.nl)
Willemijn A. van Gastel (W.vanGastel@umcutrecht.nl)
James H. MacCabe (james.maccabe@kcl.ac.uk)
René S. Kahn (R.Kahn@umcutrecht.nl)
Marco P. Boks (M.P.M.Boks@umcutrecht.nl)

Version: 4 Date: 2 June 2014

Author's response to reviews: see over
Dear editor,

Thank you for reviewing our manuscript entitled “Delayed school progression and mental health problems in adolescence: a population-based study in 10,803 adolescents”. We are grateful with the opportunity to revise and improve the manuscript.

Please find detailed discussion and alterations based on the comments made by the reviewers.

Yours sincerely,

Wanda Tempelaar, on behalf of all other co-authors

W. Tempelaar
MD, PhDstudent
Dept. of Psychiatry
University Medical Center Utrecht
Reviewer: Hugh Ramsay

Reviewer's report:

Thank you for the opportunity to review this interesting and well-written paper. I have a number of general and specific comments outlined below.

We thank the reviewer for this positive assessment.

Major compulsory revisions:

General comments:

1. This cross-sectional study design is not appropriate overall to examine for causality in the association between DSP and mental health problems and this should be acknowledged. While we know that DSP has arisen in the past, there is no information on the role of mental health problems themselves as a cause of DSP. It is therefore impractical to try to examine causal risk with this method. Rather, the method can examine DSP as a marker of risk.

   We agree with the reviewer that a cross-sectional design is not the appropriate design to study causal relationships. The manuscript was never intended to suggest this. The avoid confusion we further emphasized this limitation in ‘Strengths and Limitations’ at page 11 in the manuscript:

   “This study is also limited by its cross-sectional design. DSP and mental health problems are measured at the same moment, thus we cannot rule out that mental health problems could have preceded DSP earlier in youth.)

   Furthermore, we agree that the term ‘risk factor’ does not reflect this limitation and we are keen to avoid confusion. Therefore, we decided to change “risk factor” and “risk indicator” in the ‘Discussion’ (page 10) and ‘Conclusion’ (page 3, 12) to “risk marker” for clarification.

2. The study population is very heterogeneous in terms of age (and therefore types of mental health problems). This should be acknowledged in the limitations. Alternatively, it may be more appropriate to examine specific ages or specific mental health problems as the outcome of interest.
We thank the reviewer for pointing this out, we have added this to the Strengths and Limitations.'

Manuscript, page 11:

“Additionally, the included children are aged between 10 and 19 years old (99% between 12 and 18 years) and mental health problems may differ between the younger and older participants.”

3. There are significant problems with the way confounding is included in the methodology. Firstly, as mentioned above, the method cannot assess causality. Secondly, the number of confounders included is confusing and many are associated directly with the outcome of interest (e.g. school problems and behavior problems are features of conduct disorder). It would be more appropriate to examine DSP as an independent predictor of mental health problems and control for a smaller number of well-recognised (and separate) risk factors.

As addressed in the first comment, we agree that the cross-sectional design is a limitation of the current data and as a consequence causality cannot be inferred.

Regarding the handling of potential confounders we need to clarify on the approach we have taken. The literature on mental health problems in adolescence is loaded with papers on different risk factors. We consider it a strength of the current study that we have information on a lot of different (recognised) risk factors that could potentially influence the relationship between poor school achievement and mental health problems. In order to implement a data-driven approach whilst avoiding overfitting of our model we carefully evaluated the potential confounders and conducted the analysis in three consecutive steps. We selected only the covariates that were significantly correlated with the dependent and independent factor and significantly altered the association between poor school achievement and mental health problems in the second step. This resulted in a limited number of nine covariates, which were included in the final model. This approach ensures selection of relevant covariates for this sample and adds to our insight on risk factors for mental health problems in this group. We have better clarified this in the manuscript:

Manuscript, page 8-9:

"We used the large set of potential risk factors to perform a data-driven approach."
Therefore, we divided the analysis in three different phases. First, we analyzed the crude unadjusted association between DSP and current mental health problems. The crude odds ratio (OR) with 95% confidence interval (CI) was determined using logistic regression analysis. Secondly, covariates were tested as to whether they (1) were significantly correlated with both poor school achievement and mental health problems, and (2) altered the association between the DSP and mental health problems by at least ten percent [30]. The final step in the analysis was the composition of a predictive model using multivariate logistic regression adjusting for only the confounding factors that were obtained in step two of the analysis. Significance level was set as $p=0.05$.

**Specific comments:**

**Abstract:**

1. The methods section should clarify that SDQ was the outcome of interest, DSP the primary exposure of interest and that logistic regression was used to assess the association.

   *This is a good point, we have adjusted this in the abstract.*

**Manuscript, page 3:**

“The association of DSP with the SDQ was investigated using logistic regression with SDQ as outcome and DSP as primary exposure of interest while adjusting for socio-demographic characteristics, adverse life events, school-related factors, risk taking behaviour, healthy lifestyle and physical health.”

2. It appears excessive to conclude that DSP is a “risk factor” for mental health problems. It would be more appropriate to refer to it as a “risk marker”.

   *This is adjusted in the abstract.*

**Manuscript, page 3:**

*Conclusion: Delayed school progression is associated with general mental health problems in adolescence, and may be considered a risk marker for mental health problems. However the association is confounded by other factors making a causal relationship between impaired cognitive function such as poor scholastic performance and general mental health at adolescence less likely.*

**Introduction:**
3. It would be useful to more clearly explain how DSP could act as a risk factor for mental health problems. Is there evidence that DSP reflects cognitive function?

There is indirect evidence that scholastic achievement such as DSP reflects cognitive function. ‘Several studies showed that performance on intelligence tests is correlated with school achievement. On average, children with higher IQ scores do better on standardized achievement tests, have higher school grades, and complete more years of education. In other words, IQ scores often do predict school achievement, albeit imprecisely.’ (Educational Psychology, J.E. Ormrod, 2008, 155-156.) Furthermore, findings in a Dutch study confirm that a significant amount of the variation in school performance at Dutch secondary schools is explained by intelligence (Kort et al (2005). WISC-III-NL Handleiding. Nederlandse bewerking. Harcourt, NIP, Amsterdam.)

Moreover, repeating a grade is previously used as an indicator of scholastic underachievement, this is added to the text in the Introduction.

Manuscript, page 4:

“Repeating a grade is previously used as an indicator of scholastic underachievement and is for instance related to the risk of schizophrenia and other psychoses [14].”

4. It is important to justify the inclusion of confounders in the analysis at the introductory stage. Why were these confounders chosen for inclusion?

We agree with the reviewer that it is important to justify the inclusion of potential confounders in the analysis. Potential confounders were determined based on the literature and availability in the school survey that assessed poor school achievement and mental health problems. We added only the confounders to the final model based on the results of the first part of our analysis (significant correlation with the dependent and independent factor and significant alteration of the association between poor school achievement and mental health problems). The reason we have not included information on those confounders in the introduction is that we first wanted to present how we selected these particular confounders in the last model. Another option would have been to provide information in the introduction on all the potential confounders assessed, but this would be confusing. We therefore chose to mention the references to the literature in ‘Potential
confounders’ page 6-8 and made a selection of the most important or common covariates in our references since we were limited in the number of references. Apparently the method we described was not clear enough, therefore we have adjusted the ‘Statistical Analysis’ section by referring to the selection of potential confounders.

Manuscript, page 8:

“As described above, potential confounders were selected based on literature. We used the large set of potential risk factors to perform a data-driven approach. Therefore, we divided the analysis in three different phases.”

Study population and methods:

5. Is the sample a random representative sample? Does it reflect the Dutch population in terms of ethnic background, gender and socioeconomic status? (This is addressed later in the manuscript but may benefit from being stated earlier.)

We thank the reviewer for this remark, we have adjusted this in the ‘Results’ at page 9: “Gender was equally distributed. Most participants were Dutch, of high affluence and living with both parents.”

Measurements:

6. Has the definition of DSP used here been used elsewhere in the past?

Not exactly this definition of DSP, usually grade repetition is measured as parental report of the child having to repeat one or more grades. The advantage of the current measure is that it bypasses interviewing of the parents and can be routinely obtained. This definition identifies all pupils who repeated a grade at school regardless of reason.

Potential confounders:

7. Overall, the approach adopted of including numerous potential confounders confuses the reader and obscures the main message. Any confounder included should have a clear rationale for inclusion (be associated with both DSP and SDQ and not on the theoretical causal pathway) and should not be a mediating factor.
We do see this point and it echoes some of the previous comments. We refer to our previous response and further clarification in the manuscript (potential confounders, page 6-8). We also like to point out that our approach resulted in a limited number of nine covariates, which were included in the final model.

Manuscript, page 8:

“As described above, potential confounders were selected based on literature. We used the large set of potential risk factors to perform a data-driven approach. Therefore, we divided the analysis in three different phases.”

8. There is a good general rationale for including adverse life events and socio-demographic factors as potential confounders, though too many are included. I recommend reconsidering the inclusion of some factors (school level, family care, future perspectives on life and parental involvement). There are problems with including school-related variables and risk-taking behaviour. Some of these overlap with SDQ questions and concepts, particularly those related to conduct problems. Inclusion could therefore mask associations between DSP and SDQ general risk. I also fail to see the reasoning for inclusion of health and lifestyle factors as potential confounders. They further complicate an already long list of confounders.

We agree that some of the investigated potential confounders overlap with items from the SDQ, but the confounders included in the final model do not; they are mostly adverse life events (sexual abuse, molestation by parents, violence between parents, parental problems with alcohol / addiction) or socio-demographic factors.

Discussion:

9. As discussed above, the first paragraph of “Strengths and limitations” suggests that the method gives the possibility of a longitudinal perspective. However, this is not the case as there is no way to examine the outcome of interest longitudinally. Mental health problems commonly persist over time and this DSP may simply be associated with persistence rather than incident problems.

As pointed out, we acknowledge that our data does not allow causal inference. We have scrutinized the first paragraph of the ‘Discussion’ to
avoid suggestion of a longitudinal perspective.

Manuscript page 10:

"In this cross-sectional study we investigated the association between scholastic achievement and current mental health problems in adolescents. We demonstrated that delayed school progression (DSP), used as a proxy for poor school achievement, is associated with adolescent mental health problems."

10. As discussed above, it should be acknowledged that SDQ is a highly heterogeneous outcome across such a large age range. The role of DSP in younger adolescents with ADHD may be quite different to older adolescents with emotional problems.

We thank the reviewer for this comment, this issue is added to the 'Discussion' at page 11: “Additionally, the included children are aged between 10 and 19 years old (99% between 12 and 18 years) and mental health problems may differ between the younger and older children.”

Minor Essential Revisions:

Introduction:

1. Please provide a reference for how you divided the SDQ total difficulties score into normal and abnormal.

A reference is added to the manuscript, page 6:

“The SDQ total difficulties score can be divided into normal (0–15), borderline (16–19) and clinical scores (20–40) [16]. To identify those with clinically relevant symptoms, the outcome was dichotomized into normal to borderline and clinical.”

2. Table 1: please explain SDQ, SES

Explanations are added to the table.

Discretionary Revisions:

Introduction:
1. I would rephrase the term “elaborate” in the final line.

_This term is omitted:_ We therefore studied the association between a delay in school progression (DSP) and mental health problems in adolescents, and adjusted for potential confounders.

**Results:**

2. Line 2 would be clearer written “14.2 (SD=1.6) years and 49.8% male.” 3. Line 5 should read “were equally distributed by age, gender and ethnicity.”

_Good suggestions; it is changed in the manuscript, page 9._
Reviewer: Alexis Cullen

Reviewer's report:

This cross-sectional study investigates the relationship between delayed school progression and mental health problems in a general population sample of adolescents. The sample size is sufficiently large to be able to explore these associations (N=10,803) and a range of potentially confounding factors were examined. The results indicate that whilst delayed school progression is associated with mental health problems this relationship is rendered non-significant after adjusting for sociodemographic factors and life events.

We thank the reviewer for this positive assessment.

Minor revisions are recommended:

Minor essential revisions:

1. **Abstract** > Background: whilst I recognise that the authors have obtained DSP data for the entire school period (i.e., rather than current DSP only) the use of the term 'subsequent' here implies that this is a longitudinal study; although this is clarified in the following line I think it would best to omit this term.

   We thank the reviewer for this suggestion, the term ‘subsequent’ is omitted:

   Manuscript page 3:

   "This study investigates the association between scholastic underachievement and general mental health problems in adolescence, using delay in school progression (DSP) as a marker of poor scholastic performance."

2. **Measurements** > Mental health problems: regarding the validity of the Dutch SDQ, can the authors please indicate the age range here (i.e., has it been validated in participants as old as 18 – if not worth mentioning in the limitations).

   The SDQ is previously used in adolescents aged up to 18 years old, for instance: The Strengths and Difficulties Questionnaire (SDQ): the factor structure and scale validation in U.S. adolescents. He JP, Burstein M,
3. **Measurements** > Poor school achievement: In the Discussion the authors state that a stringent cut-off for DSP was applied to include only children with a definite history of scholastic underperformance. Can the authors clarify whether this is additional to the procedure described in the Methods (i.e., using individual age 1.5 yrs greater than mean grade age); if so it should be included in the Method.

*No, this is not additional, we used the procedure as described in the Methods and this is a stringent cut-off for DSP. To clarify we have changed the wording:*

**Manuscript page 6:**

“This stringent cut-off allowed us to identify adolescents who had failed to progress as expected to the next grade on at least one occasion at any stage of primary or secondary school [19].”

4. **Results:** paragraph 1: 'Non-Dutch adolescents also perceived significantly more academic difficulties', please rephrase as this is unclear. Does this mean that non-Dutch ethnicity was also associated with DSP?

*We thank the reviewer for pointing this out. Indeed it means that non-Dutch ethnicity was associated with DSP. This is rephrased in the text.*

**Manuscript page 9:**

“Adolescents with a non-Dutch ethnicity had higher odds for DSP (OR 2.77, OR 95% CI 2.27 – 3.73).”

5. **Results:** I can't seem to find a statement regarding the prevalence of DSP in this sample? This is important as it sheds light on the possible reasons as to why a significant association was not observed (i.e., if DSP is actually quite common then perhaps it is not surprising that it is not a more robust marker of mental health problems) and is also relevant when considering the generalisability of the findings.

*We thank the reviewer for this comment. The prevalence of DSP has been*
added to the Results.

*Manuscript, page 9:* “Of the pupils 35.4% were identified as delayed in their school progression.”

The prevalence in this study is representative of the Dutch population of pupils at school, which is on average 30%. (G. Juchtmans, B. Belfi, B. de Fraine, M. Goos, H. Knipprath, A. Vandenbroucke en B. Verbeeck red, Samen tot aan de meet. Garant, 2011)

6. **Discussion** > first paragraph: similar issue to the abstract, whilst the first sentence confirms that this is a cross-sectional study, I think the use of ‘prior poor school achievement’ and ‘adolescent-onset’ overstates the case somewhat.

In reality the SDQ assesses problems over the past six months and so presumably overlaps with the assessment of DSP (i.e., the authors did not state that they excluded DSP occurring during the past six months). Additionally, there is no way of confirming that the adolescents examined did not have problems during early childhood and so ‘adolescent-onset’ is not accurate. I would recommend removing these terms.

We thank the reviewer for these suggestions, we have adjusted the first paragraph of the Discussion and removed the terms ‘prior’ and ‘onset’:

*Manuscript page 10:*

"In this cross-sectional study we investigated the association between scholastic achievement and current mental health problems in adolescents. We demonstrated that delayed school progression (DSP), used as a proxy for poor school achievement, is associated with adolescent mental health problems.”

7. **Supplementary table:** the table title indicates that these data relate to the analyses examining the association between all variables and DSP, however, the table heading says ‘Association SDQ’. Please clarify which is correct.

We thank the reviewer for pointing this out, we accidentally wrote ‘SDQ’ in the table heading instead of ‘DSP’, this is changed.

*Discretionary revisions:*
1. **Abstract** > Results: the abstract would be strengthened if the authors could clarify which potential confounders were included in the analysis, at least just in vague terms (e.g., sociodemographic characteristics and life events). Additionally, I would argue that the use of the term ‘risk factor’ in the Conclusions section of the abstract is too strong given the cross-sectional nature of the study, perhaps ‘marker’ would be more appropriate.

*We thank the reviewer for this comment. We have adjusted the abstract.*

**Manuscript, page 3:** “After adjusting for other risk factors (socio-demographic factors and life events) in a logistic regression model the association between DSP and mental health problems was attenuated (OR 1.33, 95% CI 0.86 – 2.05).”

*Furthermore, the term ‘risk factor’ has been rephrased to ‘risk marker’ in the conclusion.*

2. **Potential confounders**: would suggest rewording ‘.....the school survey addressing both poor school achievement and mental health problems’ to ‘......the school survey which assessed both poor school achievement and mental health problems’ or something similar in order to clarify that this refers to what was assessed in the school survey (and not the literature).

*This is changed in the manuscript, page 6:*

“Potential confounders were determined based on the literature and availability in the school survey which assessed poor school achievement and mental health problems.”

3. **Results**: I think it would be worth reiterating the specifics of the analyses in the results section. For example, in paragraph 2 it would be helpful to clarify that the potential confounders examined in the adjusted model were those that were associated with both exposure and outcome and additionally led to a 10% change in the association between DSP and SDQ problems.

*We thank the reviewer for pointing this out. This is added to the manuscript, page 9:*

“We selected only those confounders that were correlated with DSP and SDQ and altered the association between DSP and SDQ.”
Minor issues not for publication:

1. **Abstract** > Method > line 3: ‘....was used to assess mental health problems and their associations’ makes it sounds as though multiple outcomes were examined when in fact only one outcome (i.e., SDQ abnormal score yes vs. no) was examined. Perhaps rephrase to make clearer.

   *We thank the reviewer for this suggestion and we rephrased the sentence.*

   **Manuscript, page 3:** “Participants completed the Strengths and Difficulties Questionnaire (SDQ) to assess mental health problems.”

2. **Introduction** > paragraph 3 > line 3: please change ‘school achievement as independent risk factor’ to ‘school achievement as an independent risk factor’.

   *This is changed at page 4.*

3. **Study population and methods** > paragraph 2: The authors state that for the purposes of the study the group was divided into those preparing for polytechnic or academic training and those focussing on vocational training. These groups are not discussed again; however, a ‘low level of education’ was associated with SDQ problems (Table 1) which presumably relates to the vocational training group? Can the authors please employ the same group labels in the methods and results to make this clearer.

   *We thank the reviewer for pointing this out, indeed a low level of education relates to the vocational training group. We changed the group labels in Table 1 and in the appendix to vocational education.*

4. **Results** > paragraph 1 > last line: this should be changed to ‘mental health problems’.

   *This is changed at page 9.*

5. **Table 2**: please confirm that the results relating to each potential confounder are after adjustment for all other confounders (i.e., that all potential confounders were entered together).
All potential confounders were entered together in the logistic regression model. We added this to the manuscript, page 9: “In the adjusted model all the potential confounding factors were entered together.”

6. **Supplementary file**: for school level (e.g., ‘LWOO’) please spell these out in full.

*Low educational level in the supplementary file is changed to ‘vocational education’.*

7. **Supplementary table**: ‘Harddrugs use last month’ please change to ‘Hard drug use last month’.

*This is changed.*
Reviewer: Sarah Sullivan

Reviewer's report:

Minor Revisions

1/ Abstract - school-related what?

*We thank the reviewer for pointing this out; it should have been ‘school related factors’, this is changed in the manuscript.*

2/ Methods - confounders - school level what?

*‘Educational level at secondary school’, this is changed in the manuscript.*

3/ family care, future perspective on life and parental involvement would not usually be classed as demographics

*We agree with the reviewer, we attempted to avoid more categories but have now categorized these as miscellaneous.*

Manuscript, page 8:

"Miscellaneous factors
Family care, future perspective on life and parental involvement."

4/ Results line 1 - was should be were

*This is changed.*

5/ Results line 8 distribution OF variables

*This is changed.*

6/ Results Para 3 - reword "In the adjusted model............" this sentence doesn't make sense to me

*We thank the reviewer for this comment, the sentence is rephrased in the*
manuscript, page 9-10:

“In the adjusted model all the potential confounding factors were entered together. The adjusted logistic regression model resulted in a non-significant association between DSP and clinical SDQ (OR 1.38, 95% CI 0.91-2.10).”

Major Compulsory Revisions

1/ The conclusion section of the abstract needs to be revised. After adjustment for confounders the 95% confidence interval includes 1 - this means that now DSP could be associated with either a reduction in mental health problems or an increase.

We thank the reviewer for this comment, we certainly do not want to overstate the conclusions and we have adjusted the abstract.

Manuscript, page 3:

“Delayed school progression is associated with general mental health problems in adolescence, but this relationship is heavily confounded by other factors. A causal relationship between impaired cognitive function such as poor scholastic performance and general mental health at adolescence is less likely and delayed school progression may merely be considered an indicator of risk for mental health problems.”

2/ A couple of comments about the Introduction - since the investigation is about adolescent mental health I feel that the Introduction should only review literature relevant to this - there should be no mention of childhood MH.

We did not intend to report on childhood mental health problems but we first mention the debilitating effects of mental health problems in children and adolescents (reference 1) and stress the importance to study this topic. The second paper we refer to (on the prevalence of mental health problems) concerns both children and adolescents.

3/There is a study missing from this literature review - Hameed et al 2013 Sz Res Vol 145 which investigates childhood literacy and later psychotic experiences - this may be worth mentioning here. It is also longitudinal and so would be important to acknowledge it. In fact, at the risk of self-promotion, there is also another paper Sullivan et al 2013 where I investigated one domain of the SDQ in childhood and adolescent psychotic experiences - this is in Cog Neuro 2013. Both of these papers should also be mentioned in the conclusion.
We agree with the reviewer that both papers are very interesting and relevant to the topic of this paper. We added the study written by Hameed et al to the Introduction (page 4) and the Conclusion, page 12.

Furthermore, we do appreciate the suggestion of the paper on the longitudinal association between social functioning and theory of mind in first-episode psychosis. This is indeed a fascinating topic. However, since this paper specifically concerns the domain of social functioning as outcome variable we did not include it in our study because we chose to focus on broader mental health problems (and did not look at separate domains of the SDQ, also mentioned at remark no 10).

4/ I think the non-responders are crucial and that this should probably be better acknowledged in the paper, although the response rate was generally good. According to the Methods - 16% of all students did not respond and this was because of absence due to illness or truancy - these students are also probably more likely to have had DSP and MH problems - therefore the association may have been slightly stronger if it had been possible to include them.

We have added this point to the strengths and limitations section at page 10: “The response rate was generally good. Moreover, the association between DSP and mental health problems might have been stronger when non-responders would have joined the study since truants or frequently ill students are probably more likely to have high SDQ scores and problems with scholastic achievement.”

5/ Use of the SDQ as dichotomous - I am aware that this is sometimes done but there needs to be a reference included here to justify this cut off - this is also a crucial point - where this cut off comes can make extreme differences to the finding and therefore needs to be justified somehow. I am not sure why the SDQ was dichotomised in this way because of the loss of so much information - perhaps it would be wise to also present the findings using a continuous variable as well.

We have also discussed this point thoroughly and we agree with the reviewer that generally a continuous variable provides more power than a categorized variable. However, this study was designed from a clinical perspective and one of its aims was to investigate easily measurable markers in children that contribute to the risk of mental health problems or to the detection of mental health problems. We therefore chose to use a dichotomized score on the SDQ, since this can be used in clinical practice and has been used in several studies now (Van Gastel 2013, Nijs 2014, Rothon 2009, Havas 2009). We did not
present the results from a linear regression model with a continuous SDQ outcome variable in this paper to avoid confusion.

6/ Potential confounders - the issue of confounding or mediating is a crucial one but it is more a case of interpretation rather than caution as stated here - after all there is no statistical difference - more it depends on our interpretation of the associations between these variables

We agree with the reviewer that this sentence causes confusion, we therefore removed it in the manuscript.

7/ Adverse life events - if these were self-report (and that is not really clear) would a 12 year old child know about a parental alcohol problem or domestic violence for instance?

The mentioned adverse life events are indeed self-reported. We agree that information on parental addiction from youth is not/hardly a reliable measurement of parental disorders but we considered it still interesting to include this information.

8/ Statistically is not a good idea to include a large number of confounders in logistic regression models - where as it is more acceptable to do this in linear regressions. See point 5 above.

We agree with the reviewer that linear models are more robust to overfitting. However, as the current analysis are conducted in a sample of over 10,000 participant, the logistic model is well capable of handling the current number of covariates. Moreover, we took particular care to include only those covariates that potentially can confound the relationship of interest.

9/ Generally it is now poor practice to decide on the inclusion of a set statistically significant level - p values should only be used in conjunction with confidence intervals - especially with a sample this large. Very small associations and changes will be significant but also meaningless.

We appreciate the point about the relevance of p-values in large epidemiological samples. In the current selection of covariates this is tackled by the criterion of over ten percent change in the coefficient (Beta).
10/ There may be merit in looking at the individual domains of the SDQ (as continuous variables) which would give more information on the association between DSP and different kinds of MH problems i.e. peer problems, emotional problems, conduct disorder etc - these associations may also be different in boys and girls - which would also be interesting - i.e. DSP in boys may be associated with externalising disorders and with internalising in girls?

We thank the reviewer for this interesting suggestion, but in this paper our primarily focus was on general mental health problems and not on specific traits. We choose not to go into detail in post-hoc analyses to keep our main message clear.

11/ Again the use of non-significant is not appropriate here. Also the assumption that the interpretation of confounding/mediating is not necessary because of the non-significance of associations.

We thank the reviewer for this remark, the wording of 'non-significant' was not appropriate and we have adjusted this in the manuscript.

Manuscript, page 10:

‘In the adjusted logistic regression model the association between DSP and clinical SDQ was not statistically significant (OR 1.38, 95% CI 0.91-2.10). Since the association is not longer significant, analysis of factors that might also act as mediating factors is redundant.’

12/ Would be interesting to discuss here the most influential mediators/confounders i.e. looks like it is economic factors?

This is an interesting idea, economic factors show the largest odds ratio in the multivariate model. However, in this paper we chose to concentrate on the relationship between scholastic achievement and mental health problems and we did not focus on other factors that may be also important to study by themselves.

13/ Discussion - first paragraph - we don't actually know that DSP is associated with MH because of the importance of the confounders or mediators - neither would we ever know from this study whether they are causally associated because of the cross-sectional design - this paragraph needs to be re-worded.
The last sentence in this paragraph is strange and I think not necessary

We thank the reviewer for these comments, we rephrased the first paragraph in the Discussion and deleted the last sentence.

14/ Interpretation of findings para 2 - elaborate more to make your meaning clear

We have adjusted this paragraph, manuscript page 12:

“In addition, poor school achievement can be seen as an event that might mediate the impact of other risk factors for mental health problems. Thus, poor school achievement might function as a trigger for alterations in the causal pathway of genetic and environmental factors underlying neurobiological changes leading to mental disorders.”

Tables

I didn't understand some parts of Table 1- there seemed to be some missing variables here i.e. future perspective, family care etc.

We thank the reviewer for these useful comments. Future perspective, family care and parental involvement did not alter the association between the DSP and mental health problems by at least ten percent and were therefore left out of the first table to keep it more accessible.

Didn't understand the N ratio in the brackets in the variables column.

We adjusted the table with the number of missing in the variables column.

The column headings also were not clear - i.e. Association SDQ OR.

The column headings are adjusted.

It would be better practice to put the p values in the table rather than have asterisks to indicate those above 0.05 (which is too large for such a large sample anyway)

We chose to report the confidence intervals of the associations to indicate the
reliability of the OR. Moreover, to keep the table accessible we used asterics to indicate in what range the p-values lie.

Table 2 - what is Low Age and how was this derived? It is interesting that a couple of the confounders look as if they may be protective

Low age is 14 years and younger. Indeed it seems that low age and Dutch ethnicity seem to be protective.

We have clarified in the manuscript page 7:

“Several socio-demographic factors were measured: gender, age (defined as low: under 14 years old), immigrant status, marital status of parents and socioeconomic position [20] and urbanization [21].”

There are a lot of unexplained abbreviations in the Appendix

These are adjusted.