# Using the Orchestrating Numeracy and the Executive (THE ONE) Programme to improve maths attainment, a two-armed cluster randomised trial

Further Appendices

July 2025

Anna Brian, James Merewood, Elena Rosa Speciani, Fin Oades, Emily MacLeod, Merrilyn Groom

The Education Endowment Foundation is an independent charity dedicated to breaking the link between family income and education achievement. We support schools, nurseries and colleges to improve teaching and learning for 2 – 19-year-olds through better use of evidence.

We do this by:

- **Summarising evidence.** Reviewing the best available evidence on teaching and learning and presenting in an accessible way.
- **Finding new evidence.** Funding independent evaluations of programmes and approaches that aim to raise the attainment of children and young people from socio-economically disadvantaged backgrounds.
- **Putting evidence to use.** Supporting education practitioners, as well as policymakers and other organisations, to use evidence in ways that improve teaching and learning.

We were set-up in 2011 by the Sutton Trust partnership with Impetus with a founding £125m grant from the Department for Education. In 2022, we were re-endowed with an additional £137m, allowing us to continue our work until at least 2032.

For more information about the EEF or this report please contact:

Education Endowment Foundation
5th Floor, Millbank Tower
21–24 Millbank
SW1P 4QP

info@eefoundation.org.uk

www.educationendowmentfoundation.org.uk

Department for Education

What Works Network

# Contents

# Appendix F: Interview Schedule (Trainers)

Section 1: Fidelity and adaptation

1. How many of The ONE settings have you worked with?

a. What did this involve? E.g., training, ongoing support

2. To what extent did you feel you had all the preparation you needed to be able to work with the settings?

a. Regarding both the in-person training and the follow-up calls?

b. If yes, what helped you to feel confident?

c. If not, what additional preparation might have helped?

3. What resources and training were you provided with to support with providing training (both in-person and the follow up calls)?

4. To what extent did you feel that you had all the resources and training you needed to be able to deliver The ONE confidently?

a. If not, what additional resources or training might have helped?

b. Where there any resources or training that were particularly helpful?

5. Did you encounter any barriers to delivering training and support for The ONE (for example tech challenges or other logistical difficulties)?

a. Were these one-off challenges, or recurring?

b. To what extent were you able to overcome these challenges?

6. To what extent did you have sufficient time to support the settings with training and delivery?

a. Was a full 30 minutes made available for each training session?

b. Was four 30-minute sessions enough to cover all the planned content in sufficient detail without rushing?

c. Did you have the time you needed to make your follow-ups?

7. To what extent did you make any adaptations to the training materials while delivering them?

a. Did you make any adaptations to your delivery based on participants' existing knowledge? (e.g. did you spend more time on new or more difficult concepts, or speed up for more familiar ones?)

b. Did you make any adaptions for practical reasons, e.g. lack of materials, time pressures etc.?

c. Did you add any additional content/information to that included in the slides and information sheet (for example additional tips, anecdotes etc.)?

d. Did you make any resources to model the content?

8. Is there anything that you would consider changing about The ONE if you had the chance?

a. Is there any content you would add or remove?

b. Are there any concepts you'd present in a different way?

c. Is there anything you'd change in terms of format (for example the number or length of sessions, the degree of interactivity of sessions)?

Section 2: Programme differentiation

9. How knowledgeable do you feel participants were about the concepts you were training them on (e.g., maths, embedding maths activities, adapting activities, and executive functioning)?

a. Were they familiar with the concepts presented, or was the content all new to them?

b. Was their level of prior knowledge as you expected? Was it higher or lower?

c. Did you note any significant differences in participants' knowledge across different settings?

10. How do you feel the content of the sessions was received by participants?

a. What activities were most used/well received? Were these the ones you gave examples of, or ones that were built upon things they were already doing?

b. Do you feel there were any components or aspects of The ONE that were particularly challenging for participants?

c. Did participants appear engaged in The ONE? Did you experience any challenges in encouraging participation?

11. Is there anything else that you'd like to add about your experience of delivering training on The ONE?

# Appendix G: Interview Schedule (Practitioners)

Section 1: Delivery of the intervention

1. Before I start can I confirm that you are a practitioner at X setting?

2. How have you found delivering the ONE overall?

a. How many staff were involved in your setting?

b. How many children/how many rooms?

c. Are there any aspects that are particularly challenging?

d. Is there anything you have found particularly helpful for delivering the activities?

e. How easy or difficult have you found it to include at least three of the activities per week as part of your normal play and learning?

3. Were there any surprising benefits or consequences of delivering the ONE intervention in your setting?

a. For example 'spill over' into other areas, or less time to spend on other aspects of the curriculum, like English?

4. Were there any other challenges you came across in implementing the ONE in your setting?

5. Were there any other factors that you felt made it easier to implement the ONE in your setting?

6. In your view, what further support could have improved the delivery of the ONE in your setting?

7. How have you tended to fit the activities into your day-to-day routine?

a. For example as small group activities, one-to-one, in outdoor play or free play?

b. Do different children engage with the activities to different extents, or do all children participate roughly the same amount?

c. Have you noticed any differences in the impact of the activities on children depending on how much they engage with them?

8. Do you feel that the activities are pitched at the right level for the children involved?

a. Why/why not?

9. Are there any groups of children that you feel particularly engage with the activities or, on the other hand, that struggle to engage, for example children with English as an additional language, children with SEND or children from disadvantaged backgrounds?

a. Are there any particular barriers to children with EAL engaging with the activities, or anything that can help them to engage?

b. Are there any particular barriers to children from disadvantaged backgrounds engaging with the activities, or anything that can help them to engage?

c. Are there any particular barriers to children with SEND engaging with the activities, or anything that can help them to engage?

d. (to what extent) did you overcome these barriers? Were these barriers (and overcoming them) something that you discussed with the trainer?

10. Have you found it necessary to adapt the material provided or instructions provided to better meet a child's needs at any point?

a. If so, why and how?

b. If not, have there been changes you wish you could make but feel you can't?

Section 2: Training and support

1. Did you attend the four professional development sessions?

a. If so, how did you find them? Did you feel they prepared you well to deliver the activities?

b. If not, how did you learn about delivering the activities? Did you feel well-prepared?

2. Did you receive any additional support to deliver The ONE?

a. If so, how did you find them? Did you feel they prepared you well to deliver the activities?

b. If not, how did you learn about delivering the activities? Did you feel well-prepared?

3. To what extent (if at all) do you feel The ONE training, support, and resources improved:

a. your knowledge of executive functions and their relationship with early numeracy?

b. your ability to incorporate play-based maths activities into your daily routine?

c. your ability to adapt activities to the right level for different children?

4. Is there anything about the support you received to deliver The ONE, that you would suggest changing?

a. Is there any content you would have liked the trainers to cover that wasn't included, or anything they should have covered in more/less depth?

b. Was the number and length of training sessions about right, or would more/less training have been helpful?

5. To what extent (if at all) has participating in The ONE affected your ability to participate in other professional development?

a. Did The ONE training crowd out your other professional development activities?

b. Did The ONE complement, or work alongside, your other professional development activities? In what ways?

Section 3: Impact on your wider role/work

1. To what extent did The ONE impact upon your other teaching activities, (e.g., activities designed to target language development and early literacy)?

a. Did The ONE training crowd out or reduce other teaching activities?

b. Did The ONE complement, or work alongside, other teaching activities? In what ways?

2. Has participating in The ONE increased overall the number of times per week you do play-based numeracy activities with children in your setting?

a. If not, why?

3. Have you had to make changes to the number of times per week that you do other activities to accommodate The ONE activities?

a. If yes, what activities have you had to cut down on?

4. How has delivering The ONE activities impacted on you as a practitioner?

a. Has it helped you to strengthen your practice in any ways?

b. Has it put any additional pressure on you?

c. Did it add to your workload? Or the workload of your colleagues?

5. To what extent has delivering The ONE impacted upon your intention to remain working in your setting and/or remain working in the profession?

a. Were you planning to leave your setting previously? And now? Why?

b. Were you planning to stay in your setting previously? And now? Why?

6. To what extent (if at all) has participating in The ONE affected your ability to participate in other professional development?

a. Did The ONE training crowd out your other professional development activities?

b. Did The ONE complement, or work alongside, your other professional development activities? In what ways?

7. Is there anything else that you'd like to tell us about your setting's experience of The ONE?

Finally, a reminder that you have been sent 1) links to a survey an 2) a request to complete your setting's EYPP and attendance data. These are the last things needed before your setting receives its payment for being involved in the intervention. Would you like to receive an reminder email about either/both of these?

# Appendix H: Interview Schedule (Managers)

Section 1: Delivery of the intervention

1. Before I start can I confirm that you are the manager/Head at X setting?

a. (If yes) In terms of your role, how involved were you in The ONE? Did you attend the trainings yourself, and did you deliver any of the activities? [Some of these questions could be skipped, depending on their level of involvement]

2. How have you found The ONE overall?

a. How many staff were involved in your setting?

b. How many children/how many rooms?

c. Are there any aspects that were particularly challenging?

d. Is there anything you have found particularly helpful?

3. Were there any surprising benefits or consequences of delivering The ONE intervention in your setting?

4. Were there any other factors that you felt made it easier to implement The ONE in your setting?

5. Were there any other challenges you came across in implementing The ONE in your setting?

6. In your view, what further support could have improved the delivery of The ONE in your setting?

7. How do you feel The ONE has fitted into your setting's day-to-day routine?

a. For example as small group activities, one-to-one, in outdoor play or free play?

8. Do you feel that the activities are pitched at the right level for the children involved?

a. Why/why not?

9. Are there any groups of children that you feel particularly engage with the activities?

a. Or, on the other hand, that struggle to engage?

b. For example children with English as an additional language, or those from disadvantaged backgrounds?

10. Are you aware of any changes or adaptations that were necessary so that the material provided or instructions provided to better meet a child's needs at any point?

Section 2: Training and support materials

1. Is there anything about The ONE itself that you have felt confused or unsure about, or that your team have expressed uncertainty about?

2. Is there any additional training, support or resources you would have liked to have received to help your team deliver the activities confidently?

3. To what extent (if at all) do you feel the training sessions improved your team's knowledge of executive functions and their relationship with early numeracy?

4. To what extent (if at all) do you feel the training sessions improved your team's ability to incorporate play-based maths activities into your setting's daily routine?

5. To what extent (if at all) do you feel the training sessions improved your team's ability to adapt activities to the right level for different children?

Section 3: Impact on your wider role/work

1. Has participating in The ONE increased the overall number of times per week staff in your setting do numeracy activities with children in your setting?

2. Have you had to make changes to the number of times per week that staff in your setting do other activities to accommodate The ONE activities?

a. If yes, what activities have you had to cut down on?

3. How has delivering the ONE activities impacted on the everyday practice of members of your staff?

a. Has it strengthened their practice in any ways?

b. Has it put any additional pressure on members of your team?

4. To what extent have The ONE training and activities impacted upon the workload of your team?

5. Did your setting's involvement in The ONE impact upon other professional development for your staff?

a. Did The ONE training crowd out other professional development activities?

b. Did The ONE complement, or work alongside, other professional development activities? In what ways?

6. To what extent has The ONE resulted in any changes to staff retention?

a. a. Do you think that The ONE has encouraged staff to stay, who may otherwise have left?

b. b. Do you think that being involved in The ONE may have encouraged staff to leave, who may otherwise have stayed?

7. Is there anything else that you'd like to tell us about your setting's experience of The ONE?

Finally, a reminder that you have been sent 1) links to a survey an 2) a request to complete your setting's EYPP and attendance data. These are the last things needed before your setting receives its payment for being involved in the intervention. Would you like to receive an reminder email about either/both of these?

# Appendix I: Baseline Survey

A. Background information

1. Please type the name of your nursery or pre-school in the space below. [open response]*

2. Please enter your nursery or pre-school's postal code. [open response]

3. This survey is intended for the Early Years Practitioners (staff at the nursery/pre-school who work directly with children on a regular basis) who will be nominated for training under The ONE.

Please select one answer.

a. I am an Early Years Practitioner who works with children aged 3 – 4 years old. .

b. I am an Early Years Practitioner who does not work with children aged 3 – 4 years old (i.e., I only work with younger children).

c. I am NOT an Early Years practitioner in my nursery or pre-school

Note: If the respondent answered 'B' or 'C' to Q4, SURVEY ENDS. Show the following message: 'Please forward this survey to the Early Years Practitioner who will be nominated for training in The ONE programme' This should be a staff member (not setting manager) who works with children aged 3-4 years old.

4. How many years of experience do you have working with children in early childhood education and care? Please input a number, no more than two digits.

5. What are your current working hours at this nursery or pre-school?

Please select one answer.

a. Full-time working hours (working 35 hours or more per week)

b. Part-time working hours, working 17.5 – 34 hours per week

c. Part-time working hours, working fewer than 17.5 hours per week

6. Which of the following qualifications have you received, which enable you to work in early childhood education and care?

Please select all that apply.

a. Level 2 qualification (e.g., Level 2 diploma for Early Years Education and Care)

b. Level 3 qualification (e.g., Level 3 diploma in Early Years Education and Care)

c. Certificate of Higher Education (e.g. a certificate in Early Years Education and Care from a university)

d. Undergraduate diploma (e.g., a diploma in the Early Years Education and Care from a university)

e. Undergraduate degree (e.g., BA in Early Childhood, BEd in Primary or Early Years, BSc in Educational and Developmental Psychology)

f. Postgraduate Certificate in Education (PGCE)

g. Postgraduate Diploma in Education (PGDE)

h. Postgraduate Teaching Apprenticeship

i. Postgraduate Degree (e.g., MA in early childhood education, PhD in Education)

j. Alternative early years or teacher training (e.g., Teach First, Now Teach)

k. Other type of training (please specify: _____)

7. In your initial early years qualifications (those you selected in the previous question), which of the following topics were covered?

Please select all that apply.

a. Child development (e.g., socio-emotional development, fine- and gross-motor development, cognitive development)

b. Executive function (e.g., working memory, inhibitory control, cognitive flexibility) and its development in childhood

c. Early language and literacy (e.g., speech and oral language development, phonemic awareness, phonics)

d. Early years' maths (e.g., early numeracy, pattern recognition)

e. Early years' science

f. None of the above

8. Have you participated in professional development/training over the last two years?

Please select one answer.

a. Yes

b. No

IF Q8= 'Yes', ASK Q9 and Q10; IF Q8 = 'No', SKIP TO Q11

9. Approximately how many hours of professional development/training have you participated in over the last two years?

Please input a number only.

10. In your professional development/training over the last two years, which of the following topics were covered?

Please select all that apply.

a. Child development (e.g., socio-emotional development, fine- and gross-motor development, cognitive development)

b. Executive function (e.g., working memory, impulse control, cognitive flexibility) and its development in childhood

c. Early language and literacy (e.g., speech and oral language development, phonemic awareness, phonics)

d. Early years' maths (e.g., early numeracy, pattern recognition)

e. Early years' science

f. None of the above

B. Practices with Children in Early Maths, Executive Function and Early Childhood Development

Section B delves into the structure of your lessons at the setting and the types of activities you tend to do with the children.

11. Over a typical working week, how often do you do the following with the children in your classroom or playgroup:

| | Never | Once a week | 2-3 times a week | 4-5 times a week | More than once a day |
|---|---|---|---|---|---|
| I play games involving counting sets of objects (e.g., counting cups at meal times) | | | | | |
| I play games involving recognising numbers (e.g., matching numbers to dots) | | | | | |
| I play games involving sorting objects into groups (e.g., sorting leaves by size) | | | | | |
| I play games involving making or extending patterns (e.g., making towers with blocks in a red-blue-red-blue pattern) | | | | | |
| I play games involving recognising, naming or matching shapes (e.g., making or talking about shapes during craft activities) | | | | | |
| I encourage children to use prepositions to talk about space (e.g., asking children to describe locations using words like under/on/between instead of pointing) | | | | | |
| I ask children to complete activities that require multiple steps or paying attention to lengthy explanations | | | | | |
| I ask children to wait or take turns during everyday activities or games. | | | | | |
| I invite children to come up with new ways to solve a problem in an activity or game. | | | | | |
| I read to the children | | | | | |
| I play games involving letters with the children (e.g., tracing out letter | | | | | |

| | | | | | |
|---|---|---|---|---|---|
| shapes in the air, finding letters in a word) | | | | | |
| I play games involving different letter sounds with children (e.g., picking out certain sounds in a word) | | | | | |
| I play games focussed on understanding the world around us (e.g., activities about life cycles or the environment, exploring the weather) | | | | | |

12. In a typical day, how much time do you spend in activities or play with children in the following group sizes:

| | Never | Occasionally, but not every day | Less than 30 minutes a day | Between 30 minutes – 1 hour a day | Between 1 – 2 hours a day | More than 2 hours a day |
|---|---|---|---|---|---|---|
| Whole-group play or activities (typically involves all children in the classroom or playgroup) | | | | | | |
| Small-group play or activities (typically involves a small subset of the normal classroom or playgroup) | | | | | | |
| One-to-one play or activities (typically involves working with just one child) | | | | | | |

C. Early Maths and Executive Function instruction, programmes and support available in your setting

Section C asks for your opinion on a variety of statements about the role of maths in learning, your confidence in teaching maths, the manner and frequency with which maths is taught in your setting, and the importance of certain skills in early maths learning. There will also be questions surrounding your experience (if any) with early maths and/or executive function programmes or interventions.

13. For each of the following statements, rate your agreement by checking the appropriate box.

| | Strongly Disagree | Disagree | Neutral (neither agree nor disagree) | Agree | Strongly Agree |
|---|---|---|---|---|---|
| Maths is an important part of the early year's curriculum. | | | | | |
| Maths activities are a very important part of the early year's education. | | | | | |
| Development in academics such as maths is the goal of early year's education. | | | | | |
| Early year's maths activities help young children to be interested in maths when they go to school. | | | | | |

| | | | | | |
|---|---|---|---|---|---|
| Maths activities help improve young children's language skills. | | | | | |

14. For each of the following statements, rate your agreement by checking the appropriate box.

| | Strongly Disagree | Disagree | Neutral (neither agree nor disagree) | Agree | Strongly Agree |
|---|---|---|---|---|---|
| I feel comfortable planning maths activities for young children in my setting. | | | | | |
| I feel comfortable doing maths activities in my setting. | | | | | |
| I have enough maths knowledge to explore maths with young children. | | | | | |
| I feel comfortable using tools such as scales, measuring cups and rulers with young children in my setting. | | | | | |
| I enjoy doing maths activities with young children in my setting. | | | | | |
| I am afraid that young children might ask me a question about maths that I can't answer. | | | | | |
| I feel comfortable adapting maths activities to extend children's maths knowledge based on my observations of young children. | | | | | |

15. For each of the following statements, rate your agreement by checking the appropriate box.

| | Strongly Disagree | Disagree | Neutral (neither agree nor disagree) | Agree | Strongly Agree |
|---|---|---|---|---|---|
| It is important for my setting to have a maths area that can be freely explored by children. | | | | | |
| There is not enough time in the day to explore maths with young children. | | | | | |
| I use all kinds of materials for maths activities with young people. | | | | | |
| Preparation for maths education takes more time than other areas. | | | | | |
| I get ideas for maths activities from what my young children do, say and ask. | | | | | |
| I include some books about maths during story time. | | | | | |

16. How important are the following skills for early maths learning?

| | Not important | Somewhat important | Moderately important | Very important |
|---|---|---|---|---|
| Children know number facts. | | | | |
| Children understand mathematical learning. | | | | |
| Children understand how mathematics is used in the real world. | | | | |
| Children can manipulate abstract information. | | | | |
| Children can store and manipulate information in their head. | | | | |
| Children can focus on relevant information and ignore distractions. | | | | |
| Children are able to think flexibly. | | | | |
| Children have good verbal skills. | | | | |
| Children are able to provide reasons to support their solutions. | | | | |
| Children are able to think creatively. | | | | |
| Children have good spatial skills. | | | | |

17. To what extent do you disagree or agree: It would be good for all early years practitioners to have professional development/training in early years' maths.

Please select one answer.

a. Strongly disagree

b. Disagree

c. Neither agree nor disagree

d. Agree

e. Strongly agree

18. In the past year, has your setting implemented any early maths and/or executive function programmes or interventions?

Please select one answer.

a. Yes

b. No

c. I don't know

IF Q18 = 'Yes', ASK Q19; IF Q18 = 'No' or 'I don't know', SKIP TO Q26.

19. Please name the early maths and/or executive function programme or intervention that was most recently implemented in your setting and tell us when it was implemented. [open response]

20. How was the early maths and/or executive function programme or intervention you previously named delivered to children?

Please select one answer.

a. Whole class

b. Small-group instruction

c. One-to-one instruction

d. A combination of the above (please specify in the comment box below)

e. Other (please specify)

A combination of the above (please specify):

[Short open response]

21. What was the duration of the early maths and/or executive function programme or intervention you named previously?

Please select one answer.

a. Less than 4 weeks

b. 4-8 weeks

c. 9-12 weeks

d. 13-16 weeks

e. 17-20 weeks

f. 21-24 weeks

g. More than 24 weeks

22. How many minutes each week were involved in delivering the early maths and/or executive function programme or intervention you named previously?

Please input a number. [open response]

23. Please describe in more detail the early maths and/or executive function programme or intervention you named previously. For example, you may provide more detail regarding the programme's objectives, materials required, and content coverage, amongst others. [open response]

24. Did you undergo training to be able to deliver the early maths and/or executive function programme or intervention you named previously?

Please select one answer.

a. Yes

b. No

IF Q24 = 'Yes', ASK Q25; IF Q24 = 'No', SKIP TO Q26.

25. How many days of training in total did each practitioner undergo to be able to deliver the early maths and/or executive function programme or intervention you named previously?

Please select one answer.

a. Less than 1 day

b. 1-2 days

c. 3-4 days

d. More than 4 days

D. Engagement with The ONE to date and participation in interviews

Section D of the survey will ask first about your setting's role in The ONE Project, and second, your consent for an interview to inform the evaluation of The ONE.

26. Has your school been allocated to the treatment or control group?

Please select one answer.

a. Treatment

b. Control

c. Do not know yet

27. The evaluation of The ONE is intended to gather lessons learned and ensure other nurseries and preschools benefit from the insights of settings that have participated in The ONE trial. To further inform our evaluation of The ONE, we would like to invite practitioners in select settings to participate in an online interview in the coming months. We appreciate the pressure settings are under at this time, but if you would consider participating in an interview, we would be very grateful.

Please note that if you give permission to be contacted for an interview below, you can withdraw this permission at any time by emailing theone@randeurope.org.

Do you provide your consent to be contacted for the purpose of taking part in an interview to inform the evaluation of The ONE?

Please select one answer.

a. Yes

b. No

# Appendix J: Control Setting Practitioners Endline Survey

Introduction

This is an endline survey for practitioners working in nurseries or preschools participating in the evaluation of The ONE Programme. This survey is designed for those settings that did not receive The ONE intervention between January and May 2024 (control settings). If your setting did participate in the intervention, please do not fill out this survey and email theone@randeurope.org to let us know you have received the wrong link.

This survey is part of an independent evaluation of The ONE intervention by RAND Europe. This evaluation is funded by the Education Endowment Foundation (EEF), an independent charity who fund research into 'what works' to improve educational practice.

This endline survey will ask you for information about you and your setting, as well as business as usual in your setting. These endline surveys will enhance our understanding of different setting contexts, and how this may influence the implementation and effectiveness of The ONE Programme.

Survey responses will be analysed collectively.. No one setting or person will ever be identified in any analysis or report. All responses will be kept confidential. The data collected will be treated confidentially and shall not be shared with third parties. The data is processed and stored securely in accordance with the General Data Protection Regulation (GDPR). No personal data or IP addresses will be recorded or added to the dataset. For more information about how the data is processed and stored for The ONE Project, please refer to our Information Sheet for Early Years Practitioners and the Information Sheet for Early Years Settings. For any questions about the research and your rights as a participant, you may contact the data protection officer of RAND Europe. If you have any questions on this survey or your rights as a research participant, please do not hesitate to contact the evaluation team at RAND Europe on theone@randeurope.org.

We ask that you complete this endline survey within two weeks of receiving it.

By clicking 'Next Page' you consent to the evaluation team using your responses in their evaluation of The ONE Programme.

If you are happy to continue, please click 'Next Page'.

Privacy Notice

RAND Europe is an independent not-for-profit research institute based in Cambridge and Brussels whose mission is to help improve policy and decision making through evidence-based research. RAND Europe are data controllers for this project, and the Department for Education (DfE) and the Education Endowment Fund (EEF) are joint data controllers for the Stronger Practice Hubs to which the project belongs.

Below we set out how your personal information will be collected, used and looked after in accordance with the UK General Data Protection Regulations (GDPR) and Data Protection Act 2018.

What data are we collecting?

As the practitioner in your setting, we are interested in asking you about your setting and about business as usual in your setting. This survey is being circulated to all control settings.

All of your responses to the survey will be confidential. The SmartSurvey platform automatically links your responses to your name and email address, but only the research team will have access to this data. The research team will only ever analyse and report the data at an anonymised and aggregate level and no individual will ever be named in any report

What are we using the data for?

The evaluation team is collecting information on your setting to aid the evaluation of The ONE. The evaluation aims to find out more about the impact that The ONE intervention has on pupil outcomes and how settings are implementing

The ONE. This questionnaire is being conducted at the end of the trial to understand your experience of business as usual, as a control setting.

We will look at how the responses of settings assigned to take part in The ONE compare to settings that are not implementing The ONE. We will analyse the data to see if there is a difference. This will help us to understand how and if The ONE contributed to changes over time, how this compares across settings implementing The ONE and those that are not, and any contextual factors that might help us understand the programme better.

How will we collect your data?

Oxford shared your contact details with us using an encrypted file shared via a secure file sharing platform called OneDrive.

Once the questionnaire is underway, your questionnaire responses will be collected and stored on the SmartSurvey platform by RAND Europe. RAND Europe will obtain the data securely from SmartSurvey. SmartSurvey will delete your questionnaire responses and identifiable data once RAND Europe has obtained it. RAND Europe will maintain this data in confidence and use it only for the purpose of evaluating The ONE. RAND Europe will save this data in a password protected folder on their internal network. Only nominated researchers on the project with the password will be able to access the file.

Please do not provide any sensitive data in this questionnaire, such as your political persuasion. If sensitive data is provided in the questionnaire, RAND Europe will delete it before analysis.

How do we keep your data secure?

The evaluation team have put various security measures in place to keep personal data secure and to prevent any unauthorised access to or use of it in accordance with Data Protection Act (2018) and UK GDPR requirements. All data collected by RAND Europe will be stored on secure servers, accessed only by relevant project team. No data will be saved on servers or shared with processors outside the UK.

How long do we keep your data?

The data will be stored securely on RAND Europe's data servers for the duration of The ONE evaluation project – from January 2023 to June 2025. To allow us time to analyse and report the results of the trial, this period will extend beyond your setting's participation in the programme. Your responses will be used to create descriptive statistics and individual settings will not be identified in this context. Your responses shall not be passed on to any third party.

What is the legal basis for processing your data?

The legal basis for RAND Europe to process your personal data is legitimate interests detailed in Article 6(1)(f) of the UK GDPR. To ensure that all processing is fair and lawful, RAND Europe have also completed a Legitimate Interest Assessment and a Data Protection Impact Assessment and have received ethical approval from the RAND internal review board. RAND Europe will process only what is required to meet these legal bases and will ensure security and safeguards are in place to protect the information.

What are your rights?

RAND Europe operates in accordance with the Data Protection Act 2018 and UK GDPR 2016 requirements. You are provided with certain rights that you may have the right to exercise through us. In summary those rights are:

- To access your data ("data subject access request") (Article 15 of the GDPR)
- To have inaccurate personal data rectified (Article 16 of the GDPR)
- To have your data erased (Article 17 of the GDPR)
- To restrict the processing of your data (Article 18 of the GDPR)
- Request the transfer of your personal data to you or to a third party (Article 20 of the GDPR)

- Object to processing of your personal data (Article 21 of the GDPR).

How do you contact us?

If you have any questions about this questionnaire or wish to exercise any of these rights, please contact the RAND Europe study team at theone@randeurope.org. Alternatively, you may contact the Data Protection Officer by email at REdpo@randeurope.org, or in writing to Data Protection Officer, RAND Europe, Eastbrook, Shaftesbury Road, Cambridge, CB2 8BF, UK. If contacting the Data Protection Officer, please make reference to the project The ONE and project number 022807.014 in your information request.

For independent advice about data protection or to lodge a complaint about how we have handled your personal data, you can contact the Information Commissioner's Office. You can visit www.ico.org.uk, email casework@ico.org.uk, or write to Information Commissioner's Office, Wycliffe House, Water Lane, Wilmslow, Cheshire, SK9 5AF, UK

Please click 'Next Page' if you would like to proceed. In doing so, you will be confirming that you have read the above Privacy Notice, accept its terms and consent to taking part in this questionnaire.

A. Background information

1. Please type the name of your nursery or pre-school in the space below. [open response]*

2. Please enter your nursery or pre-school's postal code. [open response]

3. This survey is intended for practitioners at settings that received The ONE intervention between January and May 2024.

Please select one answer.

a. I am an Early Years practitioner who works in a nursery or pre-school that is waitlisted for The ONE intervention.

b. I am an Early Years practitioner who works in a nursery or pre-school that did receive The ONE intervention between January and May 2024.

c. I am NOT an Early Years practitioner in a nursery or pre-school.

Note: If the respondent answered 'B' or 'C' to Q3, SURVEY ENDS.

If answered 'B' show the following message: "You have indicated that your nursery or pre-school did receive The ONE intervention between January and May 2024. If this is the case, please email theone@randeurope.org to let us know you have received the incorrect link."

If answered 'C' show the following message: "Please forward the email with the link to this survey to an Early Years practitioner in your nursery or pre-school and ask them to complete it."

4. Did you receive and complete a similar survey from us during November/December 2023?

a) Yes

b) No

c) I can't remember

5. How many years of experience do you have working with children in early childhood education and care? Please input a number, no more than two digits. If you have less than one year's experience please input 0.

6. What are your current working hours at this nursery or pre-school?

Please select one answer.

a. Full-time working hours (working 35 hours or more per week)

b. Part-time working hours (working 17.5 – 34 hours per week)

c. Part-time working hours (working fewer than 17.5 hours per week)

7. Which of the following qualifications have you received?

Please select all that apply.

a. Level 2 qualification (e.g., Level 2 diploma for Early Years Education and Care)

b. Level 3 qualification (e.g., Level 3 diploma in Early Years Education and Care)

c. Certificate of Higher Education (e.g. a certificate in Early Years Education and Care from a university)

d. Undergraduate diploma (e.g., a diploma in the Early Years Education and Care from a university)

e. Undergraduate degree (e.g., BA in Early Childhood, BEd in Primary or Early Years, BSc in Educational and Developmental Psychology)

f. Postgraduate Certificate in Education (PGCE)

g. Postgraduate Diploma in Education (PGDE)

h. Postgraduate Teaching Apprenticeship

i. Postgraduate Degree (e.g., MA in early childhood education, PhD in Education)

j. Alternative early years or teacher training (e.g., Teach First, Now Teach)

k. Other type of training (please specify: _____)

8. In your initial early years qualifications (those you selected in the previous question), which of the following topics were covered?

Please select all that apply.

a. Child development (e.g., socio-emotional development, fine- and gross-motor development, cognitive development)

b. Executive function (e.g., working memory, inhibitory control, cognitive flexibility) and its development in childhood

c. Early language and literacy (e.g., speech and oral language development, phonemic awareness, phonics)

d. Early years' maths (e.g., early numeracy, pattern recognition)

e. Early years' science

f. None of the above

B. Details about Your Setting

9. Consider the classroom or playroom in which you teach. On a typical day, how many children and staff in this classroom meet the following criteria:

(Please enter a number. If there are no children or staff that meet these criteria, please put 0)

| | |
|---|---|
| Children in the classroom/playroom | |
| Children in the classroom/playroom who are eligible for Early Years Pupil Premium (EYPP) | |
| Children in the classroom/playroom who have English as an Additional Language (EAL) | |
| Children in the classroom/playroom with known or suspected Special Educational Needs and Disabilities (SEND). These include Speech, language and communication needs (SLCN); Personal, Social and Emotional Development needs (PSED); and social, emotional and mental health needs (SEMH). | |
| Staff in the classroom/playroom | |
| Staff in the classroom/playroom with a Level 3 Certificate of higher | |
| Staff in the classroom/playroom with a Bachelor's Degree or higher | |

10. Over **a typical working week**, how often do you do the following with the children in your classroom or playgroup:

| | Never | Once a week | 2-3 times a week | 4-5 times a week | More than once a day |
|---|---|---|---|---|---|
| I play games involving counting sets of objects (e.g., counting cups at meal times) | | | | | |
| I play games involving recognising numbers (e.g., matching numbers to dots) | | | | | |
| I play games involving sorting objects into groups (e.g., sorting leaves by size) | | | | | |
| I play games involving making or extending patterns (e.g., making towers with blocks in a red-blue-red-blue pattern) | | | | | |
| I play games involving recognising, naming or matching shapes (e.g., making or talking about shapes during craft activities) | | | | | |

| | | | | | |
|---|---|---|---|---|---|
| I encourage children to use prepositions to talk about space (e.g., asking children to describe locations using words like under/on/between instead of pointing) | | | | | |
| I ask children to complete activities that require multiple steps or paying attention to lengthy explanations | | | | | |
| I ask children to wait or take turns during everyday activities or games. | | | | | |
| I invite children to come up with new ways to solve a problem in an activity or game. | | | | | |
| I read to the children | | | | | |
| I play games involving letters with the children (e.g., tracing out letter shapes in the air, finding letters in a word) | | | | | |
| I play games involving different letter sounds with children (e.g., picking out certain sounds in a word) | | | | | |
| I play games focussed on understanding the world around us (e.g., activities about life cycles or the environment, exploring the weather) | | | | | |

11. In **a typical day,** how much time do you spend in activities or play with children in the following group sizes:

| | Never | Occasionally, but not every day | Less than 30 minutes a day | Between 30 minutes – 1 hour a day | Between 1 – 2 hours a day | More than 2 hours a day |
|---|---|---|---|---|---|---|
| Whole-group play or activities (typically involves all children in the classroom or playgroup) | | | | | | |
| Small-group play or activities (typically involves a small subset of the normal classroom or playgroup) | | | | | | |
| One-to-one play or activities (typically involves working with just one child) | | | | | | |

C. Early Maths and Executive Function instruction, programmes and support available in your setting

Section C asks for your opinion on a variety of statements about the role of maths in learning, your confidence in teaching maths, the manner and frequency with which maths is taught in your setting, and the importance of

certain skills in early maths learning. There will also be questions surrounding your experience (if any) with early maths and/or executive function programmes or interventions.

12. For each of the following statements, rate your agreement by checking the appropriate box.

[Skip question]

| | Strongly Disagree | Disagree | Neutral (neither agree nor disagree) | Agree | Strongly Agree |
|---|---|---|---|---|---|
| Maths is an important part of the early years curriculum. | | | | | |
| Maths activities are a very important part of the early years education. | | | | | |
| Development in academics such as maths is the goal of early years education. | | | | | |
| Early years maths activities help young children to be interested in maths when they go to school. | | | | | |
| Maths activities help improve young children's language skills. | | | | | |

13. For each of the following statements, rate your agreement by checking the appropriate box.

| | Strongly Disagree | Disagree | Neutral (neither agree nor disagree) | Agree | Strongly Agree |
|---|---|---|---|---|---|
| I feel comfortable planning maths activities for young children in my setting. | | | | | |
| I feel comfortable doing maths activities in my setting. | | | | | |
| I have enough maths knowledge to explore maths with young children. | | | | | |
| I feel comfortable using tools such as scales, measuring cups and rulers with young children in my setting. | | | | | |
| I enjoy doing maths activities with young children in my setting. | | | | | |

| | | | | | |
|---|---|---|---|---|---|
| I am afraid that young children might ask me a question about maths that I can't answer. | | | | | |
| I feel comfortable adapting maths activities to extend children's maths knowledge based on my observations of young children. | | | | | |

14. For each of the following statements, rate your agreement by checking the appropriate box.

| | Strongly Disagree | Disagree | Neutral (neither agree nor disagree) | Agree | Strongly Agree |
|---|---|---|---|---|---|
| It is important for my setting to have a maths area that can be freely explored by children. | | | | | |
| There is not enough time in the day to explore maths with young children. | | | | | |
| I use all kinds of materials for maths activities with young people. | | | | | |
| Preparation for maths education takes more time than other areas. | | | | | |
| I get ideas for maths activities from what my young children do, say and ask. | | | | | |
| I include some books about maths during story time. | | | | | |

15. How important are the following skills for early maths learning?

| | Not important | Somewhat important | Moderately important | Very important |
|---|---|---|---|---|
| Children know number facts. | | | | |
| Children understand mathematical learning. | | | | |
| Children understand how mathematics is used in the real world. | | | | |
| Children can manipulate abstract information. | | | | |
| Children can store and manipulate information in their head. | | | | |
| Children can focus on relevant information and ignore distractions. | | | | |
| Children are able to think flexibly. | | | | |
| Children have good verbal skills. | | | | |

| | | | | |
|---|---|---|---|---|
| Children are able to provide reasons to support their solutions. | | | | |
| Children are able to think creatively. | | | | |
| Children have good spatial skills. | | | | |

We're interested in knowing more about the types of maths and executive functioning activities you do with the children in your settings.

16. How often do children in your setting engage in maths activities?

Please select one answer.

a. Several times a day

b. At least once a day

c. Several times a week

d. At least once a week

e. Never

17. What types of maths activities do you run for the children in your setting?

Please select all that apply.

a. 1:1 activities

b. Small group activities

c. Whole room activities

18. What do you use to help you plan and deliver the maths activities?

[open text response]

19. How often do children in your setting engage in executive functioning activities (e.g., memory games)?

Please select one answer.

a. Several times a day

b. At least once a day

c. Several times a week

d. At least once a week

e. Never

f. I am not familiar with the term 'executive functioning skills'

20. What types of executive functioning activities do you run for the children in your setting?

Please select all that apply.

a. 1:1 activities

b. Small group activities

c. Whole room activities

d. I am not familiar with the term 'executive functioning skills'

21. What do you use to help you plan and deliver the executive functioning activities?

[open text response] If you are not familiar with the term 'executive functioning' please leave blank

22. How important do you think the following are in supporting children's early learning and development:

| | Not important | Somewhat important | Moderately important | Very important | I don't know |
|---|---|---|---|---|---|
| Early language and literacy | | | | | |
| Early maths | | | | | |
| Executive function (e.g., working memory, impulse control, cognitive flexibility) | | | | | |
| Early science | | | | | |

23. Please indicate the impact that the following **potential facilitators** had on the delivery of maths activities in your setting, in the last year:

| | Very negative impact (this stops us from delivering maths activities almost entirely) | Slightly negative impact (stops us from doing some maths activities) | No impact | Slightly positive impact (this helps us to deliver some maths activities successfully) | Very positive impact (this was a key factor that helps us to deliver maths activities successfully) | N/A or not sure |
|---|---|---|---|---|---|---|

| | | | | | | |
|---|---|---|---|---|---|---|
| Staff retention/consistency during the last year | | | | | | |
| Support from colleagues | | | | | | |
| Engagement of staff | | | | | | |
| Professional development | | | | | | |
| Availability of protected time for staff professional development | | | | | | |
| Availability of protected time for preparation and planning of classroom/playroom activities | | | | | | |
| Availability of extra staff to cover professional development, preparation and planning time | | | | | | |
| Ease of adaptability of the setting routines and structure | | | | | | |
| Protected time in the setting routine or structure for early numeracy activities | | | | | | |

24. Please indicate the impact that the following **potential barriers** had on the delivery of maths activities in your setting, in the last year:

| | Very negative impact (this stops us from delivering maths activities almost entirely) | Slightly negative impact (stops us from doing some maths activities) | No impact | Slightly positive impact (this helps us to deliver some maths activities successfully) | Very positive impact (this was a key factor that helps us to deliver maths activities successfully) | N/A or not sure |
|---|---|---|---|---|---|---|
| Staff absences | | | | | | |
| Child absences | | | | | | |
| Staff turnover | | | | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| Low levels of engagement from staff | | | | | | |
| Difficulty in scheduling professional development sessions around other commitments | | | | | | |
| Difficulty in accessing the necessary materials | | | | | | |
| Difficulty in preparing and planning | | | | | | |
| Difficulty in maths activities to the routines and structure of the day | | | | | | |
| Competing priorities in the setting making it difficult to deliver activities | | | | | | |
| Difficulty in maintaining child engagement in activities | | | | | | |
| Difficulty in understanding materials and activities needed to deliver maths activities | | | | | | |
| Inappropriate delivery space (e.g., the room could not easily accommodate maths activities, the room was too loud or noisy) | | | | | | |
| Technological difficulties or lack of resources (e.g., computers could not be made available for professional development calls, tech resources limited) | | | | | | |

26. Did you undergo training to be able to deliver the early maths and/or executive function activities that you do?

Please select one answer.

a. Yes

b. No

IF Q31 = 'Yes', ASK Q32-34; IF 31 = 'No', SKIP TO END OF SURVEY.

27. When did this training occur?

Please select all that apply.

a. In the last three months

b. In the last six months

c. In the last year

d. Over a year ago

28. How many days of training in total did you undergo to be able to deliver these early maths and/or executive function activities?

Please select one answer.

a. Less than 1 day

b. 1-2 days

c. 3-4 days

d. More than 4 days

29. What topics were covered by this training?

[open response]

Thank you very much! The survey ends here. If you have any questions about this survey or your rights as a research participant, please do not hesitate to contact the evaluation team at RAND Europe on theone@randeurope.org

# Appendix K: Treatment Setting Practitioners Endline Survey

Introduction

This survey is designed for those settings that received The ONE intervention (4 weeks professional development and 12 weeks of intervention activities delivered within the classroom/playroom) between January and May 2024 (treatment settings). If your setting did not participate in the intervention, please do not fill out this survey and email theone@randeurope.org to let us know you have received the wrong link.

This survey is part of an independent evaluation of The ONE intervention by RAND Europe. This evaluation is funded by the Education Endowment Foundation (EEF), an independent charity who fund research into 'what works' to improve educational practice. This is an endline survey for practitioners working in nurseries or preschools who delivered The ONE Programme. If you did not deliver The ONE Programme, please pass this on to one of your colleagues who did.

This endline survey will ask you for information about you and your setting, as well as your experiences of The ONE intervention and what affects this may have had on the wider setting. These endline surveys will enhance our understanding of different setting contexts, and how this may influence the implementation and effectiveness of The ONE Programme, and how much The ONE costs for settings to deliver, to help inform further evaluation and delivery of this programme.

Survey responses will be analysed collectively. No one setting or person will ever be identified in any analysis or report. All responses will be kept confidential. The data collected will be treated confidentially and shall not be shared with third parties. The data is processed and stored securely in accordance with the General Data Protection Regulation (GDPR). No personal data or IP addresses will be recorded or added to the dataset. For more information about how the data is processed and stored for The ONE Project, please refer to our Information Sheet for Early Years Practitioners and the Information Sheet for Early Years Settings. For any questions about the research and your rights as a participant, you may contact the data protection officer of RAND Europe. If you have any questions on this survey or your rights as a research participant, please do not hesitate to contact the evaluation team at RAND Europe on theone@randeurope.org.

We ask that you complete this endline survey within two weeks of receiving it.

By clicking 'Next Page' you consent to the evaluation team using your responses in their evaluation of The ONE Programme.

If you are happy to continue, please click 'Next Page'.

Privacy Notice

RAND Europe is an independent not-for-profit research institute based in Cambridge and Brussels whose mission is to help improve policy and decision making through evidence-based research. RAND Europe are data controllers for this project, and the Department for Education (DfE) and the Education Endowment Fund (EEF) are joint data controllers for the Stronger Practice Hubs to which the project belongs.

Below we set out how your personal information will be collected, used and looked after in accordance with the UK General Data Protection Regulations (GDPR) and Data Protection Act 2018.

What data are we collecting?

As the practitioner in your setting, we are interested in asking you about factors that may have impacted on delivery of The ONE your setting, as well as any unintended consequences that may have resulted from it. This survey is being circulated to all treatment settings involved in The ONE.

All of your responses to the survey will be confidential. The SmartSurvey platform automatically links your responses to your name and email address, but only the research team will have access to this data. The research team will only ever analyse and report the data at an anonymised and aggregate level and no individual will ever be named in any report.

What are we using the data for?

The evaluation team is collecting information on your setting to aid the evaluation of The ONE. The evaluation aims to find out more about the impact that The ONE intervention has on pupil outcomes and how settings are implementing The ONE. This questionnaire is being conducted at the end of the trial to understand your experience of implementing the intervention in your setting. We will ask you about factors that may have influenced the implementation of The ONE in your setting, any unintended consequences that may have resulted from it.

We will look at how the responses of settings assigned to take part in The ONE compare to settings that are not implementing The ONE. We will analyse the data to see if there is a difference. This will help us to understand how and if The ONE contributed to changes over time, how this compares across settings implementing The ONE and those that are not, and any contextual factors that might help us understand the programme better.

How will we collect your data?

Oxford shared your contact details with us using an encrypted file shared via a secure file sharing platform called OneDrive.

Once the questionnaire is underway, your questionnaire responses will be collected and stored on the SmartSurvey platform by RAND Europe. RAND Europe will obtain the data securely from SmartSurvey. SmartSurvey will delete your questionnaire responses and identifiable data once RAND Europe has obtained it. RAND Europe will maintain this data in confidence and use it only for the purpose of evaluating The ONE. RAND Europe will save this data in a password protected folder on their internal network. Only nominated researchers on the project with the password will be able to access the file.

Please do not provide any sensitive data in this questionnaire, such as your political persuasion. If sensitive data is provided in the questionnaire, RAND Europe will delete it before analysis.

How do we keep your data secure?

The evaluation team have put various security measures in place to keep personal data secure and to prevent any unauthorised access to or use of it in accordance with Data Protection Act (2018) and UK GDPR requirements. All data collected by RAND Europe will be stored on secure servers, accessed only by relevant project team. No data will be saved on servers or shared with processors outside the UK.

How long do we keep your data?

The data will be stored securely on RAND Europe's data servers for the duration of The ONE evaluation project – from January 2023 to June 2025. To allow us time to analyse and report the results of the trial, this period will extend beyond your setting's participation in the programme. Your responses will be used to create descriptive statistics and individual settings will not be identified in this context. Your responses shall not be passed on to any third party.

What is the legal basis for processing your data?

The legal basis for RAND Europe to process your personal data is legitimate interests detailed in Article 6(1)(f) of the UK GDPR. To ensure that all processing is fair and lawful, RAND Europe have also completed a Legitimate Interest Assessment and a Data Protection Impact Assessment and have received ethical approval from the RAND internal review board. RAND Europe will process only what is required to meet these legal bases and will ensure security and safeguards are in place to protect the information.

What are your rights?

RAND Europe operates in accordance with the Data Protection Act 2018 and UK GDPR 2016 requirements. You are provided with certain rights that you may have the right to exercise through us. In summary those rights are:

- To access your data ("data subject access request") (Article 15 of the GDPR)
- To have inaccurate personal data rectified (Article 16 of the GDPR)

- To have your data erased (Article 17 of the GDPR)
- To restrict the processing of your data (Article 18 of the GDPR)
- Request the transfer of your personal data to you or to a third party (Article 20 of the GDPR)
- Object to processing of your personal data (Article 21 of the GDPR).

How do you contact us?

If you have any questions about this questionnaire or wish to exercise any of these rights, please contact the RAND Europe study team at theone@randeurope.org. Alternatively, you may contact the Data Protection Officer by email at REdpo@randeurope.org, or in writing to Data Protection Officer, RAND Europe, Eastbrook, Shaftesbury Road, Cambridge, CB2 8BF, UK. If contacting the Data Protection Officer, please make reference to the project The ONE and project number 022807.014 in your information request.

For independent advice about data protection or to lodge a complaint about how we have handled your personal data, you can contact the Information Commissioner's Office. You can visit www.ico.org.uk, email casework@ico.org.uk, or write to Information Commissioner's Office, Wycliffe House, Water Lane, Wilmslow, Cheshire, SK9 5AF, UK

Please click 'Next Page' if you would like to proceed. In doing so, you will be confirming that you have read the above Privacy Notice, accept its terms and consent to taking part in this questionnaire.

A. Background information

1. Please type the name of your nursery or pre-school in the space below. [open response]*

2. Please enter your nursery or pre-school's postal code. [open response]

3. This survey is intended for practitioners at settings that received The ONE intervention between January and May 2024.

Please select one answer.

a. I am an Early Years practitioner who works in a nursery or pre-school that received The ONE intervention between January and May 2024.

b. I am an Early Years practitioner who works in a nursery or pre-school that did not receive The ONE intervention between January and May 2024.

c. I am NOT an Early Years practitioner in a nursery or pre-school.

Note: If the respondent answered 'B' or 'C' to Q3, SURVEY ENDS.

If answered 'B' show the following message: "You have indicated that your nursey or pre-school did not receive The ONE intervention between January and May 2024. If this is the case, please email theone@randeurope.org to let us know you have received the incorrect link."

If answered 'C' show the following message: "Please forward the email with the link to this survey to an Early Years partitioner in your nursery or pre-school and ask them to complete it."

4. Did you receive and compete a similar survey from us during November/December 2023?

a) Yes

b) No

c) I can't remember

5. How many years of experience do you have working with children in early childhood education and care? Please input a number, no more than two digits. If you have less than one year's experience please input 0.

6. What are your current working hours at this nursery or pre-school?

Please select one answer.

a. Full-time working hours (working 35 hours or more per week)

b. Part-time working hours (working 17.5 – 34 hours per week)

c. Part-time working hours (working fewer than 17.5 hours per week)

7. Which of the following qualifications have you received?

Please select all that apply.

a. Level 2 qualification (e.g., Level 2 diploma for Early Years Education and Care)

b. Level 3 qualification (e.g., Level 3 diploma in Early Years Education and Care)

c. Certificate of Higher Education (e.g. a certificate in Early Years Education and Care from a university)

d. Undergraduate diploma (e.g., a diploma in the Early Years Education and Care from a university)

e. Undergraduate degree (e.g., BA in Early Childhood, BEd in Primary or Early Years, BSc in Educational and Developmental Psychology)

f. Postgraduate Certificate in Education (PGCE)

g. Postgraduate Diploma in Education (PGDE)

h. Postgraduate Teaching Apprenticeship

i. Postgraduate Degree (e.g., MA in early childhood education, PhD in Education)

j. Alternative early years or teacher training (e.g., Teach First, Now Teach)

k. Other type of training (please specify: _____)

8. In your initial early years qualifications (those you selected in the previous question), which of the following topics were covered?

Please select all that apply.

a. Child development (e.g., socio-emotional development, fine- and gross-motor development, cognitive development)

b. Executive function (e.g., working memory, inhibitory control, cognitive flexibility) and its development in childhood

c. Early language and literacy (e.g., speech and oral language development, phonemic awareness, phonics)

d. Early years' maths (e.g., early numeracy, pattern recognition)

e. Early years' science

f. None of the above

B. Details about Your Setting

Consider the classroom/s or playroom/s in which The ONE intervention was run in your setting. On a typical day, please indicate the number of children and staff in your setting that meet the following criteria:

| | |
|---|---|
| Children in the classroom/playroom which received The ONE | |
| Children in the classroom/playroom who are eligible for Early Years Pupil Premium (EYPP) | |
| Children in the classroom/playroom who have English as an Additional Language (EAL) | |
| Children in the classroom/playroom with known or suspected Special Educational Needs and Disabilities (SEND).. These include Speech, language and communication needs (SLCN); Personal, Social and Emotional Development needs (PSED); and social, emotional and mental health needs (SEMH) | |
| Staff in the classroom/playroom which received The ONE | |
| Staff in the classroom/playroom with a Level 3 Certificate of higher | |
| Staff in the classroom/playroom with a Bachelor's Degree or higher | |

17. Over **a typical working week**, how often do you do the following with the children in your classroom or playgroup:

| | Never | Once a week | 2-3 times a week | 4-5 times a week | More than once a day |
|---|---|---|---|---|---|
| I play games involving counting sets of objects (e.g., counting cups at meal times) | | | | | |
| I play games involving recognising numbers (e.g., matching numbers to dots) | | | | | |

| | | | | | |
|---|---|---|---|---|---|
| I play games involving sorting objects into groups (e.g., sorting leaves by size) | | | | | |
| I play games involving making or extending patterns (e.g., making towers with blocks in a red-blue-red-blue pattern) | | | | | |
| I play games involving recognising, naming or matching shapes (e.g., making or talking about shapes during craft activities) | | | | | |
| I encourage children to use prepositions to talk about space (e.g., asking children to describe locations using words like under/on/between instead of pointing) | | | | | |
| I ask children to complete activities that require multiple steps or paying attention to lengthy explanations | | | | | |
| I ask children to wait or take turns during everyday activities or games. | | | | | |
| I invite children to come up with new ways to solve a problem in an activity or game. | | | | | |
| I read to the children | | | | | |
| I play games involving letters with the children (e.g., tracing out letter shapes in the air, finding letters in a word) | | | | | |
| I play games involving different letter sounds with children (e.g., picking out certain sounds in a word) | | | | | |
| I play games focussed on understanding the world around us (e.g., activities about life cycles or the environment, exploring the weather) | | | | | |

18. In **a typical day,** how much time do you spend in activities or play with children in the following group sizes:

| | Never | Occasionally, but not every day | Less than 30 minutes a day | Between 30 minutes – 1 hour a day | Between 1 – 2 hours a day | More than 2 hours a day |
|---|---|---|---|---|---|---|
| Whole-group play or activities (typically involves all children in the classroom or playgroup) | | | | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| Small-group play or activities (typically involves a small subset of the normal classroom or playgroup) | | | | | | |
| One-to-one play or activities (typically involves working with just one child) | | | | | | |

C. Early Maths and Executive Function instruction, programmes and support available in your setting

Section C asks for your opinion on a variety of statements about the role of maths in learning, your confidence in teaching maths, the manner and frequency with which maths is taught in your setting, and the importance of certain skills in early maths learning. There will also be questions surrounding your experience (if any) with early maths and/or executive function programmes or interventions.

11. For each of the following statements, rate your agreement by checking the appropriate box.

[Skip question]

| | Strongly Disagree | Disagree | Neutral (neither agree nor disagree) | Agree | Strongly Agree |
|---|---|---|---|---|---|
| Maths is an important part of the early year's curriculum. | | | | | |
| Maths activities are a very important part of the early year's education. | | | | | |
| Development in academics such as maths is the goal of early year's education. | | | | | |
| Early year's maths activities help young children to be interested in maths when they go to school. | | | | | |
| Maths activities help improve young children's language skills. | | | | | |

12. For each of the following statements, rate your agreement by checking the appropriate box.

| | Strongly Disagree | Disagree | Neutral (neither agree nor disagree) | Agree | Strongly Agree |
|---|---|---|---|---|---|

| | | | | | |
|---|---|---|---|---|---|
| I feel comfortable planning maths activities for young children in my setting. | | | | | |
| I feel comfortable doing maths activities in my setting. | | | | | |
| I have enough maths knowledge to explore maths with young children. | | | | | |
| I feel comfortable using tools such as scales, measuring cups and rulers with young children in my setting. | | | | | |
| I enjoy doing maths activities with young children in my setting. | | | | | |
| I am afraid that young children might ask me a question about maths that I can't answer. | | | | | |
| I feel comfortable adapting maths activities to extend children's maths knowledge based on my observations of young children. | | | | | |

13. For each of the following statements, rate your agreement by checking the appropriate box.

| | Strongly Disagree | Disagree | Neutral (neither agree nor disagree) | Agree | Strongly Agree |
|---|---|---|---|---|---|
| It is important for my setting to have a maths area that can be freely explored by children. | | | | | |
| There is not enough time in the day to explore maths with young children. | | | | | |
| I use all kinds of materials for maths activities with young people. | | | | | |
| Preparation for maths education takes more time than other areas. | | | | | |
| I get ideas for maths activities from what my young children do, say and ask. | | | | | |
| I include some books about maths during story time. | | | | | |

14. How important are the following skills for early maths learning?

| | Not important | Somewhat important | Moderately important | Very important |
|---|---|---|---|---|
| Children know number facts. | | | | |
| Children understand mathematical learning. | | | | |

| | | | | |
|---|---|---|---|---|
| Children understand how mathematics is used in the real world. | | | | |
| Children can manipulate abstract information. | | | | |
| Children can store and manipulate information in their head. | | | | |
| Children can focus on relevant information and ignore distractions. | | | | |
| Children are able to think flexibly. | | | | |
| Children have good verbal skills. | | | | |
| Children are able to provide reasons to support their solutions. | | | | |
| Children are able to think creatively. | | | | |
| Children have good spatial skills. | | | | |

D. Implementing The ONE in your setting

15. For each of the following statements, rate your agreement by checking the appropriate box:

| | Strongly Disagree | Disagree | Neutral (neither agree nor disagree) | Agree | Strongly Agree |
|---|---|---|---|---|---|
| I enjoyed taking part in The ONE | | | | | |
| The ONE activities were easy to administer | | | | | |
| The children appeared to enjoy The ONE activities | | | | | |
| The ONE activities were suited to all 3-4 year olds in my setting | | | | | |
| The training prepared me to deliver The ONE | | | | | |
| I was able to deliver The ONE activities as intended | | | | | |

I was able to deliver The ONE activities as intended

16. Typically, how many The ONE activities did you administer in an average week?

a) One

b) Two

c) Three

d) Four

e) Five

f) Six

g) Seven

h) Eight

i) Nine

j) Ten

17. Typically, how much time did you spend on The ONE activities in the following group sizes:

| | Never | Occasionally, but not every day | Less than 30 minutes a day | Between 30 minutes – 1 hour a day | Between 1 – 2 hours a day | More than 2 hours a day |
|---|---|---|---|---|---|---|
| Whole-group play or activities (typically involves all children in the classroom or playgroup) | | | | | | |
| Small-group play or activities (typically involves a small subset of the normal classroom or playgroup) | | | | | | |
| One-to-one play or activities (typically involves working with just one child) | | | | | | |

18. In your professional experience, which was the most effective way to deliver The ONE programme to children?

Please select one answer.

a) One-to-one

b) In small groups

c) To the whole class

d) It depended on the activity

e) I don't know

19. Having implemented The ONE programme, which of the following statements would you agree with?;

Please select all that apply. You can select more than one.

a) I think that The ONE activities focused on improving counting and number skills

b) I think that The ONE activities focused on improving space and shape skills

c) I think that The ONE activities focused on improving patterns and ordering skills

20. How frequently did you embed tailored executive challenge into the activities in The ONE programme?

21. Please select one answer.

a) Occasionally, but not every day

b) For less than 30 minutes a day

c) For between 30 minutes – 1 hour a day

d) For between 1 – 2 hours a day

e) For more than 2 hours a day

22. At what point in the intervention did the embedding of executive challenge occur?

Please select all that apply.

a) During weeks 1-3

b) During week 4-6

c) During week 7-9

d) During week 10-12

23. To what extent do you disagree or agree: Adaptations beyond those noted in the activity sheets were needed for administering The ONE activities

Please select one answer.

a. Strongly disagree

b. Disagree

c. Neither agree nor disagree

d. Agree

e. Strongly agree

24. What adaptations did you make?

[Open response]

25. If you made adaptations, why did you make them?

[Open response]

26. To what extent did adaptations improve pupil responsiveness and outcomes?

Please select one answer.

a. To a great extent

b. Somewhat

c. Very little

d. Not at all

27. To what extent do you disagree or agree: Children in my setting were engaged with The ONE activities

Please select one answer.

f. Strongly disagree

g. Disagree

h. Neither agree nor disagree

i. Agree

j. Strongly agree

28. Which of the following facilitated or hindered engagement of children in The ONE activities?

| | Strongly facilitated | Facilitated | Neutral | Hindered | Strongly hindered |
|---|---|---|---|---|---|
| Activities being implemented in one-to-one environment | | | | | |
| Activities being implemented in small-group situations | | | | | |
| Activities being implemented in whole-class situations | | | | | |
| Frequency children are in the setting | | | | | |
| Distractions within the classroom | | | | | |
| Availability of other activities | | | | | |
| Children unable to engage with activities because of learning or language factors (i.e., SEND, EAL). | | | | | |

29. For each of the following statements, rate your agreement by checking the appropriate box:

| | Strongly Disagree | Disagree | Neutral (neither agree nor disagree) | Agree | Strongly Agree |
|---|---|---|---|---|---|
| I was able to carry out their normal duties whilst taking part in The ONE intervention | | | | | |
| The ONE intervention added additional pressure to my workload | | | | | |
| The ONE activities required significant time in preparation and planning | | | | | |
| As a result of The ONE intervention, I feel more confident to introduce numeracy concepts in everyday activities | | | | | |
| As a result of The ONE, I feel more confident to tailor other activities (unrelated to early maths) to a child's level of development | | | | | |
| As a result of The ONE, I feel more confident in building executive challenge into other activities (unrelated to early maths) | | | | | |
| Administering The ONE activities reduced the amount of time spent on early language and literacy activities | | | | | |
| Administering The ONE activities reduced the amount of time spent on early science activities | | | | | |
| The amount of time spent interacting with children in one-to-one conversations increased due to The ONE. | | | | | |
| The amount of time spent interacting with children in small group activities increased due to The ONE. | | | | | |
| The amount of time spent interacting with children in | | | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| whole-class activities increased due to The ONE. | | | | | | |

30. Please indicate the impact that the following potential facilitators had on delivery of the ONE in your setting:

| | Very negative impact (this stopped us from delivering The ONE almost entirely) | Slightly negative impact (stopped us from doing some sessions) | No impact | Slightly positive impact (this helped us to deliver some sessions successfully) | Very positive impact (this was a key factor that helped us to deliver The ONE successfully) | N/A or don't know |
|---|---|---|---|---|---|---|
| Staff retention/consistency during the intervention period | | | | | | |
| Support from the ONE team | | | | | | |
| Engagement of staff trained in the ONE | | | | | | |
| Professional development sufficiently explained the ONE intervention | | | | | | |
| Availability of protected time for staff professional development | | | | | | |
| Availability of protected time for preparation and planning of classroom/playroom activities | | | | | | |
| Availability of extra staff to cover professional development, preparation and planning time | | | | | | |
| Ease of adaptability of the setting routines and structure to accommodate the ONE activities | | | | | | |
| Protected time in the setting routine or | | | | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| structure for early numeracy activities | | | | | | |
| Materials and activities presented in the ONE were easy to understand | | | | | | |
| Sufficient materials were provided to run the ONE activities | | | | | | |
| Additional materials or resources required were readily available or easy to find | | | | | | |
| Technology required to participate in the ONE was readily available in the setting | | | | | | |

31. Please indicate the impact that the following potential barriers had on delivery of the ONE in your setting:

| | Very negative impact (this stopped us from delivering The ONE almost entirely) | Slightly negative impact (stopped us from doing some sessions) | No impact | Slightly positive impact (this helped us to deliver some sessions successfully) | Very positive impact (this was a key factor that helped us to deliver The ONE successfully) | N/A or don't know |
|---|---|---|---|---|---|---|
| Staff absences | | | | | | |
| Child absences | | | | | | |
| Staff turnover during the intervention period | | | | | | |
| Lack of support from the ONE team | | | | | | |
| Low levels of engagement from staff delivering The ONE | | | | | | |
| Difficulty in scheduling professional development sessions around other commitments | | | | | | |
| Difficulty in accessing the necessary materials needed to | | | | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| deliver the intervention | | | | | | |
| Difficulty in preparing and planning The ONE activities | | | | | | |
| Difficulty in adapting The ONE activities to the routines and structure of the day | | | | | | |
| Competing priorities in the setting making it difficult to deliver the required number of activities | | | | | | |
| Difficulty in maintaining child engagement in the activities | | | | | | |
| Difficulty in understanding the materials and activities presented in The ONE | | | | | | |
| Inappropriate delivery space (e.g., the room could not easily accommodate the activities in The ONE, the room was too loud or noisy) | | | | | | |
| Technological difficulties or lack of resources (e.g., computers could not be made available for professional development calls, tech resources limited preparation and administration of The ONE in the setting) | | | | | | |

32. Were there any other challenges you came across in implementing The ONE in your setting? [open response]

33. Were there any other facilitators that helped you to implement The ONE in your setting? [open response]

Thank you very much! The survey ends here. If you have any questions about this survey or your rights as a research participant, please do not hesitate to contact the evaluation team at RAND Europe on theone@randeurope.org.

# Appendix L: Control Setting Managers Endline Survey

Introduction

This survey is designed for those settings that did not receive The ONE between January and May 2024 (control settings). If your setting did participate in the intervention, please do not fill out this survey and email theone@randeurope.org to let us know you have received the wrong link.

This survey is part of an independent evaluation of the ONE intervention by RAND Europe. This evaluation is funded by the Education Endowment Foundation (EEF), an independent charity who fund research into 'what works' to improve educational practice. This is an endline survey for Setting Managers or Setting Headteachers working in nurseries or preschools participating in the evaluation of The ONE Programme.

This endline survey will ask you for information about you and your setting, as well as your business-as-usual in your setting. These endline surveys will enhance our understanding of different setting contexts, and how this may influence the implementation and effectiveness of The ONE Programme.

Survey responses will be analysed collectively.. No one setting or person will ever be identified in any analysis or report. All responses will be kept confidential. The data collected will be treated confidentially and shall not be shared with third parties. The data is processed and stored securely in accordance with the General Data Protection Regulation (GDPR). No personal data or IP addresses will be recorded or added to the dataset. For more information about how the data is processed and stored for The ONE Project, please refer to our Information Sheet for Early Years Practitioners and the Information Sheet for Early Years Settings. For any questions about the research and your rights as a participant, you may contact the data protection officer of RAND Europe. If you have any questions on this survey or your rights as a research participant, please do not hesitate to contact the evaluation team at RAND Europe on theone@randeurope.org.

We ask that you complete this endline survey within two weeks of receiving it.

By clicking 'Next Page' you consent to the evaluation team using your responses in their evaluation of The ONE Programme.

If you are happy to continue, please click 'Next Page'.

Privacy Notice

RAND Europe is an independent not-for-profit research institute based in Cambridge and Brussels whose mission is to help improve policy and decision making through evidence-based research. RAND Europe are data controllers for this project, and the Department for Education (DfE) and the Education Endowment Fund (EEF) are joint data controllers for the Stronger Practice Hubs to which the project belongs.

Below we set out how your personal information will be collected, used and looked after in accordance with the UK General Data Protection Regulations (GDPR) and Data Protection Act 2018.

What data are we collecting?

As the manager, head teacher, or person with a high-level overview of your setting, we are interested in asking you about your setting and about business as usual in your setting. This survey is being circulated to all control settings.

All of your responses to the survey will be confidential. The SmartSurvey platform automatically links your responses to your name and email address, but only the research team will have access to this data. The research team will only ever analyse and report the data at an anonymised and aggregate level and no individual will ever be named in any report

What are we using the data for?

The evaluation team is collecting information on your setting to aid the evaluation of The ONE. The evaluation aims to find out more about the impact that The ONE intervention has on pupil outcomes and how settings are implementing

The ONE. This questionnaire is being conducted at the end of the trial to understand your experience of business as usual, as a control setting.

We will look at how the responses of settings assigned to take part in The ONE compare to settings that are not implementing The ONE. We will analyse the data to see if there is a difference. This will help us to understand how and if The ONE contributed to changes over time, how this compares across settings implementing The ONE and those that are not, and any contextual factors that might help us understand the programme better.

How will we collect your data?

Oxford shared your contact details with us using an encrypted file shared via a secure file sharing platform called OneDrive.

Once the questionnaire is underway, your questionnaire responses will be collected and stored on the SmartSurvey platform by RAND Europe. RAND Europe will obtain the data securely from SmartSurvey. SmartSurvey will delete your questionnaire responses and identifiable data once RAND Europe has obtained it. RAND Europe will maintain this data in confidence and use it only for the purpose of evaluating The ONE. RAND Europe will save this data in a password protected folder on their internal network. Only nominated researchers on the project with the password will be able to access the file.

Please do not provide any sensitive data in this questionnaire, such as your political persuasion. If sensitive data is provided in the questionnaire, RAND Europe will delete it before analysis.

How do we keep your data secure?

The evaluation team have put various security measures in place to keep personal data secure and to prevent any unauthorised access to or use of it in accordance with Data Protection Act (2018) and UK GDPR requirements. All data collected by RAND Europe will be stored on secure servers, accessed only by relevant project team. No data will be saved on servers or shared with processors outside the UK.

How long do we keep your data?

The data will be stored securely on RAND Europe's data servers for the duration of The ONE evaluation project – from January 2023 to June 2025. To allow us time to analyse and report the results of the trial, this period will extend beyond your setting's participation in the programme. Your responses will be used to create descriptive statistics and individual settings will not be identified in this context. Your responses shall not be passed on to any third party.

What is the legal basis for processing your data?

The legal basis for RAND Europe to process your personal data is legitimate interests detailed in Article 6(1)(f) of the UK GDPR. To ensure that all processing is fair and lawful, RAND Europe have also completed a Legitimate Interest Assessment and a Data Protection Impact Assessment and have received ethical approval from the RAND internal review board. RAND Europe will process only what is required to meet these legal bases and will ensure security and safeguards are in place to protect the information.

What are your rights?

RAND Europe operates in accordance with the Data Protection Act 2018 and UK GDPR 2016 requirements. You are provided with certain rights that you may have the right to exercise through us. In summary those rights are:

- To access your data ("data subject access request") (Article 15 of the GDPR)
- To have inaccurate personal data rectified (Article 16 of the GDPR)
- To have your data erased (Article 17 of the GDPR)
- To restrict the processing of your data (Article 18 of the GDPR)
- Request the transfer of your personal data to you or to a third party (Article 20 of the GDPR)

- Object to processing of your personal data (Article 21 of the GDPR).

How do you contact us?

If you have any questions about this questionnaire or wish to exercise any of these rights, please contact the RAND Europe study team at theone@randeurope.org. Alternatively, you may contact the Data Protection Officer by email at REdpo@randeurope.org, or in writing to Data Protection Officer, RAND Europe, Eastbrook, Shaftesbury Road, Cambridge, CB2 8BF, UK. If contacting the Data Protection Officer, please make reference to the project The ONE and project number 022807.014 in your information request.

For independent advice about data protection or to lodge a complaint about how we have handled your personal data, you can contact the Information Commissioner's Office. You can visit www.ico.org.uk, email casework@ico.org.uk, or write to Information Commissioner's Office, Wycliffe House, Water Lane, Wilmslow, Cheshire, SK9 5AF, UK

Please click 'Next Page' if you would like to proceed. In doing so, you will be confirming that you have read the above Privacy Notice, accept its terms and consent to taking part in this questionnaire.

A. Background information

1) Please type the name of your nursery or pre-school in the space below. [open response]*

2) Please enter your nursery or pre-school's postal code. [open response]

3) This survey is intended for the Early Years Managers or Head Teachers at a setting that is waitlisted for The ONE intervention but has not yet received the intervention.

Please select one answer.

a. I am an Early Years Manager or Head who runs a nursery or pre-school that is waitlisted for The ONE.

b. I am an Early Years Manager or Head who runs a nursery or pre-school that did receive The ONE intervention between January and May 2024.

c. I am NOT an Early Years Manager or Head in my nursery or pre-school.

Note: If the respondent answered 'B' or 'C' to Q3, SURVEY ENDS.

If answered 'B' show the following message: "You have indicated that your nursery or pre-school did receive The ONE intervention between January and May 2024. If this is the case, please email theone@randeurope.org to let us know you have received the incorrect link."

If answered 'C' show the following message: "Please forward the email with the link to this survey to the Manager or Head of your nursery or pre-school and ask them to complete it."

4) How many years of experience do you have working with children in early childhood education and care? Please input a whole number, no more than two digits. If you have less than one year's experience please input 0.

5) How many years of experience do you have working as a manager or headteacher of an early years setting? Please input a whole number, no more than two digits. If you have less than one year's experience please input 0.

6) What are your current working hours at this nursery or pre-school?

Please select one answer.

a. Full-time working hours (working 35 hours or more per week)

b. Part-time working hours (working 17.5 – 34 hours per week)

c. Part-time working hours (working fewer than 17.5 hours per week)

7) Which of the following qualifications have you received?

Please select all that apply.

a. Level 2 qualification (e.g., Level 2 diploma for Early Years Education and Care)

b. Level 3 qualification (e.g., Level 3 diploma in Early Years Education and Care)

c. Certificate of Higher Education (e.g. a certificate in Early Years Education and Care from a university)

d. Undergraduate diploma (e.g., a diploma in the Early Years Education and Care from a university)

e. Undergraduate degree (e.g., BA in Early Childhood, BEd in Primary or Early Years, BSc in Educational and Developmental Psychology)

f. Postgraduate Certificate in Education (PGCE)

g. Postgraduate Diploma in Education (PGDE)

h. Postgraduate Teaching Apprenticeship

i. Postgraduate Degree (e.g., MA in early childhood education, PhD in Education)

j. Alternative early years or teacher training (e.g., Teach First, Now Teach)

k. Other type of training (please specify: _____)

B. Details about Your Setting

8) Please indicate whether you have children of the following ages in your Early Years setting (not including any attached primary or infant school)?

| | **Yes** | **No** |
|---|---|---|
| Under 1 year | ☐ | ☐ |
| 1-year-olds | ☐ | ☐ |
| 2-year-olds | ☐ | ☐ |
| 3-year-olds | ☐ | ☐ |
| 4-year-olds | ☐ | ☐ |
| Over 4 years | ☐ | ☐ |

9) In the last year, please indicate how many staff fall into the following categories:

(Please enter a number. If there are no staff that meet these criteria, please put 0.

If you head a maintained setting attached to a primary or infant school, please only include those staff working with children prior to starting Reception)

| | |
|---|---|
| Staff employed by the setting (in total, regardless of whether employed full- or part-time) | |
| Staff employed part-time in the setting | |
| Staff who have left your setting (either left employment in your setting or have gone on long-term leave, such as maternity leave) | |
| Staff who have newly joined your setting or returned from long-term leave, such as maternity leave | |

10) In the last year, please indicate the typical number of children and staff absences in your setting:

(Please input a whole number. Rough estimates are sufficient. If these numbers vary across weekdays, provide information for the weekday with the highest level of attendance for both children and staff)

| | |
|---|---|
| Number of child absences (average number of children not attending on a typical day,) | |
| Number of staff absences (average number of staff off for illness or other non-planned leave on a typical day) | |

11) Do you have the following in place for staff at your setting?

| | Yes | No |
|---|---|---|
| A training plan for staff members | | |
| A supervision plan for staff members (e.g., where staff are supervised by colleagues or management for the purposes of professional learning and development) | | |
| A budget specifically dedicated to professional development | | |

12) Over the last academic year, please indicate how many staff fall into the following categories:

(Please enter a number. If there are no staff that meet these criteria, please put 0.

If you head a maintained setting attached to a primary or infant school, please only include those staff working with children prior to starting Reception)

| | |
|---|---|
| Staff who were observed by colleagues or management for the purposes of professional learning and development | |

| | |
|---|---|
| Staff who received professional development or training | |
| Staff received professional development/training in child development (e.g., socio-emotional development, fine- and gross-motor development, cognitive development) | |
| Staff who received professional development/training in executive function (e.g., working memory, impulse control, cognitive flexibility) and its development in childhood | |
| Staff who received professional development/training in early language and literacy (e.g., speech and oral language development, phonemic awareness, phonics) | |
| Staff who received professional development/training in early years' maths (e.g., early numeracy, pattern recognition) | |
| Staff who received professional development/training in early years' science | |

C. Business as usual in your setting

We're interested in knowing more about the types of maths and executive functioning activities you do with the children in your settings.

13) How often do children in your setting engage in maths activities?

Please select one answer.

a) Several times a day

b) At least once a day

c) Several times a week

d) At least once a week

e) Never

14) What types of maths activities do you run for the children in your setting?

Please select all that apply.

a) Introducing maths into free play

b) 1:1 activities

c) Small group activities

d) Whole room activities

15) What do you use to help you plan and deliver maths activities (e.g., planning sheets, resources, professional development)?

[open text response]

16) How often do you consider children's executive functioning skills when planning for activities (e.g., memory games)?

Please select one answer.

a) Several times a day

b) At least once a day

c) Several times a week

d) At least once a week

e) Never

f) I am not familiar with the term 'executive functioning skills'

17) What types of executive functioning activities do you run for the children in your setting?

Please select all that apply.

a) 1:1 activities

b) Small group activities

c) Whole room activities

d) I am not familiar with the term 'executive functioning skills'

18) What do you use to help you plan and deliver the executive functioning activities?

[open text response] If you are not familiar with the term 'executive functioning' please leave blank.

19) How important do you think the following are in supporting children's early learning and development:

[include SKIP QUESTON option here]

| | Not important | Somewhat important | Moderately important | Very important |
|---|---|---|---|---|
| Early language and literacy | | | | |
| Early maths | | | | |

| | | | | |
|---|---|---|---|---|
| Executive function (being able to store information, being able to ignore distractions, being able to think flexibly) | | | | |
| Early science | | | | |

20) Please indicate the impact that the following potential facilitators had on the delivery of maths activities in your setting, in the last year:

| | Very negative impact (this stops us from delivering maths activities almost entirely) | Slightly negative impact (stops us from doing some maths activities) | No impact | Slightly positive impact (this helps us to deliver some maths activities successfully) | Very positive impact (this was a key factor that helps us to deliver maths activities successfully) | N/A or not sure |
|---|---|---|---|---|---|---|
| Staff retention/consistency during the last year | | | | | | |
| Support from colleagues | | | | | | |
| Engagement of staff | | | | | | |
| Professional development | | | | | | |
| Availability of protected time for staff professional development | | | | | | |
| Availability of protected time for preparation and planning of classroom/playroom activities | | | | | | |
| Availability of extra staff to cover professional development, preparation and planning time | | | | | | |
| Ease of adaptability of the setting routines and structure | | | | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| Protected time in the setting routine or structure for early numeracy activities | | | | | | |

21) Please indicate the impact that the following potential barriers had on the delivery of maths activities in your setting, in the last year:

| | Very negative impact (this stops us from delivering maths activities almost entirely) | Slightly negative impact (stops us from doing some maths activities) | No impact | Slightly positive impact (this helps us to deliver some maths activities successfully) | Very positive impact (this was a key factor that helps us to deliver maths activities successfully) | N/A or not sure |
|---|---|---|---|---|---|---|
| Staff absences | | | | | | |
| Child absences | | | | | | |
| Staff turnover | | | | | | |
| Low levels of engagement from staff | | | | | | |
| Difficulty in scheduling professional development sessions around other commitments | | | | | | |
| Difficulty in accessing the necessary materials | | | | | | |
| Difficulty in preparing and planning | | | | | | |
| Difficulty in maths activities to the routines and structure of the day | | | | | | |
| Competing priorities in the setting making it difficult to deliver activities | | | | | | |
| Difficulty in maintaining child engagement in activities | | | | | | |
| Difficulty in understanding materials and activities needed to deliver maths activities | | | | | | |
| Inappropriate delivery space (e.g., the room could not easily accommodate maths activities, the room was too loud or noisy) | | | | | | |
| Technological difficulties or lack of resources (e.g., computers could not be made available for | | | | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| professional development calls, tech resources limited) | | | | | | |

D. Business as usual costs in your setting

22) In an average week, how many hours do you spend supporting maths activities?

[Please input a whole number, no more than two digits] hours per week

23) In an average week, how many hours do you spend supporting executive functioning activities?

[Please input a number, no more than two digits] hours per week

If you are not familiar with the term 'executive functioning' please leave blank

24) We would like to understand costs associated with the resources needed to deliver your business as usual maths and executive functioning activities. Please list the relevant costs your setting has spent in the last year below, rounding to the nearest £10.

| **Description** | **Total monetary value** | **Bought for a specific intervention or programme or already had [bought/had]** |
|---|---|---|
| **e.g. CPD, resources, memberships, subscriptions, specialist teachers under costs to give them examples of what we mean.** | **e.g. £830** | **e.g. bought** |
| **[fill in]** | [fill in] | [fill in] |

Thank you very much! The survey ends here. If you have any questions about this survey or your rights as a research participant, please do not hesitate to contact the evaluation team at RAND Europe on theone@randeurope.org.

# Appendix M: Treatment Setting Managers Endline Survey

Introduction

This survey is designed for those settings that received The ONE intervention (4 weeks professional development and 12 weeks of intervention activities delivered within the classroom/playroom) between January and May 2024 (treatment settings). If your setting did not participate in the intervention, please do not fill out this survey and email theone@randeurope.org to let us know you have received the wrong link.

This survey is part of an independent evaluation of The ONE intervention by RAND Europe. This evaluation is funded by the Education Endowment Foundation (EEF), an independent charity who fund research into 'what works' to improve educational practice. This is an endline survey for managers or headteachers working in nurseries or preschools who delivered The ONE Programme.

This endline survey will ask you for information about you and your setting, as well as your experiences of The ONE intervention and what affects this may have had on the wider setting, and the costs involved in participating in The ONE. These endline surveys will enhance our understanding of different setting contexts, and how this may influence the implementation and effectiveness of The ONE Programme, and how much the ONE costs for settings to deliver, to help inform further evaluation and delivery of this programme.

Survey responses will be analysed collectively.. No one setting or person will ever be identified in any analysis or report. All responses will be kept confidential. The data collected will be treated confidentially and shall not be shared with third parties. The data is processed and stored securely in accordance with the General Data Protection Regulation (GDPR). No personal data or IP addresses will be recorded or added to the dataset. For more information about how the data is processed and stored for The ONE Project, please refer to our Information Sheet for Early Years Practitioners and the Information Sheet for Early Years Settings. For any questions about the research and your rights as a participant, you may contact the data protection officer of RAND Europe. If you have any questions on this survey or your rights as a research participant, please do not hesitate to contact the evaluation team at RAND Europe on theone@randeurope.org.

We ask that you complete this endline survey within two weeks of receiving it.

By clicking 'Next Page' you consent to the evaluation team using your responses in their evaluation of The ONE Programme.

If you are happy to continue, please click 'Next Page'.

Privacy Notice

RAND Europe is an independent not-for-profit research institute based in Cambridge and Brussels whose mission is to help improve policy and decision making through evidence-based research. RAND Europe are data controllers for this project, and the Department for Education (DfE) and the Education Endowment Fund (EEF) are joint data controllers for the Stronger Practice Hubs to which the project belongs.

Below we set out how your personal information will be collected, used and looked after in accordance with the UK General Data Protection Regulations (GDPR) and Data Protection Act 2018.

What data are we collecting?

As the manager, head teacher, or person with a high-level overview of your setting, we are interested in asking you about factors that may have impacted on delivery of The ONE your setting, as well as the associated costs and any unintended consequences that may have resulted from it. This survey is being circulated to all treatment settings involved in The ONE.

All of your responses to the survey will be confidential. The SmartSurvey platform automatically links your responses to your name and email address, but only the research team will have access to this data. The research team will only ever analyse and report the data at an anonymised and aggregate level and no individual will ever be named in any report

What are we using the data for?

The evaluation team is collecting information on your setting to aid the evaluation of The ONE. The evaluation aims to find out more about the impact that The ONE intervention has on pupil outcomes and how settings are implementing The ONE. This questionnaire is being conducted at the end of the trial to understand your experience of implementing the intervention in your setting. We will ask you about factors that may have influenced the implementation of The ONE in your setting, as well as the associated costs and any unintended consequences that may have resulted from it.

We will look at how the responses of settings assigned to take part in The ONE compare to settings that are not implementing The ONE. We will analyse the data to see if there is a difference. This will help us to understand how and if The ONE contributed to changes over time, how this compares across settings implementing The ONE and those that are not, and any contextual factors that might help us understand the programme better.

How will we collect your data?

Oxford shared your contact details with us using an encrypted file shared via a secure file sharing platform called OneDrive.

Once the questionnaire is underway, your questionnaire responses will be collected and stored on the SmartSurvey platform by RAND Europe. RAND Europe will obtain the data securely from SmartSurvey. SmartSurvey will delete your questionnaire responses and identifiable data once RAND Europe has obtained it. RAND Europe will maintain this data in confidence and use it only for the purpose of evaluating The ONE. RAND Europe will save this data in a password protected folder on their internal network. Only nominated researchers on the project with the password will be able to access the file.

Please do not provide any sensitive data in this questionnaire, such as your political persuasion. If sensitive data is provided in the questionnaire, RAND Europe will delete it before analysis.

How do we keep your data secure?

The evaluation team have put various security measures in place to keep personal data secure and to prevent any unauthorised access to or use of it in accordance with Data Protection Act (2018) and UK GDPR requirements. All data collected by RAND Europe will be stored on secure servers, accessed only by relevant project team. No data will be saved on servers or shared with processors outside the UK.

How long do we keep your data?

The data will be stored securely on RAND Europe's data servers for the duration of The ONE evaluation project – from January 2023 to June 2025. To allow us time to analyse and report the results of the trial, this period will extend beyond your setting's participation in the programme. Your responses will be used to create descriptive statistics and individual settings will not be identified in this context. Your responses shall not be passed on to any third party.

What is the legal basis for processing your data?

The legal basis for RAND Europe to process your personal data is legitimate interests detailed in Article 6(1)(f) of the UK GDPR. To ensure that all processing is fair and lawful, RAND Europe have also completed a Legitimate Interest Assessment and a Data Protection Impact Assessment and have received ethical approval from the RAND internal review board. RAND Europe will process only what is required to meet these legal bases and will ensure security and safeguards are in place to protect the information.

What are your rights?

RAND Europe operates in accordance with the Data Protection Act 2018 and UK GDPR 2016 requirements. You are provided with certain rights that you may have the right to exercise through us. In summary those rights are:

- To access your data ("data subject access request") (Article 15 of the GDPR)
- To have inaccurate personal data rectified (Article 16 of the GDPR)
- To have your data erased (Article 17 of the GDPR)
- To restrict the processing of your data (Article 18 of the GDPR)
- Request the transfer of your personal data to you or to a third party (Article 20 of the GDPR)
- Object to processing of your personal data (Article 21 of the GDPR).

How do you contact us?

If you have any questions about this questionnaire or wish to exercise any of these rights, please contact the RAND Europe study team at theone@randeurope.org. Alternatively, you may contact the Data Protection Officer by email at REdpo@randeurope.org, or in writing to Data Protection Officer, RAND Europe, Eastbrook, Shaftesbury Road, Cambridge, CB2 8BF, UK. If contacting the Data Protection Officer, please make reference to the project The ONE and project number 022807.014 in your information request.

For independent advice about data protection or to lodge a complaint about how we have handled your personal data, you can contact the Information Commissioner's Office. You can visit www.ico.org.uk, email casework@ico.org.uk, or write to Information Commissioner's Office, Wycliffe House, Water Lane, Wilmslow, Cheshire, SK9 5AF, UK

Please click 'Next Page' if you would like to proceed. In doing so, you will be confirming that you have read the above Privacy Notice, accept its terms and consent to taking part in this questionnaire.

A. Background information

1. Please type the name of your nursery or pre-school in the space below. [open response]*

2. Please enter your nursery or pre-school's postal code. [open response]

3. This survey is intended for the Early Years Managers or Head Teachers at a setting that received the ONE intervention between January and May 2024.

Please select one answer.

a. I am an Early Years Manager or Head who runs a nursery or pre-school that received the ONE intervention between January and May 2024.

b. I am an Early Years Manager or Head who runs a nursery or pre-school that did not receive the ONE intervention between January and May 2024.

c. I am NOT an Early Years Manager or Head in my nursery or pre-school.

Note: If the respondent answered 'B' or 'C' to Q3, SURVEY ENDS.

If answered 'B' show the following message: "You have indicated that your nursey or pre-school did not receive the ONE intervention between January and May 2024. If this is the case, please email theone@randeurope.org to let us know you have received the incorrect link."

If answered ‘C’ show the following message: “Please forward the email with the link to this survey to the Manager or Head of your nursery or pre-school and ask them to complete it.”

4. How many years of experience do you have working with children in early childhood education and care? Please input a whole number, no more than two digits. If you have less than one year’s experience please input 0.

5. How many years of experience do you have working as a manager or headteacher of an early years setting? Please input a whole number, no more than two digits. If you have less than one year’s experience please input 0.

6. What are your current working hours at this nursery or pre-school?

Please select one answer.

a. Full-time working hours (working 35 hours or more per week)

b. Part-time working hours (working 17.5 – 34 hours per week)

c. Part-time working hours (working fewer than 17.5 hours per week)

7. Which of the following qualifications have you received?

Please select all that apply.

a. Level 2 qualification (e.g., Level 2 diploma for Early Years Education and Care)

b. Level 3 qualification (e.g., Level 3 diploma in Early Years Education and Care)

c. Certificate of Higher Education (e.g. a certificate in Early Years Education and Care from a university)

d. Undergraduate diploma (e.g., a diploma in the Early Years Education and Care from a university)

e. Undergraduate degree (e.g., BA in Early Childhood, BEd in Primary or Early Years, BSc in Educational and Developmental Psychology)

f. Postgraduate Certificate in Education (PGCE)

g. Postgraduate Diploma in Education (PGDE)

h. Postgraduate Teaching Apprenticeship

i. Postgraduate Degree (e.g., MA in early childhood education, PhD in Education)

j. Alternative early years or teacher training (e.g., Teach First, Now Teach)

k. Other type of training (please specify: _____)

A. Details about Your Setting

8. Please indicate whether you have children of the following ages in your Early Years setting (not including any attached primary or infant school)?

| | **Yes** | **No** |
|---|---|---|
| Under 1 year | ☐ | ☐ |

| | | |
|---|---|---|
| 1-year-olds | ☐ | ☐ |
| 2-year-olds | ☐ | ☐ |
| 3-year-olds | ☐ | ☐ |
| 4-year-olds | ☐ | ☐ |
| Over 4 years | ☐ | ☐ |

9. In the last year, please indicate how many staff fall into the following categories:

(Please enter a number. If there are no staff that meet these criteria, please put 0.

If you head a maintained setting attached to a primary or infant school, please only include those staff working with children prior to starting Reception)

| | |
|---|---|
| Staff employed by the setting (in total, regardless of whether employed full- or part-time) | |
| Staff employed part-time in the setting | |
| Staff who have left your setting (either left employment in your setting or have gone on long-term leave, such as maternity leave) | |
| Staff who have newly joined your setting or returned from long-term leave, such as maternity leave | |

10. To what extent do you feel that participating in The ONE has impacted upon retention of staff in your setting?

Please select one answer.

a. The ONE was beneficial to staff retention

b. The ONE negatively impacted staff retention

c. The ONE did not make a difference to staff retention

If you selected a) or b), please provide more detail [open response]

11. In the last year, please indicate the typical number of children and staff absences in your setting:

(Please input a whole number. Rough estimates are sufficient. If these numbers vary across weekdays, provide information for the weekday with the highest level of attendance for both children and staff)

| | |
|---|---|
| Number of child absences (average number of children not attending on a typical day | |
| Number of staff absences (average number of staff off for illness or other non-planned leave on a typical day | |

12. Do you have the following in place for staff at your setting?

| | Yes | No |
|---|---|---|
| A training plan for staff members | | |

| | | |
|---|---|---|
| A supervision plan for staff members (e.g., where staff are supervised by colleagues or management for the purposes of professional learning and development) | | |
| A budget specifically dedicated to professional development | | |

13. Over the last academic year, excluding the ONE intervention, please indicate how many staff fall into the following categories:

(Please enter a number. If there are no staff that meet these criteria, please put 0.

If you head a maintained setting attached to a primary or infant school, please only include those staff working with children prior to starting Reception)

| | |
|---|---|
| Staff who were observed by colleagues or management for the purposes of professional learning and development | |
| Staff who received professional development or training | |
| Staff received professional development/training in child development (e.g., socio-emotional development, fine- and gross-motor development, cognitive development) | |
| Staff who received professional development/training in executive function (e.g., working memory, impulse control, cognitive flexibility) and its development in childhood | |
| Staff who received professional development/training in early language and literacy (e.g., speech and oral language development, phonemic awareness, phonics) | |
| Staff who received professional development/training in early years' maths (e.g., early numeracy, pattern recognition) | |
| Staff who received professional development/training in early years' science | |

B. Implementing the ONE in your setting

14. For each of the following statements, rate your agreement by checking the appropriate box:

| | Strongly Disagree | Disagree | Neutral (neither agree | Agree | Strongly Agree |
|---|---|---|---|---|---|

| | | | | | |
|---|---|---|---|---|---|
| | | | nor disagree) | | |
| The staff trained in my setting for the ONE were the best-placed staff to deliver the ONE activities to the children | | | | | |
| The staff enjoyed taking part in the ONE | | | | | |
| I have seen an improvement in the ability of the staff to support children with early numeracy and/or executive function | | | | | |
| The ONE activities were easy to deliver | | | | | |
| The ONE activities were suited to all 3-4 year olds in my setting | | | | | |

15. For each of the following statements, rate your agreement by checking the appropriate box:

| | Strongly Disagree | Disagree | Neutral (neither agree nor disagree) | Agree | Strongly Agree |
|---|---|---|---|---|---|
| The staff were able to carry out their normal duties whilst delivering The ONE | | | | | |
| The ONE intervention significantly reduced staff capacity for other activities | | | | | |
| The ONE activities required significant staff time in preparation and planning | | | | | |
| As a result of The ONE intervention, staff feel more confident to introduce numeracy concepts in everyday activities | | | | | |
| As a result of The ONE, staff feel more confident to adapt maths activities to embed more executive functioning | | | | | |
| As a result of The ONE, staff feel more confident to include play-based maths activities in weekly activities | | | | | |
| Delivering The ONE reduced the amount of time spent on | | | | | |

| | | | | | |
|---|---|---|---|---|---|
| other activities in the classroom (i.e., language and literacy, social skills) | | | | | |
| As a result of The ONE, staff feel more confident in other areas of their practice. | | | | | |

16. Please indicate the impact that the following potential facilitators had on delivery of the ONE in your setting:

| | Very negative impact (this stopped us from delivering The ONE almost entirely) | Slightly negative impact (stopped us from doing some sessions) | No impact | Slightly positive impact (this helped us to deliver some sessions successfully) | Very positive impact (this was a key factor that helped us to deliver The ONE successfully) | N/A or don't know |
|---|---|---|---|---|---|---|
| Staff retention/consistency during the intervention period | | | | | | |
| Support from the ONE team | | | | | | |
| Engagement of staff trained in the ONE | | | | | | |
| Professional development sufficiently explained the ONE intervention | | | | | | |
| Availability of protected time for staff professional development | | | | | | |
| Availability of protected time for preparation and planning of classroom/playroom activities | | | | | | |
| Availability of extra staff to cover professional development, preparation and planning time | | | | | | |
| Ease of adaptability of the setting routines and structure to | | | | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| accommodate the ONE activities | | | | | | |
| Protected time in the setting routine or structure for early numeracy activities | | | | | | |
| Materials and activities presented in the ONE were easy to understand | | | | | | |
| The materials provided to run the ONE activities (i.e., activity cards, resource box, plus additional help from delivery team member) | | | | | | |
| Technology required to participate in the ONE was readily available in the setting | | | | | | |

17. Please indicate the impact that the following potential barriers had on delivery of the ONE in your setting:

| | Very negative impact (this stopped us from delivering The ONE almost entirely) | Slightly negative impact (stopped us from doing some sessions) | No impact | Slightly positive impact (this helped us to deliver some sessions successfully) | Very positive impact (this was a key factor that helped us to deliver The ONE successfully) | N/a or don't know |
|---|---|---|---|---|---|---|
| Staff absences | | | | | | |
| Child absences | | | | | | |
| Staff turnover during the intervention period | | | | | | |
| Lack of support from the ONE team | | | | | | |
| Low levels of engagement from staff delivering The ONE | | | | | | |
| Difficulty in scheduling professional development | | | | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| sessions around other commitments | | | | | | |
| Difficulty in accessing the necessary materials needed to deliver the intervention | | | | | | |
| Difficulty in preparing and planning The ONE activities | | | | | | |
| Difficulty in adapting The ONE activities to the routines and structure of the day | | | | | | |
| Competing priorities in the setting making it difficult to deliver the required number of activities | | | | | | |
| Difficulty in maintaining child engagement in the activities | | | | | | |
| Difficulty in understanding the materials and activities presented in The ONE | | | | | | |
| Inappropriate delivery space (e.g., the room could not easily accommodate the activities in The ONE, the room was too loud or noisy) | | | | | | |
| Technological difficulties or lack of resources (e.g., computers could not be made available for professional development calls, tech resources limited preparation and administration | | | | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| of The ONE in the setting) | | | | | | |

18. Were there any other challenges you came across in implementing The ONE in your setting? [open response]

19. Were there any other facilitators that helped you to implement The ONE in your setting? [open response]

20. In your role as manager of head of the setting, to what extent were you able to support the delivery of the ONE within your setting?

Please select one answer.

a) To a great extent

b) Somewhat

c) Very little

d) Not at all

21. How important do you think the following are in supporting children's early learning and development:

[include SKIP QUESTON option here]

| | Not important | Somewhat important | Moderately important | Very important |
|---|---|---|---|---|
| Early language and literacy | | | | |
| Early maths | | | | |
| Executive function (being able to store information, being able to ignore distractions, being able to think flexibly) | | | | |
| Early science | | | | |

22. In your view, how likely are you to continue delivering the ONE activities in your setting?

Please select one answer.

a. Very likely

b. Somewhat likely

c. Not very likely

d. Very unlikely

23. In your view, how likely are to train new staff in the ONE activities?

Please select one answer.

a. Very likely

b. Somewhat likely

c. Not very likely

d. Very unlikely

24. How has participation in the ONE changed your plans for staff professional development this academic year (September 2023 – September 2024)?

a) Increased amount of professional development

b) Decreased amount of professional development

c) Changed the focus of professional development (e.g., increased the focus on early maths and decreased the focus on other topics)

d) Increased the variety of topics covered by professional development (e.g., early maths added to pre-existing planned professional development in other topics)

If you selected c) or d), please indicate which topics will be covered by professional development in your setting this academic year? [open response]

C. Costs of implementing the ONE in your setting

23. How many hours per week have you spent supporting The ONE activities?

[Please input a number, no more than two digits] hours per week

24. We would like to understand costs associated with the resources needed to deliver of The ONE including any resources you had already and any that you needed to buy specifically. Please list the relevant costs for your setting below, rounding to the nearest £10.

| **Description** | **Total monetary value** | **Bought for The ONE or already had [bought/had]** |
|---|---|---|
| **e.g. 2 computers** | **e.g. £830** | **e.g. bought** |
| **[fill in]** | [fill in] | [fill in] |

25. Apart from resources, did you have any other costs associated with delivering The ONE?

Please select one answer.

A) no

B) yes

If answered ‘A’ show the following message: “Thank you very much! The survey ends here. If you have any questions about this survey or your rights as a research participant, please do not hesitate to contact the evaluation team at RAND Europe on theone@randeurope.org.”

If answered ‘B’ ask the following question:

26. In addition to the resources you listed above, what other costs have you incurred? Please include a description of the cost(s) and how much you spent, rounding to the nearest £10.

[open text]

Thank you very much! The survey ends here. If you have any questions about this survey or your rights as a research participant, please do not hesitate to contact the evaluation team at RAND Europe on theone@randeurope.org.

# Appendix N: Analysis Syntax

Dd

```
#####################  The ONE Analysis: Secondary Analysis ##################

## [PURPOSE OF SCRIPT]:

####+  This script conducts the main secondary analysis for the ONE.
####+  We are exploring the impact of the intervention on 2 secondary outcome:
####+    1. Results of the Heads-Toes-Knees-Shoulders (HTKS-R) test for executive function
####+    2. Results of Corsi Blocks test for executive function

####+  These outcomes will be used to estimate 3 models
####+    1. A mixed-measures model, which uses raw HTKS-R scores at endline for the outcome and raw corsi blocks scores at baseline
####+    as the explanatory variable
####+    2. A HKTS-R model, which uses raw HTKS-R scores at endline for the outcome and raw HKTS-R scores at baseline
####+    as the explanatory variable
####+    3. A corsi blocks model, which uses raw corsi blocks scores at endline for the outcome and raw corsi blocks scores at baseline
####+    as the explanatory variable

####+    For each of these models, we need to do the following analysis
####+  1. Descriptive statistics (mean, sd, min/max) for
#+          a. Overall sample
#+          b. Analytical sample (i.e., sample which has no missing values for any
#+           of the model covariates)
#+      2. Run a mixed-effects random intercept model
#+      3. Test that key assumptions of the model are satisfied
#+      4. Generate an effect size
#+      5. Account for multiple comparisons problem of having multiple secondary outcomes using Romano-Wolf correction


#############################

# Clear workspace
rm(list=ls())
set.seed(999) #NB this is the same seed set in the randomisation code at the start of the project so we shall retain it here for consistency.

library("haven")
library("plyr")
library("dplyr")
library("ggplot2")
library("lme4")
library("sjstats")
library("Hmisc")
library("performance")
library("lmtest")
library("eeptools")
```

```
library("lmerTest")
library("officer")
library("boot")
library("boot.pval")
library("crctStepdown")

# Load cleaned data
data <- read_dta("//Nevis/RE/Projects/In_Progress/022807.014_P2899_EEF_The_ONE/WIP/06. Data collection/02.
Endline/02. Data_NotMirrored/BaselineEndlineComplianceClean2.dta")
data <- filter(data, endlineonly != 1)

######################################################################
###################### Mixed Measures Model ######################
######################################################################

# Section A: Descriptive statistics

#1. Histograms of outcomes

#Use this to check distribution of endline scores for our outcome

#We are mainly checking for ceiling effects, as flagged at baseline stage.

# In overall data
HTKS_ovr_hist <- hist(data$E_HTKS_total,
                      main="Histogram of HTKS subtest at endline (overall sample)",
                      xlab = "HTKS Scores",
                      xlim = c(0, 120),
                      breaks = seq(0, 120, 10),
                      xaxp = c(0, 120, 6))

#+In analytical data (the analytical data is defined as the data used in the multi-level regression model.
#+It includes the observations where we have complete data for all covariates included in the outcome model)
Sec1_model_hist <- hist(dplyr::filter(data,
                                      !is.na(B_Corsi_total) &
                                        !is.na(Settingtype) &
                                        !is.na(SettingRegion) &
                                        !is.na(SettingID))$E_HTKS_total,
                        main="Histogram of EYTN subtest at endline (analytical sample)",
                        xlab = "Early numeracy scores",
                        xlim = c(0, 120),
                        breaks = seq(0, 120, 10),
                        xaxp = c(0, 120, 6))

##No real floor or ceiling effects.
baseline_hist1 <- hist(data$B_Corsi_total,
                       main="Histogram of Corsi Blocks subtest at endline (overall sample)",
                       xlab = "Corsi Blocks Scores",
                       xlim = c(0, 15),
                       breaks = seq(0, 15, 3),
                       xaxp = c(0, 15, 5))

#+Some potential floor effects in the baseline control. Does warrant additional sensistivity analysis, but some
descriptive stats below.
```

```
sd <- sd(data$B_Corsi_total, na.rm = T)

# Create df which is a copy of analytical data
data_sd <- dplyr::filter(data,
                         !is.na(B_Corsi_total) &
                           !is.na(Settingtype) &
                           !is.na(SettingRegion) &
                           !is.na(SettingID))

data_sd <- data_sd %>%
  mutate(sd_dummy = 0)

data_sd$sd_dummy <- ifelse(data_sd$B_Corsi_total < sd, 1, 0)

mean(data_sd$sd_dummy)

# 40.04% within a standard deviation of the floor.

Hmisc::describe(dplyr::filter(data,
                          !is.na(E_HTKS_total))$E_HTKS_total)
sd(dplyr::filter(data,
              !is.na(E_HTKS_total))$E_HTKS_total)


##Analytical sample
Hmisc::describe(dplyr::filter(data,
                          !is.na(B_Corsi_total) &
                            !is.na(Settingtype) &
                            !is.na(SettingRegion) &
                            !is.na(E_HTKS_total) &
                            !is.na(SettingID))$E_HTKS_total)
sd(dplyr::filter(data,
              !is.na(B_Corsi_total) &
                !is.na(Settingtype) &
                !is.na(SettingRegion) &
                !is.na(E_HTKS_total) &
                !is.na(SettingID))$E_HTKS_total)


#And by treatment and control groups:
Hmisc::describe(dplyr::filter(data,  #treatment
                          treatment == 1 &
                            !is.na(B_Corsi_total) &
                            !is.na(Settingtype) &
                            !is.na(SettingRegion) &
                            !is.na(E_HTKS_total) &
                            !is.na(SettingID))$E_HTKS_total)

sd(dplyr::filter(data, #treatment
              treatment == 1 &
                !is.na(B_Corsi_total) &
                !is.na(Settingtype) &
                !is.na(SettingRegion) &
```

```
              !is.na(E_HTKS_total) &
              !is.na(SettingID))$E_HTKS_total)

Hmisc::describe(dplyr::filter(data,  #control
                        treatment == 0 &
                          !is.na(B_Corsi_total)  &
                          !is.na(Settingtype) &
                          !is.na(SettingRegion) &
                          !is.na(E_HTKS_total) &
                          !is.na(SettingID))$E_HTKS_total)
sd(dplyr::filter(data, #control
            treatment == 0 &
              !is.na(B_Corsi_total) &
              !is.na(Settingtype) &
              !is.na(SettingRegion) &
              !is.na(E_HTKS_total) &
              !is.na(SettingID))$E_HTKS_total)

##################################################
##################################################
##################################################

# Section B: Secondary outcome model 1

### 1. Run the multi-level model.

sec1_model <- lmer(E_HTKS_total ~
              treatment +
              B_Corsi_total +
              SettingRegion +
              Settingtype +
              (1 | SettingID),
            data=data, REML = FALSE)

summary(sec1_model) #Produce the results

performance::icc(sec1_model) #ICC

#########

### 2. Testing OLS assumptions

#Residual diagnostics: Testing normality of residuals OLS assumption
resid_pr <- resid(sec1_model) #Create object which stores the residuals of the model
plot_resid_pr <- plot(density(resid_pr)) #Kernel density plot to explore normality
shapiro.test(resid_pr) #Shapiro-wilk test for normality. !FLAG potentially K-S test better with larger sample sizes!
(p<0.05 rejects H0 that data is normally distributed)
ks.test(resid_pr, "pnorm") #K-S test (p<0.05 rejects H0 that data is normally distributed)

##Both of these tests reject H0

# QQ line
qq_line <- ggplot(data = data.frame(resid = resid_pr), aes(sample = resid)) +
  stat_qq() +
```

```
  stat_qq_line() +
  labs(title = "Q-Q Plot of Residuals", x = "Theoretical Quantiles", y = "Sample Quantiles") +
  theme_minimal()


#Plotting residual density against a normal distribution for comparison. NB this is using the residual objects created
from the secondary model above
sec1_resid_df <- as.data.frame(resid_pr)
sd_resid <- sd(resid_pr)
max_resid <- max(resid_pr)
min_resid <- min(resid_pr)

resid_sec1_kd<- ggplot(sec1_resid_df, aes(x = resid_pr)) +
  geom_density(aes(color = "Kernel density"), linewidth = 1) +
  stat_function(aes(color = "Normal density"),
                fun = function(x) dnorm(x, mean = 0, sd = sd_resid),
                linetype = "dotted", linewidth = 1) +
  theme_minimal() +
  scale_x_continuous(limits = c(min_resid, max_resid)) +
  scale_color_manual(name = "Legend", values = c("Kernel density" = "#000080", "Normal density" = "red")) +
  labs(
    title = "Executive Function Model 1 - Residuals Density Plot",
    x = "Residuals",
    y = "Density"
  )

#Testing linearity and assumption that residual errors have a mean of 0
plot(sec1_model, col = "red") #We want the line here to be horizontal and at 0

### Figures to export:
resid_sec1_kd
qq_line

plots <- list(resid_sec1_kd, qq_line)

### Export graphs to word

# Create a Word document
doc <- read_docx()

# Save each plot as an image and add to the Word document
for (i in seq_along(plots)) {
  # Export the plot to a temporary file
  file_name <- tempfile(fileext = ".png")
  ggplot2::ggsave(filename = file_name, plot = plots[[i]])

  # Add the plot to the Word document
  doc <- doc %>%
    body_add_par(value = paste("Plot", i), style = "heading 1") %>%
    body_add_img(src = file_name, width = 6, height = 4)
}

# Save the document
```

```
print(doc, target = "//Nevis/RE/Projects/In_Progress/022807.014_P2899_EEF_The_ONE/WIP/09. Impact/Graphs & Summary Tables/secondary_outcome_model1.docx")

### 3.  Bootstrapping CIs and p-value
#+ The plots and normality test results suggest
#+ that the residuals are not normally distributed. We will
#+ therefore need to re-estimate CIs and p-values using bootstrapping.

set.seed(999)
boot_sec1_model <- boot.pval::boot_summary(sec1_model,
                                type = "norm",
                                method = NULL,
                                conf.level = 0.95)
boot_sec1_model

######################



### 4.  Effect-size estimation

hedges.g <- function(c, n, m, v, w){
  c/sqrt((((n - 1)*v)+((m - 1)* w))/(n + m - 2))
}

#We use the output from the model above to calculate the effect size: (NB: This needs to be done for confidence intervals too!)
#We need to define the objects needed to be inputted into the Hedges G function created above (c, n, m, v, & w)

coefs <- data.frame(summary(sec1_model)$coefficients) #create data frame of all coefficients from secondary outcome model

c <- coefs["treatment", "Estimate"] #Extract the treatment coefficient

n <- nrow(dplyr::filter(data,
                  treatment == 1 &
                    !is.na(E_HTKS_total) &
                    !is.na(B_Corsi_total) &
                    !is.na(SettingRegion) &
                    !is.na(Settingtype) &
                    !is.na(SettingID))) #Number of individuals in treatment group in the model

m <- nrow(dplyr::filter(data,
                  treatment == 0 &
                    !is.na(E_HTKS_total) &
                    !is.na(B_Corsi_total) &
                    !is.na(SettingRegion) &
                    !is.na(Settingtype) &
                    !is.na(SettingID))) #Number of individuals in control group in the model

v <- var(dplyr::filter(data,
                 treatment == 1 &
                   !is.na(B_Corsi_total) &
```

```
                  !is.na(SettingRegion) &
                  !is.na(Settingtype) &
                  !is.na(E_HTKS_total) &
                  !is.na(SettingID))$E_HTKS_total) #Variance in outcome among treatment group

w <- var(dplyr::filter(data,
                  treatment == 0 &
                  !is.na(B_Corsi_total) &
                  !is.na(SettingRegion) &
                  !is.na(Settingtype) &
                  !is.na(E_HTKS_total) &
                  !is.na(SettingID))$E_HTKS_total)  #Variance in outcome among control group

secondary1_effect_size <- hedges.g(c, n, m, v, w)
secondary1_effect_size

#LOWER CI EFFECT SIZE using lower 95% CI from the bootstrapped model:

#We redefine c to equal the treatment coefficient minus the SE created above
c <- boot_sec1_model$Lower.bound[2]

sec1_effect_size_low <- hedges.g(c, n, m, v, w) #Re-run hedges G on lower CI treatment coefficient
sec1_effect_size_low #This is the lower confidence interval for the treatment effect on the following directions subtest.



#UPPER CI EFFECT SIZE using bootrapped 95% upper CI from the above model:

#We redefine c again to equal the treatment coefficient plus the SE created above. Remember c has already been
defined to be the lower confidence interval, we therefore need to add 2 SEs to it.
c <- boot_sec1_model$Upper.bound[2]

sec1_effect_size_high <- hedges.g(c, n, m, v, w) #Re-run hedges G on upper CI treatment coefficient
sec1_effect_size_high #This is the upper confidence interval for the treatment effect on the following directions
subtest.

## P-value
p_values <- boot_sec1_model$p.value[2]
p_value <- round(p_values, digits = 2)


### 5. Creating secondary analysis table (to paste output into report)

##+ This part of the code is to create a formatted output table which can just be copied
##+ into the table empty table in word

## Find missing numbers for model (number of obs with outcome but missing covariates)

# Treatment
n_mis <- nrow(dplyr::filter(data,
                    treatment == 1 &
                      !is.na(E_HTKS_total) &
                      (is.na(B_Corsi_total) |
                         is.na(SettingRegion) |
```

```
                        is.na(Settingtype) |
                        is.na(SettingID))))

# Control
m_mis <- nrow(dplyr::filter(data,
                        treatment == 0 &
                         !is.na(E_HTKS_total) &
                         (is.na(B_Corsi_total) |
                            is.na(SettingRegion) |
                            is.na(Settingtype) |
                            is.na(SettingID))))

# Paste in numbers (non-missing and missing) for table
num_t <- paste0(as.character(n), " (", as.character(n_mis), ")")
num_c <- paste0(as.character(m), " (", as.character(m_mis), ")")

# Total numbers
total_n <- n + m
total_n_t_c <- paste0(as.character(total_n), " (", n, ";", m, ")")

## Generate means and CIs

# Treatment
me_t <- t.test(dplyr::filter(data,
                        treatment == 1 &
                          !is.na(E_HTKS_total) &
                          !is.na(B_Corsi_total)  &
                          !is.na(SettingRegion) &
                          !is.na(Settingtype) &
                          !is.na(SettingID))$E_HTKS_total) #Run t-test on outcome in analytical sample

mean_t <- as.character(round(me_t$estimate, digits = 2)) #extract mean from t-test object (me), to 2dp and as
character to facilitate pasting in CIs.

me_ci_t <- as.character(round(me_t$conf.int[c(1, 2)], digits = 2)) #extract CIs (95%) from t-test object (me), to 2dp and
as character to facilitate pasting

mean_ci_t <- paste0(mean_t, " (", me_ci_t[1], " - ", me_ci_t[2], ")") #Paste CIs in brackets next to mean figure

# Control
me_c <- t.test(dplyr::filter(data,
                        treatment == 0 &
                          !is.na(E_HTKS_total) &
                          !is.na(B_Corsi_total)  &
                          !is.na(SettingRegion) &
                          !is.na(Settingtype) &
                          !is.na(SettingID))$E_HTKS_total)

mean_c <- as.character(round(me_c$estimate, digits = 2))
me_ci_c <- as.character(round(me_c$conf.int[c(1, 2)], digits = 2))
mean_ci_c <- paste0(mean_c, " (", me_ci_c[1], " - ", me_ci_c[2], ")")


## Effect size with CIs
```

```
ES_CI <- paste0(as.character(round(secondary1_effect_size, digits = 3)),
                " (", as.character(round(sec1_effect_size_low, digits = 2)),
                " -- ",
                as.character(round(sec1_effect_size_high, digits = 2)),
                ")")

## Create table

secondary_analysis_table_1 <- t(as.data.frame(c(num_t, mean_ci_t, num_c, mean_ci_c, total_n_t_c, ES_CI,
p_value)))
colnames(secondary_analysis_table_1) <- c("N (intervention)", "Mean (intervention", "N (control)", "Mean (control)",
"Total (T;C)", "Hedges g (95% CIs)", "p-value")

rownames(secondary_analysis_table_1) <- "HTKS Executive Function Test"

## Display results - TABLE 1
View(secondary_analysis_table_1)

############

### 6. Creating effect size estimation table. Most of the required fields have already been created

u_m <- as.character(round((mean(dplyr::filter(data, #Unadjusted difference in means
                                    treatment == 1 &
                                      !is.na(B_Corsi_total)  &
                                      !is.na(SettingRegion) &
                                      !is.na(Settingtype) &
                                      !is.na(SettingID))$E_HTKS_total, na.rm = TRUE)) - (mean(dplyr::filter(data,
                                                                    treatment == 0 &
                                                                      !is.na(B_Corsi_total) &
                                                                      !is.na(SettingRegion) &
                                                                      !is.na(Settingtype) &
                                                                      !is.na(SettingID))$E_HTKS_total, na.rm =
TRUE)), digits = 2))

c <- round(coefs["treatment", "Estimate"], digits = 2) #Adjusted difference in means

var_t <- round(v, digits = 2) #Variance of outcome in treatment

var_c <- round(w, digits = 2) #Variance of outcome in control

pooled_var <- round(((((n - 1)*v)+((m - 1)* w))/(n + m - 2)), digits = 2) #pooled standard deviation



# Create table 2
sec1_es_est <- t(as.data.frame(c(u_m, c, num_t, var_t, num_c, var_c, pooled_var)))
colnames(sec1_es_est) <- c("Unadj. diff in means", "Adj diff in means", "N (treatment)", "Variance (treatment)", "N
(control)", "Variance (control)", "Pooled Variance")
rownames(sec1_es_est) <- "HTKS Executive Function Test"
View(sec1_es_est)


#########################################################
```

```
#####################################################
################### HTKS-R Model ####################
#####################################################

# Section A: Descriptive statistics

#1. Histograms of outcomes

#Use this to check distribution of endline scores for our outcome

#We are mainly checking for ceiling effects, as flagged at baseline stage.

# Already have histogram for overall data

#+In analytical data (the analytical data is defined as the data used in the multi-level regression model.
#+It includes the observations where we have complete data for all covariates included in the outcome model)
##Endline
Sec2_model_hist <- hist(dplyr::filter(data,
                                      !is.na(B_HTKS_total) &
                                        !is.na(Settingtype) &
                                        !is.na(SettingRegion) &
                                        !is.na(SettingID))$E_HTKS_total,
                        main="Histogram of HTKS subtest at endline (analytical sample)",
                        xlab = "HTKS scores",
                        xlim = c(0, 120),
                        breaks = seq(0, 120, 10),
                        xaxp = c(0, 120, 6))
##Basline
Sec2_model_histB <- hist(dplyr::filter(data,
                                      !is.na(B_HTKS_total) &
                                        !is.na(Settingtype) &
                                        !is.na(SettingRegion) &
                                        !is.na(SettingID))$B_HTKS_total,
                        main="Histogram of HTKS subtest at baseline (analytical sample)",
                        xlab = "HTKS scores",
                        xlim = c(0, 120),
                        breaks = seq(0, 120, 10),
                        xaxp = c(0, 120, 6))

##No real floor or ceiling effects.

Hmisc::describe(dplyr::filter(data,
                              !is.na(E_HTKS_total))$E_HTKS_total)
sd(dplyr::filter(data,
                 !is.na(E_HTKS_total))$E_HTKS_total)


##Analytical sample
Hmisc::describe(dplyr::filter(data,
                              !is.na(B_HTKS_total) &
                                !is.na(Settingtype) &
                                !is.na(SettingRegion) &
                                !is.na(E_HTKS_total) &
```

```
                          !is.na(SettingID))$E_HTKS_total)
sd(dplyr::filter(data,
                 !is.na(B_HTKS_total) &
                   !is.na(Settingtype) &
                   !is.na(SettingRegion) &
                   !is.na(E_HTKS_total) &
                   !is.na(SettingID))$E_HTKS_total)


#And by treatment and control groups:
Hmisc::describe(dplyr::filter(data,  #treatment
                          treatment == 1 &
                            !is.na(B_HTKS_total) &
                            !is.na(Settingtype) &
                            !is.na(SettingRegion) &
                            !is.na(E_HTKS_total) &
                            !is.na(SettingID))$E_HTKS_total)

sd(dplyr::filter(data, #treatment
                 treatment == 1 &
                   !is.na(B_HTKS_total) &
                   !is.na(Settingtype) &
                   !is.na(SettingRegion) &
                   !is.na(E_HTKS_total) &
                   !is.na(SettingID))$E_HTKS_total)

Hmisc::describe(dplyr::filter(data,  #control
                          treatment == 0 &
                            !is.na(B_HTKS_total) &
                            !is.na(Settingtype) &
                            !is.na(SettingRegion) &
                            !is.na(E_HTKS_total) &
                            !is.na(SettingID))$E_HTKS_total)
sd(dplyr::filter(data, #control
                 treatment == 0 &
                   !is.na(B_HTKS_total) &
                   !is.na(Settingtype) &
                   !is.na(SettingRegion) &
                   !is.na(E_HTKS_total) &
                   !is.na(SettingID))$E_HTKS_total)

##################################################
##################################################
##################################################

# Section B: Secondary outcome model 1

### 1. Run the multi-level model.

sec2_model <- lmer(E_HTKS_total ~
                     treatment +
                     B_HTKS_total +
                     SettingRegion +
                     Settingtype +
```

```
                (1 | SettingID),
              data=data, REML = FALSE)

summary(sec2_model) #Produce the results

performance::icc(sec2_model) #ICC

#########

### 2. Testing OLS assumptions

#Residual diagnostics: Testing normality of residuals OLS assumption
resid_sec2 <- resid(sec2_model) #Create object which stores the residuals of the model
plot_resid_sec2 <- plot(density(resid_sec2)) #Kernel density plot to explore normality
shapiro.test(resid_sec2) #Shapiro-wilk test for normality. !FLAG potentially K-S test better with larger sample sizes! (p<0.05 rejects H0 that data is normally distributed)
ks.test(resid_sec2, "pnorm") #K-S test (p<0.05 rejects H0 that data is normally distributed)

##Both of these tests reject H0

# QQ line
qq_line2 <- ggplot(data = data.frame(resid = resid_sec2), aes(sample = resid)) +
  stat_qq() +
  stat_qq_line() +
  labs(title = "Q-Q Plot of Residuals", x = "Theoretical Quantiles", y = "Sample Quantiles") +
  theme_minimal()


#Plotting residual density against a normal distribution for comparison. NB this is using the residual objects created from the secondary model above
sec2_resid_df <- as.data.frame(resid_sec2)
sd_resid_sec2 <- sd(resid_sec2)
max_resid_sec2 <- max(resid_sec2)
min_resid_sec2 <- min(resid_sec2)

resid_sec2_kd<- ggplot(sec2_resid_df, aes(x = resid_sec2)) +
  geom_density(aes(color = "Kernel density"), linewidth = 1) +
  stat_function(aes(color = "Normal density"),
                fun = function(x) dnorm(x, mean = 0, sd = sd_resid_sec2),
                linetype = "dotted", linewidth = 1) +
  theme_minimal() +
  scale_x_continuous(limits = c(min_resid_sec2, max_resid_sec2)) +
  scale_color_manual(name = "Legend", values = c("Kernel density" = "#000080", "Normal density" = "red")) +
  labs(
    title = "Executive Function Model 2 - Residuals Density Plot",
    x = "Residuals",
    y = "Density"
  )

#Testing linearity and assumption that residual errors have a mean of 0
plot(sec2_model, col = "red") #We want the line here to be horizontal and at 0

### Figures to export:
resid_sec2_kd
```

```
qq_line2

plots <- list(resid_sec2_kd, qq_line2)

### Export graphs to word

# Create a Word document
doc <- read_docx()

# Save each plot as an image and add to the Word document
for (i in seq_along(plots)) {
  # Export the plot to a temporary file
  file_name <- tempfile(fileext = ".png")
  ggplot2::ggsave(filename = file_name, plot = plots[[i]])

  # Add the plot to the Word document
  doc <- doc %>%
    body_add_par(value = paste("Plot", i), style = "heading 1") %>%
    body_add_img(src = file_name, width = 6, height = 4)
}

# Save the document
print(doc, target = "//Nevis/RE/Projects/In_Progress/022807.014_P2899_EEF_The_ONE/WIP/09. Impact/Graphs &
Summary Tables/secondary_outcome_model2.docx")

### 3.  Bootstrapping CIs and p-value
#+ The plots and normality test results suggest
#+ that the residuals are not normally distributed. We will
#+ therefore need to re-estimate CIs and p-values using bootstrapping.

set.seed(999)
boot_sec2_model <- boot.pval::boot_summary(sec2_model,
                                  type = "norm",
                                  method = NULL,
                                  conf.level = 0.95)
boot_sec2_model

######################



### 4.  Effect-size estimation

coefs2 <- data.frame(summary(sec2_model)$coefficients) #create data frame of all coefficients from secondary
outcome model

c <- coefs2["treatment", "Estimate"] #Extract the treatment coefficient

n <- nrow(dplyr::filter(data,
                  treatment == 1 &
                    !is.na(E_HTKS_total) &
                    !is.na(B_HTKS_total) &
                    !is.na(SettingRegion) &
```

```
                  !is.na(Settingtype) &
                  !is.na(SettingID))) #Number of individuals in treatment group in the model

m <- nrow(dplyr::filter(data,
                   treatment == 0 &
                    !is.na(E_HTKS_total) &
                    !is.na(B_HTKS_total) &
                    !is.na(SettingRegion) &
                    !is.na(Settingtype) &
                    !is.na(SettingID))) #Number of individuals in control group in the model

v <- var(dplyr::filter(data,
                  treatment == 1 &
                    !is.na(B_HTKS_total) &
                    !is.na(SettingRegion) &
                    !is.na(Settingtype) &
                    !is.na(E_HTKS_total) &
                    !is.na(SettingID))$E_HTKS_total) #Variance in outcome among treatment group

w <- var(dplyr::filter(data,
                  treatment == 0 &
                    !is.na(B_HTKS_total) &
                    !is.na(SettingRegion) &
                    !is.na(Settingtype) &
                    !is.na(E_HTKS_total) &
                    !is.na(SettingID))$E_HTKS_total)  #Variance in outcome among control group

secondary2_effect_size <- hedges.g(c, n, m, v, w)
secondary2_effect_size

#LOWER CI EFFECT SIZE using lower 95% CI from the bootstrapped model:

#We redefine c to equal the treatment coefficient minus the SE created above
c <- boot_sec2_model$Lower.bound[2]

sec2_effect_size_low <- hedges.g(c, n, m, v, w) #Re-run hedges G on lower CI treatment coefficient
sec2_effect_size_low #This is the lower confidence interval for the treatment effect on the following directions subtest.



#UPPER CI EFFECT SIZE using bootrapped 95% upper CI from the above model:

#We redefine c again to equal the treatment coefficient plus the SE created above. Remember c has already been
defined to be the lower confidence interval, we therefore need to add 2 SEs to it.
c <- boot_sec2_model$Upper.bound[2]

sec2_effect_size_high <- hedges.g(c, n, m, v, w) #Re-run hedges G on upper CI treatment coefficient
sec2_effect_size_high #This is the upper confidence interval for the treatment effect on the following directions
subtest.

## P-value
p_values <- boot_sec2_model$p.value[2]
p_value <- round(p_values, digits = 2)
```

```
### 5. Creating secondary analysis table (to paste output into report)

##+ This part of the code is to create a formatted output table which can just be copied
##+ into the table empty table in word

## Find missing numbers for model (number of obs with outcome but missing covariates)

# Treatment
n_mis <- nrow(dplyr::filter(data,
                            treatment == 1 &
                              !is.na(E_HTKS_total) &
                              (is.na(B_HTKS_total) |
                                 is.na(SettingRegion) |
                                 is.na(Settingtype) |
                                 is.na(SettingID))))

# Control
m_mis <- nrow(dplyr::filter(data,
                            treatment == 0 &
                              !is.na(E_HTKS_total) &
                              (is.na(B_HTKS_total) |
                                 is.na(SettingRegion) |
                                 is.na(Settingtype) |
                                 is.na(SettingID))))

# Paste in numbers (non-missing and missing) for table
num_t <- paste0(as.character(n), " (", as.character(n_mis), ")")
num_c <- paste0(as.character(m), " (", as.character(m_mis), ")")

# Total numbers
total_n <- n + m
total_n_t_c <- paste0(as.character(total_n), " (", n, ";", m, ")")

## Generate means and CIs

# Treatment
me_t <- t.test(dplyr::filter(data,
                             treatment == 1 &
                               !is.na(E_HTKS_total) &
                               !is.na(B_HTKS_total) &
                               !is.na(SettingRegion) &
                               !is.na(Settingtype) &
                               !is.na(SettingID))$E_HTKS_total) #Run t-test on outcome in analytical sample

mean_t <- as.character(round(me_t$estimate, digits = 2)) #extract mean from t-test object (me), to 2dp and as character to facilitate pasting in CIs.

me_ci_t <- as.character(round(me_t$conf.int[c(1, 2)], digits = 2)) #extract CIs (95%) from t-test object (me), to 2dp and as character to facilitate pasting

mean_ci_t <- paste0(mean_t, " (", me_ci_t[1], " - ", me_ci_t[2], ")") #Paste CIs in brackets next to mean figure

# Control
```

```
me_c <- t.test(dplyr::filter(data,
                        treatment == 0 &
                          !is.na(E_HTKS_total) &
                          !is.na(B_HTKS_total) &
                          !is.na(SettingRegion) &
                          !is.na(Settingtype) &
                          !is.na(SettingID))$E_HTKS_total)

mean_c <- as.character(round(me_c$estimate, digits = 2))
me_ci_c <- as.character(round(me_c$conf.int[c(1, 2)], digits = 2))
mean_ci_c <- paste0(mean_c, " (", me_ci_c[1], " - ", me_ci_c[2], ")")


## Effect size with CIs
ES_CI <- paste0(as.character(round(secondary2_effect_size, digits = 3)),
            " (", as.character(round(sec2_effect_size_low, digits = 2)),
            " -- ",
            as.character(round(sec2_effect_size_high, digits = 2)),
            ")")

## Create table

secondary_analysis_table_2 <- t(as.data.frame(c(num_t, mean_ci_t, num_c, mean_ci_c, total_n_t_c, ES_CI,
p_value)))
colnames(secondary_analysis_table_2) <- c("N (intervention)", "Mean (intervention", "N (control)", "Mean (control)",
"Total (T;C)", "Hedges g (95% CIs)", "p-value")

rownames(secondary_analysis_table_2) <- "HTKS Executive Function Test"

## Display results - TABLE 1
View(secondary_analysis_table_2)

############

### 6. Creating effect size estimation table. Most of the required fields have already been created

u_m <- as.character(round((mean(dplyr::filter(data, #Unadjusted difference in means
                                    treatment == 1 &
                                      !is.na(B_HTKS_total) &
                                      !is.na(SettingRegion) &
                                      !is.na(Settingtype) &
                                      !is.na(SettingID))$E_HTKS_total, na.rm = TRUE)) - (mean(dplyr::filter(data,
                                                                                            treatment == 0 &
                                                                                              !is.na(B_HTKS_total) &
                                                                                              !is.na(SettingRegion) &
                                                                                              !is.na(Settingtype) &
                                                                                              !is.na(SettingID))$E_HTKS_total, na.rm =
TRUE)), digits = 2))

c <- round(coefs2["treatment", "Estimate"], digits = 2) #Adjusted difference in means

var_t <- round(v, digits = 2) #Variance of outcome in treatment

var_c <- round(w, digits = 2) #Variance of outcome in control
```

```
pooled_var <- round(((((n - 1)*v)+((m - 1)* w))/(n + m - 2)), digits = 2) #pooled standard deviation



# Create table 2
sec2_es_est <- t(as.data.frame(c(u_m, c, num_t, var_t, num_c, var_c, pooled_var)))
colnames(sec2_es_est) <- c("Unadj. diff in means", "Adj diff in means", "N (treatment)", "Variance (treatment)", "N
(control)", "Variance (control)", "Pooled Variance")
rownames(sec2_es_est) <- "HTKS Executive Function Test"
View(sec2_es_est)


########


#########################################################

#########################################################
#################### Corsi Model ####################
#########################################################

# Section A: Descriptive statistics

#1. Histograms of outcomes

#Use this to check distribution of endline scores for our outcome

#We are mainly checking for ceiling effects, as flagged at baseline stage.

# Overall Corsi Endline
Corsi_ovr_hist <- hist(data$E_Corsi_total,
                       main="Histogram of Corsi Block subtest at endline (overall sample)",
                       xlab = "Corsi Block scores",
                       xlim = c(0, 15),
                       breaks = seq(0, 15, 3),
                       xaxp = c(0, 15, 5))

#+In analytical data (the analytical data is defined as the data used in the multi-level regression model.
#+It includes the observations where we have complete data for all covariates included in the outcome model)

Sec3_model_hist <- hist(dplyr::filter(data,
                                      !is.na(B_Corsi_total) &
                                        !is.na(Settingtype) &
                                        !is.na(SettingRegion) &
                                        !is.na(SettingID))$E_Corsi_total,
                        main="Histogram of Corsi Block subtest at endline (analytical sample)",
                        xlab = "Corsi Block scores",
                        xlim = c(0, 15),
                        breaks = seq(0, 15, 3),
                        xaxp = c(0, 15, 5))

##No real floor or ceiling effects.
```

```
Hmisc::describe(dplyr::filter(data,
                        !is.na(E_Corsi_total))$E_Corsi_total)
sd(dplyr::filter(data,
              !is.na(E_Corsi_total))$E_Corsi_total)


##Analytical sample
Hmisc::describe(dplyr::filter(data,
                        !is.na(B_Corsi_total) &
                          !is.na(Settingtype) &
                          !is.na(SettingRegion) &
                          !is.na(E_Corsi_total) &
                          !is.na(SettingID))$E_Corsi_total)
sd(dplyr::filter(data,
              !is.na(B_Corsi_total) &
                !is.na(Settingtype) &
                !is.na(SettingRegion) &
                !is.na(E_Corsi_total) &
                !is.na(SettingID))$E_Corsi_total)


#And by treatment and control groups:
Hmisc::describe(dplyr::filter(data,  #treatment
                        treatment == 1 &
                          !is.na(B_Corsi_total) &
                          !is.na(Settingtype) &
                          !is.na(SettingRegion) &
                          !is.na(E_Corsi_total) &
                          !is.na(SettingID))$E_Corsi_total)

sd(dplyr::filter(data, #treatment
              treatment == 1 &
                !is.na(B_Corsi_total) &
                !is.na(Settingtype) &
                !is.na(SettingRegion) &
                !is.na(E_Corsi_total) &
                !is.na(SettingID))$E_Corsi_total)

Hmisc::describe(dplyr::filter(data,  #control
                        treatment == 0 &
                          !is.na(B_Corsi_total) &
                          !is.na(Settingtype) &
                          !is.na(SettingRegion) &
                          !is.na(E_Corsi_total) &
                          !is.na(SettingID))$E_Corsi_total)
sd(dplyr::filter(data, #control
              treatment == 0 &
                !is.na(B_Corsi_total) &
                !is.na(Settingtype) &
                !is.na(SettingRegion) &
                !is.na(E_Corsi_total) &
                !is.na(SettingID))$E_Corsi_total)

#####################################################
```

```
##################################################
##################################################

# Section B: Secondary outcome model 1

### 1. Run the multi-level model.

sec3_model <- lmer(E_Corsi_total ~
                     treatment +
                     B_Corsi_total  +
                     SettingRegion +
                     Settingtype +
                     (1 | SettingID),
                   data=data, REML = FALSE)

summary(sec3_model) #Produce the results

performance::icc(sec3_model) #ICC

#########

### 2. Testing OLS assumptions

#Residual diagnostics: Testing normality of residuals OLS assumption
resid_sec3 <- resid(sec3_model) #Create object which stores the residuals of the model
plot_resid_sec3 <- plot(density(resid_sec3)) #Kernel density plot to explore normality
shapiro.test(resid_sec3) #Shapiro-wilk test for normality. !FLAG potentially K-S test better with larger sample sizes!
(p<0.05 rejects H0 that data is normally distributed)
ks.test(resid_sec3, "pnorm") #K-S test (p<0.05 rejects H0 that data is normally distributed)

##Both of these tests reject H0

# QQ line
qq_line3 <- ggplot(data = data.frame(resid = resid_sec3), aes(sample = resid)) +
  stat_qq() +
  stat_qq_line() +
  labs(title = "Q-Q Plot of Residuals", x = "Theoretical Quantiles", y = "Sample Quantiles") +
  theme_minimal()


#Plotting residual density against a normal distribution for comparison. NB this is using the residual objects created
from the secondary model above
sec3_resid_df <- as.data.frame(resid_sec3)
sd_resid_sec3 <- sd(resid_sec3)
max_resid_sec3 <- max(resid_sec3)
min_resid_sec3 <- min(resid_sec3)

resid_sec3_kd<- ggplot(sec3_resid_df, aes(x = resid_sec3)) +
  geom_density(aes(color = "Kernel density"), linewidth = 1) +
  stat_function(aes(color = "Normal density"),
                fun = function(x) dnorm(x, mean = 0, sd = sd_resid_sec3),
                linetype = "dotted", linewidth = 1) +
  theme_minimal() +
  scale_x_continuous(limits = c(min_resid_sec3, max_resid_sec3)) +
```

```
  scale_color_manual(name = "Legend", values = c("Kernel density" = "#000080", "Normal density" = "red")) +
  labs(
    title = "Executive Function Model 3 - Residuals Density Plot",
    x = "Residuals",
    y = "Density"
  )

#Testing linearity and assumption that residual errors have a mean of 0
plot(sec3_model, col = "red") #We want the line here to be horizontal and at 0

### Figures to export:
resid_sec3_kd
qq_line3

plots <- list(resid_sec3_kd, qq_line3)

### Export graphs to word

# Create a Word document
doc <- read_docx()

# Save each plot as an image and add to the Word document
for (i in seq_along(plots)) {
  # Export the plot to a temporary file
  file_name <- tempfile(fileext = ".png")
  ggplot2::ggsave(filename = file_name, plot = plots[[i]])

  # Add the plot to the Word document
  doc <- doc %>%
    body_add_par(value = paste("Plot", i), style = "heading 1") %>%
    body_add_img(src = file_name, width = 6, height = 4)
}

# Save the document
print(doc, target = "//Nevis/RE/Projects/In_Progress/022807.014_P2899_EEF_The_ONE/WIP/09. Impact/Graphs &
Summary Tables/secondary_outcome_model3.docx")

### 3.  Bootstrapping CIs and p-value
#+ The plots and normality test results suggest
#+ that the residuals are not normally distributed. We will
#+ therefore need to re-estimate CIs and p-values using bootstrapping.

set.seed(999)
boot_sec3_model <- boot.pval::boot_summary(sec3_model,
                                  type = "norm",
                                  method = NULL,
                                  conf.level = 0.95)
boot_sec3_model

######################
```

```
### 4.  Effect-size estimation

coefs3 <- data.frame(summary(sec3_model)$coefficients) #create data frame of all coefficients from secondary outcome model

c <- coefs3["treatment", "Estimate"] #Extract the treatment coefficient

n <- nrow(dplyr::filter(data,
                        treatment == 1 &
                          !is.na(E_Corsi_total) &
                          !is.na(B_Corsi_total) &
                          !is.na(SettingRegion) &
                          !is.na(Settingtype) &
                          !is.na(SettingID))) #Number of individuals in treatment group in the model

m <- nrow(dplyr::filter(data,
                        treatment == 0 &
                          !is.na(E_Corsi_total) &
                          !is.na(B_Corsi_total) &
                          !is.na(SettingRegion) &
                          !is.na(Settingtype) &
                          !is.na(SettingID))) #Number of individuals in control group in the model

v <- var(dplyr::filter(data,
                       treatment == 1 &
                         !is.na(B_Corsi_total) &
                         !is.na(SettingRegion) &
                         !is.na(Settingtype) &
                         !is.na(E_Corsi_total) &
                         !is.na(SettingID))$E_Corsi_total) #Variance in outcome among treatment group

w <- var(dplyr::filter(data,
                       treatment == 0 &
                         !is.na(B_Corsi_total) &
                         !is.na(SettingRegion) &
                         !is.na(Settingtype) &
                         !is.na(E_Corsi_total) &
                         !is.na(SettingID))$E_Corsi_total)  #Variance in outcome among control group

secondary3_effect_size <- hedges.g(c, n, m, v, w)
secondary3_effect_size

#LOWER CI EFFECT SIZE using lower 95% CI from the bootstrapped model:

#We redefine c to equal the treatment coefficient minus the SE created above
c <- boot_sec3_model$Lower.bound[2]

sec3_effect_size_low <- hedges.g(c, n, m, v, w) #Re-run hedges G on lower CI treatment coefficient
sec3_effect_size_low #This is the lower confidence interval for the treatment effect on the following directions subtest.



#UPPER CI EFFECT SIZE using bootrapped 95% upper CI from the above model:
```

```
#We redefine c again to equal the treatment coefficient plus the SE created above. Remember c has already been defined to be the lower confidence interval, we therefore need to add 2 SEs to it.
c <- boot_sec3_model$Upper.bound[2]

sec3_effect_size_high <- hedges.g(c, n, m, v, w) #Re-run hedges G on upper CI treatment coefficient
sec3_effect_size_high #This is the upper confidence interval for the treatment effect on the following directions subtest.

## P-value
p_values <- boot_sec3_model$p.value[2]
p_value <- round(p_values, digits = 2)


### 5. Creating secondary analysis table (to paste output into report)

##+ This part of the code is to create a formatted output table which can just be copied
##+ into the table empty table in word

## Find missing numbers for model (number of obs with outcome but missing covariates)

# Treatment
n_mis <- nrow(dplyr::filter(data,
                            treatment == 1 &
                              !is.na(E_Corsi_total) &
                              (is.na(B_Corsi_total) |
                                 is.na(SettingRegion) |
                                 is.na(Settingtype) |
                                 is.na(SettingID))))

# Control
m_mis <- nrow(dplyr::filter(data,
                            treatment == 0 &
                              !is.na(E_Corsi_total) &
                              (is.na(B_Corsi_total) |
                                 is.na(SettingRegion) |
                                 is.na(Settingtype) |
                                 is.na(SettingID))))

# Paste in numbers (non-missing and missing) for table
num_t <- paste0(as.character(n), " (", as.character(n_mis), ")")
num_c <- paste0(as.character(m), " (", as.character(m_mis), ")")

# Total numbers
total_n <- n + m
total_n_t_c <- paste0(as.character(total_n), " (", n, ";", m, ")")

## Generate means and CIs

# Treatment
me_t <- t.test(dplyr::filter(data,
                             treatment == 1 &
                               !is.na(E_Corsi_total) &
                               !is.na(B_Corsi_total) &
                               !is.na(SettingRegion) &
```

```
                    !is.na(Settingtype) &
                    !is.na(SettingID))$E_Corsi_total) #Run t-test on outcome in analytical sample

mean_t <- as.character(round(me_t$estimate, digits = 2)) #extract mean from t-test object (me), to 2dp and as character to facilitate pasting in CIs.

me_ci_t <- as.character(round(me_t$conf.int[c(1, 2)], digits = 2)) #extract CIs (95%) from t-test object (me), to 2dp and as character to facilitate pasting

mean_ci_t <- paste0(mean_t, " (", me_ci_t[1], " - ", me_ci_t[2], ")") #Paste CIs in brackets next to mean figure

# Control
me_c <- t.test(dplyr::filter(data,
                      treatment == 0 &
                        !is.na(E_Corsi_total) &
                        !is.na(B_Corsi_total) &
                        !is.na(SettingRegion) &
                        !is.na(Settingtype) &
                        !is.na(SettingID))$E_Corsi_total)

mean_c <- as.character(round(me_c$estimate, digits = 2))
me_ci_c <- as.character(round(me_c$conf.int[c(1, 2)], digits = 2))
mean_ci_c <- paste0(mean_c, " (", me_ci_c[1], " - ", me_ci_c[2], ")")


## Effect size with CIs
ES_CI <- paste0(as.character(round(secondary3_effect_size, digits = 3)),
          " (", as.character(round(sec3_effect_size_low, digits = 2)),
          " -- ",
          as.character(round(sec3_effect_size_high, digits = 2)),
          ")")

## Create table

secondary_analysis_table_3 <- t(as.data.frame(c(num_t, mean_ci_t, num_c, mean_ci_c, total_n_t_c, ES_CI,
p_value)))
colnames(secondary_analysis_table_3) <- c("N (intervention)", "Mean (intervention", "N (control)", "Mean (control)",
"Total (T;C)", "Hedges g (95% CIs)", "p-value")

rownames(secondary_analysis_table_3) <- "HTKS Executive Function Test"

## Display results - TABLE 1
View(secondary_analysis_table_3)

############

### 6. Creating effect size estimation table. Most of the required fields have already been created

u_m <- as.character(round((mean(dplyr::filter(data, #Unadjusted difference in means
                                  treatment == 1 &
                                    !is.na(B_Corsi_total) &
                                    !is.na(SettingRegion) &
                                    !is.na(Settingtype) &
                                    !is.na(SettingID))$E_Corsi_total, na.rm = TRUE)) - (mean(dplyr::filter(data,
```

```
                                                                    treatment == 0 &
                                                                      !is.na(B_Corsi_total) &
                                                                      !is.na(SettingRegion) &
                                                                      !is.na(Settingtype) &
                                                                      !is.na(SettingID))$E_Corsi_total, na.rm =
TRUE)), digits = 2))


c <- round(coefs3["treatment", "Estimate"], digits = 2) #Adjusted difference in means

var_t <- round(v, digits = 2) #Variance of outcome in treatment

var_c <- round(w, digits = 2) #Variance of outcome in control

pooled_var <- round(((((n - 1)*v)+((m - 1)* w))/(n + m - 2)), digits = 2)


# Create table 2
sec3_es_est <- t(as.data.frame(c(u_m, c, num_t, var_t, num_c, var_c, pooled_var)))
colnames(sec3_es_est) <- c("Unadj. diff in means", "Adj diff in means", "N (treatment)", "Variance (treatment)", "N
(control)", "Variance (control)", "Pooled Variance")
rownames(sec3_es_est) <- "Corsi Blocks Executive Function Test"
View(sec3_es_est)

##################################################

##################################################
############ Romano-Wolf Correction #############
##################################################

complete_secondary_data <- as.data.frame(dplyr::filter(data,
                                !is.na(E_Corsi_total),
                                !is.na(E_HTKS_total),
                                !is.na(B_Corsi_total),
                                !is.na(B_HTKS_total)))

##Encoding Text Vars into dummies
complete_secondary_data <- complete_secondary_data %>%
mutate(IsPVI = ifelse(Settingtype=="PVI", 1, 0))

complete_secondary_data$SettingRegion <- revalue(complete_secondary_data$SettingRegion,
                                      c("East Midlands" = "EMids",
                                        "London" = "L",
                                        "East of England"  = "E",
                                        "Yorkshire and Humber" = "YH"))

complete_secondary_data <- complete_secondary_data %>%
  mutate(IsEast = ifelse(SettingRegion=="E", 1, 0))

complete_secondary_data <- complete_secondary_data %>%
  mutate(IsLondon = ifelse(SettingRegion=="L", 1, 0))

complete_secondary_data <- complete_secondary_data %>%
  mutate(IsYH = ifelse(SettingRegion=="YH", 1, 0))
```

```
##SettingID needs to be converted to an integer
complete_secondary_data$SettingID <- substring(complete_secondary_data$SettingID, 2, 4)
complete_secondary_data$SettingID <- as.double(complete_secondary_data$SettingID)

complete_secondary_data$treatment <- as.double(complete_secondary_data$treatment)

##Mixed_measures model on complete subset
sec1_2 <- lmer(E_HTKS_total ~
                treatment +
                B_Corsi_total +
                IsEast +
                IsLondon+
                IsYH+
                IsPVI +
                (1 | SettingID),
              data=complete_secondary_data, REML = FALSE)

##HTKS model complete data subset
sec2_2 <- lmer(E_HTKS_total ~
                treatment +
                B_HTKS_total +
                IsEast +
                IsLondon+
                IsYH+
                IsPVI +
                (1 | SettingID),
              data=complete_secondary_data, REML = FALSE)

##Corsi block model complete data subset
sec3_2 <- lmer(E_Corsi_total ~
                treatment +
                B_Corsi_total +
                IsEast +
                IsLondon+
                IsYH+
                IsPVI +
                (1 | SettingID),
              data=complete_secondary_data, REML = FALSE)


results <- stepdown(
  list(sec1_2, sec2_2, sec3_2),
  tr_var = "treatment",
  cl_var = "SettingID",
  data = complete_secondary_data,
  alpha = 0.05,
  plots = FALSE,
  n_permute = 1000,
  nsteps = 1000,
  type = "rw",
  confint = TRUE,
  verbose = TRUE
)
print(results)
```

```
#####################  The ONE Analysis: Compliance Analysis ##################

## [PURPOSE OF SCRIPT]:
###+ Conducting the compliance analysis for the Primary outcome of the ONE. We want to show to what degree
compliance
###+ (or lack of) contributed to the null effect found in the primary analysis.
###+
###+ We have three measures of compliance
###+    Setting-level Compliance: Attendance at professional development sessions
###+    Setting-level Dosage: Intervention activities offered to the children
###+    Child-level Dosage: Attendance patterns at setting
###+      - This will have to be created from setting reported data
###+
###+ This will involve a descriptive summary of compliance followed by a 2SLS model of the effect of compliance on
###+ early numeracy ability

rm(list = ls())
set.seed(999)

library("haven")
library("dplyr")
library("ggplot2")
library("lme4")
library("sjstats")
library("Hmisc")
library("pastecs")
library("performance")
library("lmtest")
library("eeptools")
library("dplyr")
library("lmerTest")
library("officer")
library("magrittr")

data <- read_dta("//Nevis/RE/Projects/In_Progress/022807.014_P2899_EEF_The_ONE/WIP/06. Data collection/02.
Endline/02. Data_NotMirrored/BaselineEndlineComplianceClean2.dta")
data <- filter(data, endlineonly != 1)

#########################################################################################################
#################

##Create binary compliance measure for child level dosage
##From E_CurrentAttendancePatterns
data <- data %>%
  mutate(ChildDosageBinary = ifelse(E_CurrentAttendancePatterns >= 15, 1, 0))

#########################################################################################################
#################

##Ensure compliance is 0 for control settings
#Setting compliance
data$`_Practitioner_Consistently_Pres` <- ifelse(data$treatment == 0, 0, data$`_Practitioner_Consistently_Pres`)
table(data$`_Practitioner_Consistently_Pres`, data$treatment)
```

```
##966 children in compliant settings. None in control settings

#Setting Dosage
data$Activity_Adherence_Binary <- ifelse(data$treatment == 0, 0, data$Activity_Adherence_Binary)
table(data$Activity_Adherence_Binary, data$treatment)
##226 children in compliant settings. None in control settings.

#For the continuous dosage variable
data$Activity_Adherence <- ifelse(data$treatment == 0, 0, data$Activity_Adherence)

#Child Dosage
data$ChildDosageBinary <- ifelse(data$treatment == 0, 0, data$ChildDosageBinary)
table(data$ChildDosageBinary, data$treatment)
##860 children in compliant settings. None in control settings

#For the continuous child-level dosage variable. create a new variable
data <- data %>%
  mutate(ChildDosage = ChildDosageBinary * E_CurrentAttendancePatterns)

##############################################################################################
#################

### Descriptive tables of compliance within treatment group
#Setting-level compliance (_Practitioner_Consistently_Pres)
table(dplyr::filter(data,
                    treatment == 1)$`_Practitioner_Consistently_Pres`)

# Setting-level Dosage (Activity_Adherence_Binary)
table(dplyr::filter(data,
                    treatment == 1)$Activity_Adherence_Binary)

# Summary of number of activities in the treatment group
describe(dplyr::filter(data,
                    treatment == 1)$Activity_Adherence)

hist(dplyr::filter(data,
                   treatment == 1)$Activity_Adherence,
     main = "Total Activities",
     xlab = "Activities",
     col = "#440154FF")

# Child-level Dosage (ChildDosageBinary)
table(dplyr::filter(data,
                    treatment == 1)$ChildDosageBinary)

describe(dplyr::filter(data,
                    treatment == 1)$ChildDosage)

##############################################################################################
#################

#############################################################
################### Compliance Analysis ###################
#############################################################
```

```
### 2SLS (CACE) compliance analysis using setting-level compliance analysis binary

compl_data <- dplyr::filter(data,
                            !is.na(`_Practitioner_Consistently_Pres`))

### First Stage Regression

complianceFS <- lmer(`_Practitioner_Consistently_Pres` ~
                       treatment +
                       SettingRegion +
                       Settingtype +
                       (1 | SettingID),
                     data=compl_data, REML = FALSE)

summary(complianceFS)

# Predicted compliance- ignoring random effects
compl_data$predicted_compliance <- predict(complianceFS, compl_data, re.form = ~(1|SettingID),
                                           allow.new.levels = T)

describe(compl_data$predicted_compliance)

## Second Stage Regression
complianceSS <- lmer(E_score ~
                       predicted_compliance +
                       B_score +
                       SettingRegion +
                       Settingtype +
                       (1 | SettingID),
                     data=compl_data, REML = FALSE)

summary(complianceSS)

##+ Binary compliance measure does not significantly affect early numeracy.
##+ Given most settings complied by this measure, and that the primary analysis model also finds a null effect
##+ this isn't particularly surprising.

##############################################################################################
################

##Testing OLS assumptions.

#Residual diagnostics: Testing normality of residuals OLS assumption
resid_comp <- resid(complianceSS) #Create object which stores the residuals of the model
plot_resid_comp <- plot(density(resid_comp)) #Kernel density plot to explore normality
shapiro.test(resid_comp) #Shapiro-wilk test for normality. !FLAG potentially K-S test better with larger sample sizes!
(p<0.05 rejects H0 that data is normally distributed)
ks.test(resid_comp, "pnorm") #K-S test (p<0.05 rejects H0 that data is normally distributed)

##Like in the primary analysis the residuals look very normally distributed but both of the tests reject normality
##So will need to bootstrap confidence intervals

# QQ line
```

```
qq_line <- ggplot(data = data.frame(resid = resid_comp), aes(sample = resid)) +
  stat_qq() +
  stat_qq_line() +
  labs(title = "Q-Q Plot of Residuals", x = "Theoretical Quantiles", y = "Sample Quantiles") +
  theme_minimal()

# Deviation from normality at extreme quantiles

#Plotting residual density against a normal distribution for comparison. NB this is using the residual objects created
from the primary model above
comp_resid_df <- as.data.frame(resid_comp)
sd_resid <- sd(resid_comp)
max_resid <- max(resid_comp)
min_resid <- min(resid_comp)

resid_comp_kd <- ggplot(comp_resid_df, aes(x = resid_comp)) +
  geom_density(aes(color = "Kernel density"), linewidth = 1) +
  stat_function(aes(color = "Normal density"),
                fun = function(x) dnorm(x, mean = 0, sd = sd_resid),
                linetype = "dotted", linewidth = 1) +
  theme_minimal() +
  scale_x_continuous(limits = c(min_resid, max_resid)) +
  scale_color_manual(name = "Legend", values = c("Kernel density" = "#000080", "Normal density" = "red")) +
  labs(
    title = "Early Numeracy - Residuals Density Plot",
    x = "Residuals",
    y = "Density"
  )

#Testing linearity and assumption that residual errors have a mean of 0
plot(complianceSS, col = "red") #We want the line here to be horizontal and at 0

plots <- list(resid_comp_kd, qq_line)

### Export graphs to word

# Create a Word document
doc <- read_docx()

# Save each plot as an image and add to the Word document
for (i in seq_along(plots)) {
  # Export the plot to a temporary file
  file_name <- tempfile(fileext = ".png")
  ggplot2::ggsave(filename = file_name, plot = plots[[i]])

  # Add the plot to the Word document
  doc <- doc %>%
    body_add_par(value = paste("Plot", i), style = "heading 1") %>%
    body_add_img(src = file_name, width = 6, height = 4)
}

# Save the document
print(doc, target = "//Nevis/RE/Projects/In_Progress/022807.014_P2899_EEF_The_ONE/WIP/09. Impact/Graphs &
Summary Tables/compliance_analysis_model.docx")
```

```
#########################################################################################################
#################

## Bootstrapping CIs and p-values

boot_compSS <- boot.pval::boot_summary(complianceSS,
                                       type = "norm",
                                       method = NULL,
                                       conf.level = 0.95)
boot_compSS

#########################################################################################################
#################

## Calculating effect sizes

hedges.g <- function(c, n, m, v, w) {
  c / sqrt((((n - 1) * v) + ((m - 1) * w)) / (n + m - 2))
}

# Extract coefficients and calculate effect size
coefs <- data.frame(summary(complianceSS)$coefficients)

c <- coefs["predicted_compliance", "Estimate"] # Extract the compliance coefficient

# For saving the number of obs and variance across T/C groups, we need to make sure this is among obs not missing
compliance information
n <- nrow(dplyr::filter(compl_data,
                        treatment== 1 &
                          !is.na(E_score) &
                          !is.na(B_score) &
                          !is.na(SettingRegion) &
                          !is.na(Settingtype) &
                          !is.na(SettingID)))

m <- nrow(dplyr::filter(compl_data,
                        treatment == 0 &
                          !is.na(E_score) &
                          !is.na(B_score) &
                          !is.na(SettingRegion) &
                          !is.na(Settingtype) &
                          !is.na(SettingID)))

 v <- var(dplyr::filter(compl_data,
                        treatment == 1 &
                          !is.na(E_score) &
                          !is.na(B_score) &
                          !is.na(SettingRegion) &
                          !is.na(Settingtype) &
                          !is.na(SettingID))$E_score)

 w <- var(dplyr::filter(compl_data,
                        treatment == 0 &
```

```
                !is.na(E_score) &
                !is.na(B_score) &
                !is.na(SettingRegion) &
                !is.na(Settingtype) &
                !is.na(SettingID))$E_score)

comp_effect_size <- hedges.g(c, n, m, v, w)
comp_effect_size

#LOWER CI EFFECT SIZE using lower 95% CI from the bootstrapped model:
c <- boot_compSS$Lower.bound[2]
comp_effect_size_low <- hedges.g(c, n, m, v, w)

# Upper 95% CI
c <- boot_compSS$Upper.bound[2]# Extract the compliance coefficient
comp_effect_size_high <- hedges.g(c, n, m, v, w)

# P-values
p_values <- boot_compSS$p.value[2]
p_value <- round(p_values, digits = 2)


#########################################################################################################
################

### Creating Analysis Output Tables

## Missingness

# Treatment
n_mis <- nrow(dplyr::filter(compl_data,
                  treatment == 1 &
                    !is.na(E_score) &
                    (is.na(B_score) |
                       is.na(SettingRegion) |
                       is.na(Settingtype) |
                       is.na(SettingID))))

# Control
m_mis <- nrow(dplyr::filter(compl_data,
                  treatment == 0 &
                    !is.na(E_score) &
                    (is.na(B_score) |
                       is.na(SettingRegion) |
                       is.na(Settingtype) |
                       is.na(SettingID))))

# Paste in numbers (non-missing and missing) for table
num_t <- paste0(as.character(n), " (", as.character(n_mis), ")")
num_c <- paste0(as.character(m), " (", as.character(m_mis), ")")

# Total numbers
total_n <- n + m
total_n_t_c <- paste0(as.character(total_n), " (", n, ";", m, ")")
```

```
## Generate means and CIs

# Treatment
me_t <- t.test(dplyr::filter(compl_data,
                        treatment == 1 &
                          !is.na(E_score) &
                          !is.na(B_score) &
                          !is.na(SettingRegion) &
                          !is.na(Settingtype) &
                          !is.na(SettingID))$E_score) #Run t-test on outcome in analytical sample

mean_t <- as.character(round(me_t$estimate, digits = 2)) #extract mean from t-test object (me), to 2dp and as character to facilitate pasting in CIs.

me_ci_t <- as.character(round(me_t$conf.int[c(1, 2)], digits = 2)) #extract CIs (95%) from t-test object (me), to 2dp and as character to facilitate pasting

mean_ci_t <- paste0(mean_t, " (", me_ci_t[1], " - ", me_ci_t[2], ")") #Paste CIs in brackets next to mean figure

# Control
me_c <- t.test(dplyr::filter(compl_data,
                        treatment == 0 &
                          !is.na(E_score) &
                          !is.na(B_score) &
                          !is.na(SettingRegion) &
                          !is.na(Settingtype) &
                          !is.na(SettingID))$E_score)

mean_c <- as.character(round(me_c$estimate, digits = 2))
me_ci_c <- as.character(round(me_c$conf.int[c(1, 2)], digits = 2))
mean_ci_c <- paste0(mean_c, " (", me_ci_c[1], " - ", me_ci_c[2], ")")


## Effect size with CIs
ES_CI <- paste0(as.character(round(comp_effect_size, digits = 2)),
            " (", as.character(round(comp_effect_size_low, digits = 2)),
            " - ",
            as.character(round(comp_effect_size_high, digits = 2)),
            ")")

## Create table

comp_analysis_table <- as.data.frame(t(c(num_t, mean_ci_t, num_c, mean_ci_c, total_n_t_c, ES_CI, p_value)))
colnames(comp_analysis_table) <- c("N (intervention)", "Mean (intervention", "N (control)", "Mean (control)", "Total (T;C)", "Hedges g (95% CIs)", "p-value")

rownames(comp_analysis_table) <- "EYTN Numeracy (Compliance)"

## Display results
View(comp_analysis_table)

##################
```

```
### Effect size estimation table
u_m <- as.character(round((mean(dplyr::filter(compl_data, #Unadjusted difference in means- within the sample with non-missing compliance data
                                              treatment == 1 &
                                                !is.na(B_score) &
                                                !is.na(SettingRegion) &
                                                !is.na(Settingtype) &
                                                !is.na(SettingID))$E_score, na.rm = TRUE)) - (mean(dplyr::filter(compl_data,
                                                                                                                treatment == 0 &
                                                                                                                  !is.na(B_score) &
                                                                                                                  !is.na(SettingRegion) &
                                                                                                                  !is.na(Settingtype) &
                                                                                                                  !is.na(SettingID))$E_score, na.rm =
TRUE)), digits = 2))

c <- round(coefs["predicted_compliance", "Estimate"], digits = 2) #Adjusted difference in means

var_t <- round(v, digits = 2) #Variance of outcome in treatment

var_c <- round(w, digits = 2) #Variance of outcome in control

pooled_var <- round(((((n - 1)*v)+((m - 1)* w))/(n + m - 2)), digits = 2) #pooled variance


# Create table
comp_es_est <- t(as.data.frame(c(u_m, c, num_t, var_t, num_c, var_c, pooled_var)))
colnames(comp_es_est) <- c("Unadj. diff in means", "Adj diff in means", "N (treatment)", "Variance (treatment)", "N (control)", "Variance (control)", "Pooled Variance")
rownames(comp_es_est) <- "EYTN Numeracy (compliance)"
View(comp_es_est)


#################################################################
############## Setting-Level Dosage Analysis ##############
#################################################################

### 2SLS (CACE) dosage analysis using setting-level activity adherence
### This is the measure with the lowest compliance so if we find an effect anywhere I imagine it will be here

dos1_data <- dplyr::filter(data,
                           !is.na(Activity_Adherence))

### First Stage Regression

dos1FS <- lmer(Activity_Adherence ~
                 treatment +
                 SettingRegion +
                 Settingtype +
                 (1 | SettingID),
               data=dos1_data, REML = FALSE)

summary(dos1FS)

# Predicted compliance- ignoring random effects
```

```
dos1_data$predicted_dosage <- predict(dos1FS, dos1_data, re.form = ~(1|SettingID),
                                      allow.new.levels = T)

describe(dos1_data$predicted_dosage)

## Second Stage Regression
dos1SS <- lmer(E_score ~
                 predicted_dosage +
                 B_score +
                 SettingRegion +
                 Settingtype +
                 (1 | SettingID),
               data=dos1_data, REML = FALSE)

summary(dos1SS)

##+ Predicted dosage measure does not significantly affect early numeracy.

#########################################################################################################
################

##Testing OLS assumptions.

#Residual diagnostics: Testing normality of residuals OLS assumption
resid_dos1 <- resid(dos1SS) #Create object which stores the residuals of the model
plot_resid_dos1 <- plot(density(resid_dos1)) #Kernel density plot to explore normality
shapiro.test(resid_dos1) #Shapiro-wilk test for normality. !FLAG potentially K-S test better with larger sample sizes!
(p<0.05 rejects H0 that data is normally distributed)
ks.test(resid_dos1, "pnorm") #K-S test (p<0.05 rejects H0 that data is normally distributed)

##Residuals not normally distributed so will need to bootstrap confidence intervals

# QQ line
qq_line <- ggplot(data = data.frame(resid = resid_dos1), aes(sample = resid)) +
  stat_qq() +
  stat_qq_line() +
  labs(title = "Q-Q Plot of Residuals", x = "Theoretical Quantiles", y = "Sample Quantiles") +
  theme_minimal()

# Deviation from normality at extreme quantiles

#Plotting residual density against a normal distribution for comparison. NB this is using the residual objects created
from the primary model above
dos1_resid_df <- as.data.frame(resid_dos1)
sd_resid <- sd(resid_dos1)
max_resid <- max(resid_dos1)
min_resid <- min(resid_dos1)

resid_dos1_kd <- ggplot(dos1_resid_df, aes(x = resid_dos1)) +
  geom_density(aes(color = "Kernel density"), linewidth = 1) +
  stat_function(aes(color = "Normal density"),
                fun = function(x) dnorm(x, mean = 0, sd = sd_resid),
                linetype = "dotted", linewidth = 1) +
  theme_minimal() +
```

```
  scale_x_continuous(limits = c(min_resid, max_resid)) +
  scale_color_manual(name = "Legend", values = c("Kernel density" = "#000080", "Normal density" = "red")) +
  labs(
    title = "Early Numeracy - Residuals Density Plot",
    x = "Residuals",
    y = "Density"
  )

#Testing linearity and assumption that residual errors have a mean of 0
plot(dos1SS, col = "red") #We want the line here to be horizontal and at 0

plots <- list(resid_dos1_kd, qq_line)

### Export graphs to word

# Create a Word document
doc <- read_docx()

# Save each plot as an image and add to the Word document
for (i in seq_along(plots)) {
  # Export the plot to a temporary file
  file_name <- tempfile(fileext = ".png")
  ggplot2::ggsave(filename = file_name, plot = plots[[i]])

  # Add the plot to the Word document
  doc <- doc %>%
    body_add_par(value = paste("Plot", i), style = "heading 1") %>%
    body_add_img(src = file_name, width = 6, height = 4)
}

# Save the document
print(doc, target = "//Nevis/RE/Projects/In_Progress/022807.014_P2899_EEF_The_ONE/WIP/09. Impact/Graphs &
Summary Tables/dosage_analysis_model1.docx")

##########################################################################################################
#################

## Bootstrapping CIs and p-values

boot_dos1SS <- boot.pval::boot_summary(dos1SS,
                                       type = "norm",
                                       method = NULL,
                                       conf.level = 0.95)
boot_dos1SS

##########################################################################################################
#################

## Calculating effect sizes
# Extract coefficients and calculate effect size
coefs <- data.frame(summary(dos1SS)$coefficients)

c <- coefs["predicted_dosage", "Estimate"] # Extract the compliance coefficient
```

```
# For saving the number of obs and variance across T/C groups, we need to make sure this is among obs not missing
compliance information
n <- nrow(dplyr::filter(dos1_data,
                        treatment== 1 &
                          !is.na(E_score) &
                          !is.na(B_score) &
                          !is.na(SettingRegion) &
                          !is.na(Settingtype) &
                          !is.na(SettingID)))

m <- nrow(dplyr::filter(dos1_data,
                        treatment == 0 &
                          !is.na(E_score) &
                          !is.na(B_score) &
                          !is.na(SettingRegion) &
                          !is.na(Settingtype) &
                          !is.na(SettingID)))

 v <- var(dplyr::filter(dos1_data,
                        treatment == 1 &
                         !is.na(E_score) &
                         !is.na(B_score) &
                         !is.na(SettingRegion) &
                         !is.na(Settingtype) &
                         !is.na(SettingID))$E_score)

 w <- var(dplyr::filter(dos1_data,
                        treatment == 0 &
                         !is.na(E_score) &
                         !is.na(B_score) &
                         !is.na(SettingRegion) &
                         !is.na(Settingtype) &
                         !is.na(SettingID))$E_score)

dos1_effect_size <- hedges.g(c, n, m, v, w)
dos1_effect_size

#LOWER CI EFFECT SIZE using lower 95% CI from the bootstrapped model:
c <- boot_dos1SS$Lower.bound[2]
dos1_effect_size_low <- hedges.g(c, n, m, v, w)

# Upper 95% CI
c <- boot_dos1SS$Upper.bound[2]# Extract the compliance coefficient
dos1_effect_size_high <- hedges.g(c, n, m, v, w)

# P-values
p_values <- boot_dos1SS$p.value[2]
p_value <- round(p_values, digits = 2)


##############################################################################################
################

### Creating Analysis Output Tables
```

```
## Missingess

# Treatment
n_mis <- nrow(dplyr::filter(dos1_data,
                            treatment == 1 &
                              !is.na(E_score) &
                              (is.na(B_score) |
                                 is.na(SettingRegion) |
                                 is.na(Settingtype) |
                                 is.na(SettingID))))

# Control
m_mis <- nrow(dplyr::filter(dos1_data,
                            treatment == 0 &
                              !is.na(E_score) &
                              (is.na(B_score) |
                                 is.na(SettingRegion) |
                                 is.na(Settingtype) |
                                 is.na(SettingID))))

# Paste in numbers (non-missing and missing) for table
num_t <- paste0(as.character(n), " (", as.character(n_mis), ")")
num_c <- paste0(as.character(m), " (", as.character(m_mis), ")")

# Total numbers
total_n <- n + m
total_n_t_c <- paste0(as.character(total_n), " (", n, ";", m, ")")

## Generate means and CIs

# Treatment
me_t <- t.test(dplyr::filter(dos1_data,
                             treatment == 1 &
                               !is.na(E_score) &
                               !is.na(B_score) &
                               !is.na(SettingRegion) &
                               !is.na(Settingtype) &
                               !is.na(SettingID))$E_score) #Run t-test on outcome in analytical sample

mean_t <- as.character(round(me_t$estimate, digits = 2)) #extract mean from t-test object (me), to 2dp and as
character to facilitate pasting in CIs.

me_ci_t <- as.character(round(me_t$conf.int[c(1, 2)], digits = 2)) #extract CIs (95%) from t-test object (me), to 2dp and
as character to facilitate pasting

mean_ci_t <- paste0(mean_t, " (", me_ci_t[1], " - ", me_ci_t[2], ")") #Paste CIs in brackets next to mean figure

# Control
me_c <- t.test(dplyr::filter(dos1_data,
                             treatment == 0 &
                               !is.na(E_score) &
                               !is.na(B_score) &
                               !is.na(SettingRegion) &
```

```
                    !is.na(Settingtype) &
                    !is.na(SettingID))$E_score)

mean_c <- as.character(round(me_c$estimate, digits = 2))
me_ci_c <- as.character(round(me_c$conf.int[c(1, 2)], digits = 2))
mean_ci_c <- paste0(mean_c, " (", me_ci_c[1], " - ", me_ci_c[2], ")")


## Effect size with CIs
ES_CI <- paste0(as.character(round(dos1_effect_size, digits = 4)),
          " (", as.character(round(dos1_effect_size_low, digits = 4)),
          " - ",
          as.character(round(dos1_effect_size_high, digits = 4)),
          ")")

## Create table

dos1_analysis_table <- as.data.frame(t(c(num_t, mean_ci_t, num_c, mean_ci_c, total_n_t_c, ES_CI, p_value)))
colnames(dos1_analysis_table) <- c("N (intervention)", "Mean (intervention", "N (control)", "Mean (control)", "Total
(T;C)", "Hedges g (95% CIs)", "p-value")

rownames(dos1_analysis_table) <- "EYTN Numeracy (Compliance)"

## Display results
View(dos1_analysis_table)

##################

### Effect size estimation table
u_m <- as.character(round((mean(dplyr::filter(dos1_data, #Unadjusted difference in means- within the sample with
non-missing compliance data
                                treatment == 1 &
                                  !is.na(B_score) &
                                  !is.na(SettingRegion) &
                                  !is.na(Settingtype) &
                                  !is.na(SettingID))$E_score, na.rm = TRUE)) - (mean(dplyr::filter(dos1_data,
                                                                                  treatment == 0 &
                                                                                    !is.na(B_score) &
                                                                                    !is.na(SettingRegion) &
                                                                                    !is.na(Settingtype) &
                                                                                    !is.na(SettingID))$E_score, na.rm = TRUE)),
digits = 2))

c <- round(coefs["predicted_dosage", "Estimate"], digits = 3) #Adjusted difference in means

var_t <- round(v, digits = 2) #Variance of outcome in treatment

var_c <- round(w, digits = 2) #Variance of outcome in control

pooled_var <- round(((((n - 1)*v)+((m - 1)* w))/(n + m - 2)), digits = 2) #pooled variance


# Create table
dos1_es_est <- t(as.data.frame(c(u_m, c, num_t, var_t, num_c, var_c, pooled_var)))
```

```
colnames(dos1_es_est) <- c("Unadj. diff in means", "Adj diff in means", "N (treatment)", "Variance (treatment)", "N (control)", "Variance (control)", "Pooled Variance")
rownames(dos1_es_est) <- "EYTN Numeracy (compliance)"
View(dos1_es_est)

############################################################
############## Child-Level Dosage Analysis ##############
############################################################

### 2SLS (CACE) dosage analysis using child-level attendance

dos2_data <- dplyr::filter(data,
                           !is.na(ChildDosage))

### First Stage Regression

dos2FS <- lmer(ChildDosage ~
                 treatment +
                 B_score +
                 SettingRegion +
                 Settingtype +
                 (1 | SettingID),
               data=dos2_data, REML = FALSE)

summary(dos2FS)

# Predicted compliance- ignoring random effects
dos2_data$predicted_dosage <- predict(dos2FS, dos2_data, re.form = ~(1|SettingID),
                                      allow.new.levels = T)

describe(dos2_data$predicted_dosage)
## NB some predicted values are negative

## Second Stage Regression
dos2SS <- lmer(E_score ~
                 predicted_dosage +
                 B_score +
                 SettingRegion +
                 Settingtype +
                 (1 | SettingID),
               data=dos2_data, REML = FALSE)

summary(dos2SS)

##+ Predicted dosage measure does not significantly affect early numeracy.

############################################################################################################
#################

##Testing OLS assumptions.

#Residual diagnostics: Testing normality of residuals OLS assumption
resid_dos2 <- resid(dos2SS) #Create object which stores the residuals of the model
plot_resid_dos2 <- plot(density(resid_dos2)) #Kernel density plot to explore normality
```

```
shapiro.test(resid_dos2) #Shapiro-wilk test for normality. !FLAG potentially K-S test better with larger sample sizes! (p<0.05 rejects H0 that data is normally distributed)
ks.test(resid_dos2, "pnorm") #K-S test (p<0.05 rejects H0 that data is normally distributed)

##Residuals not normally distributed so will need to bootstrap confidence intervals

# QQ line
qq_line <- ggplot(data = data.frame(resid = resid_dos2), aes(sample = resid)) +
  stat_qq() +
  stat_qq_line() +
  labs(title = "Q-Q Plot of Residuals", x = "Theoretical Quantiles", y = "Sample Quantiles") +
  theme_minimal()

# Deviation from normality at extreme quantiles

#Plotting residual density against a normal distribution for comparison. NB this is using the residual objects created from the primary model above
dos2_resid_df <- as.data.frame(resid_dos2)
sd_resid <- sd(resid_dos2)
max_resid <- max(resid_dos2)
min_resid <- min(resid_dos2)

resid_dos2_kd <- ggplot(dos2_resid_df, aes(x = resid_dos2)) +
  geom_density(aes(color = "Kernel density"), linewidth = 1) +
  stat_function(aes(color = "Normal density"),
                fun = function(x) dnorm(x, mean = 0, sd = sd_resid),
                linetype = "dotted", linewidth = 1) +
  theme_minimal() +
  scale_x_continuous(limits = c(min_resid, max_resid)) +
  scale_color_manual(name = "Legend", values = c("Kernel density" = "#000080", "Normal density" = "red")) +
  labs(
    title = "Early Numeracy - Residuals Density Plot",
    x = "Residuals",
    y = "Density"
  )

#Testing linearity and assumption that residual errors have a mean of 0
plot(dos2SS, col = "red") #We want the line here to be horizontal and at 0

plots <- list(resid_dos2_kd, qq_line)

### Export graphs to word

# Create a Word document
doc <- read_docx()

# Save each plot as an image and add to the Word document
for (i in seq_along(plots)) {
  # Export the plot to a temporary file
  file_name <- tempfile(fileext = ".png")
  ggplot2::ggsave(filename = file_name, plot = plots[[i]])

  # Add the plot to the Word document
  doc <- doc %>%
```

```
  body_add_par(value = paste("Plot", i), style = "heading 1") %>%
  body_add_img(src = file_name, width = 6, height = 4)
}

# Save the document
print(doc, target = "//Nevis/RE/Projects/In_Progress/022807.014_P2899_EEF_The_ONE/WIP/09. Impact/Graphs &
Summary Tables/dosage_analysis_model2.docx")

##########################################################################################################
#################

## Bootstrapping CIs and p-values

boot_dos2SS <- boot.pval::boot_summary(dos2SS,
                                       type = "norm",
                                       method = NULL,
                                       conf.level = 0.95)
boot_dos2SS

##########################################################################################################
#################

## Calculating effect sizes
# Extract coefficients and calculate effect size
coefs <- data.frame(summary(dos2SS)$coefficients)

c <- coefs["predicted_dosage", "Estimate"] # Extract the compliance coefficient

# For saving the number of obs and variance across T/C groups, we need to make sure this is among obs not missing
compliance information
n <- nrow(dplyr::filter(dos2_data,
                        treatment== 1 &
                          !is.na(E_score) &
                          !is.na(B_score) &
                          !is.na(SettingRegion) &
                          !is.na(Settingtype) &
                          !is.na(SettingID)))

m <- nrow(dplyr::filter(dos2_data,
                        treatment == 0 &
                          !is.na(E_score) &
                          !is.na(B_score) &
                          !is.na(SettingRegion) &
                          !is.na(Settingtype) &
                          !is.na(SettingID)))

 v <- var(dplyr::filter(dos2_data,
                        treatment == 1 &
                         !is.na(E_score) &
                         !is.na(B_score) &
                         !is.na(SettingRegion) &
                         !is.na(Settingtype) &
                         !is.na(SettingID))$E_score)
```

```
 w <- var(dplyr::filter(dos2_data,
                        treatment == 0 &
                          !is.na(E_score) &
                          !is.na(B_score) &
                          !is.na(SettingRegion) &
                          !is.na(Settingtype) &
                          !is.na(SettingID))$E_score)

dos2_effect_size <- hedges.g(c, n, m, v, w)
dos2_effect_size

#LOWER CI EFFECT SIZE using lower 95% CI from the bootstrapped model:
c <- boot_dos2SS$Lower.bound[2]
dos2_effect_size_low <- hedges.g(c, n, m, v, w)

# Upper 95% CI
c <- boot_dos2SS$Upper.bound[2]# Extract the compliance coefficient
dos2_effect_size_high <- hedges.g(c, n, m, v, w)

# P-values
p_values <- boot_dos2SS$p.value[2]
p_value <- round(p_values, digits = 2)


##############################################################################################
################

### Creating Analysis Output Tables

## Missingness

# Treatment
n_mis <- nrow(dplyr::filter(dos2_data,
                            treatment == 1 &
                              !is.na(E_score) &
                              (is.na(B_score) |
                                 is.na(SettingRegion) |
                                 is.na(Settingtype) |
                                 is.na(SettingID))))

# Control
m_mis <- nrow(dplyr::filter(dos2_data,
                            treatment == 0 &
                              !is.na(E_score) &
                              (is.na(B_score) |
                                 is.na(SettingRegion) |
                                 is.na(Settingtype) |
                                 is.na(SettingID))))

# Paste in numbers (non-missing and missing) for table
num_t <- paste0(as.character(n), " (", as.character(n_mis), ")")
num_c <- paste0(as.character(m), " (", as.character(m_mis), ")")

# Total numbers
```

```
total_n <- n + m
total_n_t_c <- paste0(as.character(total_n), " (", n, ";", m, ")")

## Generate means and CIs

# Treatment
me_t <- t.test(dplyr::filter(dos2_data,
                             treatment == 1 &
                               !is.na(E_score) &
                               !is.na(B_score) &
                               !is.na(SettingRegion) &
                               !is.na(Settingtype) &
                               !is.na(SettingID))$E_score) #Run t-test on outcome in analytical sample

mean_t <- as.character(round(me_t$estimate, digits = 2)) #extract mean from t-test object (me), to 2dp and as character to facilitate pasting in CIs.

me_ci_t <- as.character(round(me_t$conf.int[c(1, 2)], digits = 2)) #extract CIs (95%) from t-test object (me), to 2dp and as character to facilitate pasting

mean_ci_t <- paste0(mean_t, " (", me_ci_t[1], " - ", me_ci_t[2], ")") #Paste CIs in brackets next to mean figure

# Control
me_c <- t.test(dplyr::filter(dos2_data,
                             treatment == 0 &
                               !is.na(E_score) &
                               !is.na(B_score) &
                               !is.na(SettingRegion) &
                               !is.na(Settingtype) &
                               !is.na(SettingID))$E_score)

mean_c <- as.character(round(me_c$estimate, digits = 2))
me_ci_c <- as.character(round(me_c$conf.int[c(1, 2)], digits = 2))
mean_ci_c <- paste0(mean_c, " (", me_ci_c[1], " - ", me_ci_c[2], ")")


## Effect size with CIs
ES_CI <- paste0(as.character(round(dos2_effect_size, digits = 4)),
                " (", as.character(round(dos2_effect_size_low, digits = 4)),
                " - ",
                as.character(round(dos2_effect_size_high, digits = 4)),
                ")")

## Create table

dos2_analysis_table <- as.data.frame(t(c(num_t, mean_ci_t, num_c, mean_ci_c, total_n_t_c, ES_CI, p_value)))
colnames(dos2_analysis_table) <- c("N (intervention)", "Mean (intervention", "N (control)", "Mean (control)", "Total (T;C)", "Hedges g (95% CIs)", "p-value")

rownames(dos2_analysis_table) <- "EYTN Numeracy (Compliance)"

## Display results
View(dos2_analysis_table)
```

```
##################

### Effect size estimation table
u_m <- as.character(round((mean(dplyr::filter(dos2_data, #Unadjusted difference in means- within the sample with non-missing compliance data
                                              treatment == 1 &
                                                !is.na(B_score) &
                                                !is.na(SettingRegion) &
                                                !is.na(Settingtype) &
                                                !is.na(SettingID))$E_score, na.rm = TRUE)) - (mean(dplyr::filter(dos2_data,
                                                                                                                treatment == 0 &
                                                                                                                  !is.na(B_score) &
                                                                                                                  !is.na(SettingRegion) &
                                                                                                                  !is.na(Settingtype) &
                                                                                                                  !is.na(SettingID))$E_score, na.rm = TRUE)),
digits = 2))

c <- round(coefs["predicted_dosage", "Estimate"], digits = 3) #Adjusted difference in means

var_t <- round(v, digits = 2) #Variance of outcome in treatment

var_c <- round(w, digits = 2) #Variance of outcome in control

pooled_var <- round(((((n - 1)*v)+((m - 1)* w))/(n + m - 2)), digits = 2) #pooled variance


# Create table
dos2_es_est <- t(as.data.frame(c(u_m, c, num_t, var_t, num_c, var_c, pooled_var)))
colnames(dos2_es_est) <- c("Unadj. diff in means", "Adj diff in means", "N (treatment)", "Variance (treatment)", "N (control)", "Variance (control)", "Pooled Variance")
rownames(dos2_es_est) <- "EYTN Numeracy (compliance)"
View(dos2_es_est)

#####################  The ONE Analysis: Sensitivity Analysis ##################
########################### Excluding Unblinded Settings ########################

## [PURPOSE OF SCRIPT]:

####+  This script conducts a sensitivity analysis for the ONE.
####+  We are exploring the impact of the intervention on 1 primary outcome:
####+    1. Results of EYTN test for early numeracy excluding the data from tests for the two unblinded settings:
####+       a) The Avenue
####+       b) Healdswood

####+    For this outcome, we need to create a dataset excluding these two settings
####+    The remainder of the analysis follows the same procedure as the primary data ana;usis

#############################

# Clear workspace
rm(list=ls())
set.seed(999) #NB this is the same seed set in the randomisation code at the start of the project so we shall retain it here for consistency.
```

```
# Load packages
library("haven")
library("dplyr")
library("ggplot2")
library("lme4")
library("sjstats")
library("Hmisc")
library("performance")
library("lmtest")
library("eeptools")
library("lmerTest")
library("officer")
library("boot")
library("boot.pval")


# Load cleaned data
data <- read_dta("//Nevis/RE/Projects/In_Progress/022807.014_P2899_EEF_The_ONE/WIP/06. Data collection/02.
Endline/02. Data_NotMirrored/BaselineEndlineComplianceClean2.dta")
data <- filter(data, endlineonly != 1)

#########################################################

# Section A: filtering out settings

data <- dplyr::filter(data,
                      SettingName != "Healdswood infant school" &
                      SettingName != "The Avenue Nursery")

#########################################################

# Section B: Descriptive statistics


#1. Histograms of outcomes

#Use this to check distribution of endline and baseline scores for our outcome

#We are mainly checking for ceiling and floor effects

# In overall data
EYTN_ovr_hist <- hist(data$E_score,
                      main="Histogram of EYTN subtest at endline (overall sample)",
                      xlab = "Early Numeracy Scores",
                      xlim = c(0, 120),
                      breaks = seq(0, 120, 10),
                      xaxp = c(0, 120, 6))

#+In analytical data (the analytical data is defined as the data used in the multi-level regression model.
#+It includes the observations where we have complete data for all covariates included in the outcome model)
EYTN_model_hist <- hist(dplyr::filter(data,
                                      !is.na(B_score) &
                                        !is.na(Settingtype) &
                                        !is.na(SettingRegion) &
```

```
                    !is.na(SettingID))$E_score,
     main="Histogram of EYTN subtest at endline (analytical sample)",
     xlab = "Early numeracy scores",
     xlim = c(0, 120),
     breaks = seq(0, 120, 10),
     xaxp = c(0, 120, 6))


## No floor or ceiling effects in endline test distribution
## Now for baseline
EYTN_ovr_histB <- hist(data$B_score,
                       main="Histogram of EYTN subtest at baseline (overall sample)",
                       xlab = "Early Numeracy Scores",
                       xlim = c(0, 120),
                       breaks = seq(0, 120, 10),
                       xaxp = c(0, 120, 6))

#+In analytical data (the analytical data is defined as the data used in the multi-level regression model.
#+It includes the observations where we have complete data for all covariates included in the outcome model)
EYTN_model_histB <- hist(dplyr::filter(data,
                                       !is.na(B_score) &
                                         !is.na(Settingtype) &
                                         !is.na(SettingRegion) &
                                         !is.na(SettingID))$B_score,
                         main="Histogram of EYTN subtest at baseline (analytical sample)",
                         xlab = "Early numeracy scores",
                         xlim = c(0, 120),
                         breaks = seq(0, 120, 10),
                         xaxp = c(0, 120, 6))

#+ We're unlikely to include any additional sensitivity analysis for this as any floor effects are only observed at
baseline.
#+ However, will include additional analysis here to see what proportion of the observations are within one standard
deviation of the floor.

sd <- sd(data$B_score, na.rm = T)

# Create df which is a copy of analytical data
data_sd <- dplyr::filter(data,
                         !is.na(B_score) &
                           !is.na(Settingtype) &
                           !is.na(SettingRegion) &
                           !is.na(SettingID))

data_sd <- data_sd %>%
  mutate(sd_dummy = 0)

data_sd$sd_dummy <- ifelse(data_sd$B_score < sd, 1, 0)

mean(data_sd$sd_dummy)

##Here we see that 30.82% of observations are within one sd of the floor

#2. Means, SDs, Min and Max (for whole sample and analytical sample)
```

```
#Overall sample
Hmisc::describe(dplyr::filter(data,
                        !is.na(E_score))$E_score)
sd(dplyr::filter(data,
            !is.na(E_score))$E_score)


##Analytical sample
Hmisc::describe(dplyr::filter(data,
                        !is.na(B_score) &
                          !is.na(Settingtype) &
                          !is.na(SettingRegion) &
                          !is.na(E_score) &
                          !is.na(SettingID))$E_score)
sd(dplyr::filter(data,
            !is.na(B_score) &
              !is.na(Settingtype) &
              !is.na(SettingRegion) &
              !is.na(E_score) &
              !is.na(SettingID))$E_score)


#And by treatment and control groups:
Hmisc::describe(dplyr::filter(data,  #treatment
                        treatment == 1 &
                          !is.na(B_score) &
                          !is.na(Settingtype) &
                          !is.na(SettingRegion) &
                          !is.na(E_score) &
                          !is.na(SettingID))$E_score)

sd(dplyr::filter(data, #treatment
            treatment == 1 &
              !is.na(B_score) &
              !is.na(Settingtype) &
              !is.na(SettingRegion) &
              !is.na(E_score) &
              !is.na(SettingID))$E_score)

Hmisc::describe(dplyr::filter(data,  #control
                        treatment == 0 &
                          !is.na(B_score) &
                          !is.na(Settingtype) &
                          !is.na(SettingRegion) &
                          !is.na(E_score) &
                          !is.na(SettingID))$E_score)
sd(dplyr::filter(data, #control
            treatment == 0 &
              !is.na(B_score) &
              !is.na(Settingtype) &
              !is.na(SettingRegion) &
              !is.na(E_score) &
              !is.na(SettingID))$E_score)
```

```
################################################
################################################
################################################



# Section C: Primary outcome model

### 1. Run the multi-level model.

pr_model <- lmer(E_score ~
                   treatment +
                   B_score +
                   SettingRegion +
                   Settingtype +
                   (1 | SettingID),
                 data=data, REML = FALSE)

summary(pr_model) #Produce the results

performance::icc(pr_model) #ICC

#########

### 2. Testing OLS assumptions

#Residual diagnostics: Testing normality of residuals OLS assumption
resid_pr <- resid(pr_model) #Create object which stores the residuals of the model
plot_resid_pr <- plot(density(resid_pr)) #Kernel density plot to explore normality
shapiro.test(resid_pr) #Shapiro-wilk test for normality. !FLAG potentially K-S test better with larger sample sizes!
(p<0.05 rejects H0 that data is normally distributed)
ks.test(resid_pr, "pnorm") #K-S test (p<0.05 rejects H0 that data is normally distributed

##Both of these tests reject H0

# QQ line
qq_line <- ggplot(data = data.frame(resid = resid_pr), aes(sample = resid)) +
  stat_qq() +
  stat_qq_line() +
  labs(title = "Q-Q Plot of Residuals", x = "Theoretical Quantiles", y = "Sample Quantiles") +
  theme_minimal()


#Plotting residual density against a normal distribution for comparison. NB this is using the residual objects created
from the primary model above
pr_resid_df <- as.data.frame(resid_pr)
sd_resid <- sd(resid_pr)
max_resid <- max(resid_pr)
min_resid <- min(resid_pr)

resid_pr_kd<- ggplot(pr_resid_df, aes(x = resid_pr)) +
  geom_density(aes(color = "Kernel density"), linewidth = 1) +
  stat_function(aes(color = "Normal density"),
```

```
            fun = function(x) dnorm(x, mean = 0, sd = sd_resid),
            linetype = "dotted", linewidth = 1) +
  theme_minimal() +
  scale_x_continuous(limits = c(min_resid, max_resid)) +
  scale_color_manual(name = "Legend", values = c("Kernel density" = "#000080", "Normal density" = "red")) +
  labs(
    title = "Basic Concepts - Residuals Density Plot",
    x = "Residuals",
    y = "Density"
  )

#Testing linearity and assumption that residual errors have a mean of 0
plot(pr_model, col = "red") #We want the line here to be horizontal and at 0


### Figures to export:
resid_pr_kd
qq_line

plots <- list(resid_pr_kd, qq_line)

### Export graphs to word

# Create a Word document
doc <- read_docx()

# Save each plot as an image and add to the Word document
for (i in seq_along(plots)) {
  # Export the plot to a temporary file
  file_name <- tempfile(fileext = ".png")
  ggplot2::ggsave(filename = file_name, plot = plots[[i]])

  # Add the plot to the Word document
  doc <- doc %>%
    body_add_par(value = paste("Plot", i), style = "heading 1") %>%
    body_add_img(src = file_name, width = 6, height = 4)
}

# Save the document
print(doc, target = "//Nevis/RE/Projects/In_Progress/022807.014_P2899_EEF_The_ONE/WIP/09. Impact/Graphs &
Summary Tables/Sensitivity1.docx")

##########


### 3.  Bootstrapping CIs and p-value
#+ The QQ residual plot, Shapiro-Wilk, and K-S test
#+ results suggest that the residuals are not normally distributed. We will
#+ therefore need to re-estimate CIs and p-values using bootstrapping.

set.seed(999)
boot_pr_model <- boot.pval::boot_summary(pr_model,
                                         type = "norm",
                                         method = NULL,
```

```
                                conf.level = 0.95)
boot_pr_model


######################



### 4.  Effect-size estimation

#Create Hedges g function.
hedges.g <- function(c, n, m, v, w){
  c/sqrt((((n - 1)*v)+((m - 1)* w))/(n + m-2))
}



#We use the output from the model above to calculate the effect size: (NB: This needs to be done for confidence intervals too!)
#We need to define the objects needed to be inputted into the Hedges G function created above (c, n, m, v, & w)

coefs <- data.frame(summary(pr_model)$coefficients) #create data frame of all coefficients from primary outcome model

c <- coefs["treatment", "Estimate"] #Extract the treatment coefficient

n <- nrow(dplyr::filter(data,
                        treatment == 1 &
                          !is.na(E_score) &
                          !is.na(B_score) &
                          !is.na(SettingRegion) &
                          !is.na(Settingtype) &
                          !is.na(SettingID))) #Number of individuals in treatment group in the model

m <- nrow(dplyr::filter(data,
                        treatment == 0 &
                          !is.na(E_score) &
                          !is.na(B_score) &
                          !is.na(SettingRegion) &
                          !is.na(Settingtype) &
                          !is.na(SettingID))) #Number of individuals in control group in the model

v <- var(dplyr::filter(data,
                       treatment == 1 &
                         !is.na(B_score) &
                         !is.na(SettingRegion) &
                         !is.na(Settingtype) &
                         !is.na(E_score) &
                         !is.na(SettingID))$E_score) #Variance in outcome among treatment group

w <- var(dplyr::filter(data,
                       treatment == 0 &
```

```
                !is.na(B_score) &
                !is.na(SettingRegion) &
                !is.na(Settingtype) &
                !is.na(E_score) &
                !is.na(SettingID))$E_score)  #Variance in outcome among control group

primary_effect_size <- hedges.g(c, n, m, v, w)
primary_effect_size

#LOWER CI EFFECT SIZE using lower 95% CI from the bootstrapped model:

#We redefine c to equal the treatment coefficient minus the SE created above
c <- boot_pr_model$Lower.bound[2]

pr_effect_size_low <- hedges.g(c, n, m, v, w) #Re-run hedges G on lower CI treatment coefficient
pr_effect_size_low #This is the lower confidence interval for the treatment effect on the following directions subtest.



#UPPER CI EFFECT SIZE using bootrapped 95% upper CI from the above model:

#We redefine c again to equal the treatment coefficient plus the SE created above. Remember c has already been
defined to be the lower confidence interval, we therefore need to add 2 SEs to it.
c <- boot_pr_model$Upper.bound[2]

pr_effect_size_high <- hedges.g(c, n, m, v, w) #Re-run hedges G on upper CI treatment coefficient
pr_effect_size_high #This is the upper confidence interval for the treatment effect on the following directions subtest.

## P-value
p_values <- boot_pr_model$p.value[2]
p_value <- round(p_values, digits = 2)


### 5. Creating primary analysis table (to paste output into report)

##+ This part of the code is to create a formatted output table which can just be copied
##+ into the table empty table in word

## Find missing numbers for model (number of obs with outcome but missing covariates)

# Treatment
n_mis <- nrow(dplyr::filter(data,
                  treatment == 1 &
                    !is.na(E_score) &
                    (is.na(B_score) |
                       is.na(SettingRegion) |
                       is.na(Settingtype) |
                       is.na(SettingID))))

# Control
m_mis <- nrow(dplyr::filter(data,
                  treatment == 0 &
                    !is.na(E_score) &
                    (is.na(B_score) |
```

```
                    is.na(SettingRegion) |
                    is.na(Settingtype) |
                    is.na(SettingID))))

# Paste in numbers (non-missing and missing) for table
num_t <- paste0(as.character(n), " (", as.character(n_mis), ")")
num_c <- paste0(as.character(m), " (", as.character(m_mis), ")")

# Total numbers
total_n <- n + m
total_n_t_c <- paste0(as.character(total_n), " (", n, ";", m, ")")

## Generate means and CIs

# Treatment
me_t <- t.test(dplyr::filter(data,
                    treatment == 1 &
                      !is.na(E_score) &
                      !is.na(B_score) &
                      !is.na(SettingRegion) &
                      !is.na(Settingtype) &
                      !is.na(SettingID))$E_score) #Run t-test on outcome in analytical sample

mean_t <- as.character(round(me_t$estimate, digits = 3)) #extract mean from t-test object (me), to 2dp and as
character to facilitate pasting in CIs.

me_ci_t <- as.character(round(me_t$conf.int[c(1, 2)], digits = 2)) #extract CIs (95%) from t-test object (me), to 2dp and
as character to facilitate pasting

mean_ci_t <- paste0(mean_t, " (", me_ci_t[1], " - ", me_ci_t[2], ")") #Paste CIs in brackets next to mean figure

# Control
me_c <- t.test(dplyr::filter(data,
                    treatment == 0 &
                      !is.na(E_score) &
                      !is.na(B_score) &
                      !is.na(SettingRegion) &
                      !is.na(Settingtype) &
                      !is.na(SettingID))$E_score)

mean_c <- as.character(round(me_c$estimate, digits = 3))
me_ci_c <- as.character(round(me_c$conf.int[c(1, 2)], digits = 2))
mean_ci_c <- paste0(mean_c, " (", me_ci_c[1], " - ", me_ci_c[2], ")")


## Effect size with CIs
ES_CI <- paste0(as.character(round(primary_effect_size, digits = 3)),
            " (", as.character(round(pr_effect_size_low, digits = 2)),
            " -- ",
            as.character(round(pr_effect_size_high, digits = 2)),
            ")")

## Create table
```

```
primary_analysis_table <- t(as.data.frame(c(num_t, mean_ci_t, num_c, mean_ci_c, total_n_t_c, ES_CI, p_value)))
colnames(primary_analysis_table) <- c("N (intervention)", "Mean (intervention", "N (control)", "Mean (control)", "Total (T;C)", "Hedges g (95% CIs)", "p-value")

rownames(primary_analysis_table) <- "EYTN2 Numeracy"

## Display results - TABLE 1
View(primary_analysis_table)

###########

### 6. Creating effect size estimation table. Most of the required fields have already been created

u_m <- as.character(round((mean(dplyr::filter(data, #Unadjusted difference in means
                                    treatment == 1 &
                                      !is.na(B_score) &
                                      !is.na(SettingRegion) &
                                      !is.na(Settingtype) &
                                      !is.na(SettingID))$E_score, na.rm = TRUE)) - (mean(dplyr::filter(data,
                                                                                      treatment == 0 &
                                                                                        !is.na(B_score) &
                                                                                        !is.na(SettingRegion) &
                                                                                        !is.na(Settingtype) &
                                                                                        !is.na(SettingID))$E_score, na.rm = TRUE)),
digits = 3))

c <- round(coefs["treatment", "Estimate"], digits = 3) #Adjusted difference in means

var_t <- round(v, digits = 2) #Variance of outcome in treatment

var_c <- round(w, digits = 2) #Variance of outcome in control

pooled_var <- round(((((n - 1)*v)+((m - 1)* w))/(n + m-2)), digits = 2) #pooled standard deviation



# Create table 2
pr_es_est <- t(as.data.frame(c(u_m, c, num_t, var_t, num_c, var_c, pooled_var)))
colnames(pr_es_est) <- c("Unadj. diff in means", "Adj diff in means", "N (treatment)", "Variance (treatment)", "N (control)", "Variance (control)", "Pooled Variance")
rownames(pr_es_est) <- "EYTN Numeracy Test"
View(pr_es_est)
#####################  The ONE Analysis: Sensitivity Analysis ##################
################################ Accounting for Age ############################

## [PURPOSE OF SCRIPT]:

####+  This script conducts the age sensitivity analysis for the ONE.
####+  We conduct analysis with the primary and mixed measure model adding the child's age at baseline to the


############################

# Clear workspace
```

```
rm(list=ls())
set.seed(999) #NB this is the same seed set in the randomisation code at the start of the project so we shall retain it here for consistency.

# Load packages
library("haven")
library("dplyr")
library("ggplot2")
library("lme4")
library("sjstats")
library("Hmisc")
library("performance")
library("lmtest")
library("eeptools")
library("lmerTest")
library("officer")
library("boot")
library("boot.pval")


# Load cleaned data
data <- read_dta("//Nevis/RE/Projects/In_Progress/022807.014_P2899_EEF_The_ONE/WIP/06. Data collection/02. Endline/02. Data_NotMirrored/BaselineEndlineComplianceClean2.dta")
data <- filter(data, endlineonly != 1)

#########################################################
################ EYTN Numeracy Test ################
#########################################################

# Section A: Descriptive statistics


#1. Histograms of outcomes

#Use this to check distribution of endline and baseline scores for our outcome

#We are mainly checking for ceiling and floor effects

# In overall data
EYTN_ovr_hist <- hist(data$E_score,
                      main="Histogram of EYTN subtest at endline (overall sample)",
                      xlab = "Early Numeracy Scores",
                      xlim = c(0, 120),
                      breaks = seq(0, 120, 10),
                      xaxp = c(0, 120, 6))

#+In analytical data (the analytical data is defined as the data used in the multi-level regression model.
#+It includes the observations where we have complete data for all covariates included in the outcome model)
EYTN_model_hist <- hist(dplyr::filter(data,
                                      !is.na(B_score) &
                                        !is.na(Settingtype) &
                                        !is.na(SettingRegion) &
                                        !is.na(SettingID) &
                                        !is.na(AgeAtBase))$E_score,
```

```
                main="Histogram of EYTN subtest at endline (analytical sample with age)",
                xlab = "Early numeracy scores",
                xlim = c(0, 120),
                breaks = seq(0, 120, 10),
                xaxp = c(0, 120, 6))


## No floor or ceiling effects in endline test distribution
## Now for baseline
EYTN_ovr_histB <- hist(data$B_score,
                       main="Histogram of EYTN subtest at baseline (overall sample)",
                       xlab = "Early Numeracy Scores",
                       xlim = c(0, 120),
                       breaks = seq(0, 120, 10),
                       xaxp = c(0, 120, 6))

#+In analytical data (the analytical data is defined as the data used in the multi-level regression model.
#+It includes the observations where we have complete data for all covariates included in the outcome model)
EYTN_model_histB <- hist(dplyr::filter(data,
                                       !is.na(B_score) &
                                         !is.na(Settingtype) &
                                         !is.na(SettingRegion) &
                                         !is.na(SettingID)&
                                         !is.na(AgeAtBase))$B_score,
                         main="Histogram of EYTN subtest at baseline (analytical sample)",
                         xlab = "Early numeracy scores",
                         xlim = c(0, 120),
                         breaks = seq(0, 120, 10),
                         xaxp = c(0, 120, 6))

#+ We're unlikely to include any additional sensitivity analysis for this as any floor effects are only observed at
baseline.
#+ However, will include additional analysis here to see what proportion of the observations are within one standard
deviation of the floor.

sd <- sd(data$B_score, na.rm = T)

# Create df which is a copy of analytical data
data_sd <- dplyr::filter(data,
                         !is.na(B_score) &
                           !is.na(Settingtype) &
                           !is.na(SettingRegion) &
                           !is.na(SettingID)&
                           !is.na(AgeAtBase))

data_sd <- data_sd %>%
  mutate(sd_dummy = 0)

data_sd$sd_dummy <- ifelse(data_sd$B_score < sd, 1, 0)

mean(data_sd$sd_dummy)

##Here we see that 30.57% of observations are within one sd of the floor
```

```
#2. Means, SDs, Min and Max (for whole sample and analytical sample)

#Overall sample
Hmisc::describe(dplyr::filter(data,
                        !is.na(E_score))$E_score)
sd(dplyr::filter(data,
            !is.na(E_score))$E_score)


##Analytical sample
Hmisc::describe(dplyr::filter(data,
                        !is.na(B_score) &
                          !is.na(Settingtype) &
                          !is.na(SettingRegion) &
                          !is.na(E_score) &
                          !is.na(SettingID) &
                          !is.na(AgeAtBase))$E_score)
sd(dplyr::filter(data,
            !is.na(B_score) &
              !is.na(Settingtype) &
              !is.na(SettingRegion) &
              !is.na(E_score) &
              !is.na(SettingID) &
              !is.na(AgeAtBase))$E_score)


#And by treatment and control groups:
Hmisc::describe(dplyr::filter(data,  #treatment
                        treatment == 1 &
                          !is.na(B_score) &
                          !is.na(Settingtype) &
                          !is.na(SettingRegion) &
                          !is.na(E_score) &
                          !is.na(SettingID) &
                          !is.na(AgeAtBase))$E_score)

sd(dplyr::filter(data, #treatment
            treatment == 1 &
              !is.na(B_score) &
              !is.na(Settingtype) &
              !is.na(SettingRegion) &
              !is.na(E_score) &
              !is.na(SettingID) &
              !is.na(AgeAtBase))$E_score)

Hmisc::describe(dplyr::filter(data,  #control
                        treatment == 0 &
                          !is.na(B_score) &
                          !is.na(Settingtype) &
                          !is.na(SettingRegion) &
                          !is.na(E_score) &
                          !is.na(SettingID) &
                          !is.na(AgeAtBase))$E_score)
sd(dplyr::filter(data, #control
```

```
            treatment == 0 &
              !is.na(B_score) &
              !is.na(Settingtype) &
              !is.na(SettingRegion) &
              !is.na(E_score) &
              !is.na(SettingID) &
              !is.na(AgeAtBase))$E_score)

##################################################
##################################################
##################################################



# Section B: Primary outcome model

### 1. Run the multi-level model.

pr_model <- lmer(E_score ~
                   treatment +
                   B_score +
                   SettingRegion +
                   Settingtype +
                   AgeAtBase +
                   (1 | SettingID),
                 data=data, REML = FALSE)

summary(pr_model) #Produce the results

performance::icc(pr_model) #ICC

#########

### 2. Testing OLS assumptions

#Residual diagnostics: Testing normality of residuals OLS assumption
resid_pr <- resid(pr_model) #Create object which stores the residuals of the model
plot_resid_pr <- plot(density(resid_pr)) #Kernel density plot to explore normality
shapiro.test(resid_pr) #Shapiro-wilk test for normality. !FLAG potentially K-S test better with larger sample sizes!
(p<0.05 rejects H0 that data is normally distributed)
ks.test(resid_pr, "pnorm") #K-S test (p<0.05 rejects H0 that data is normally distributed

##Both of these tests reject H0

# QQ line
qq_line <- ggplot(data = data.frame(resid = resid_pr), aes(sample = resid)) +
  stat_qq() +
  stat_qq_line() +
  labs(title = "Q-Q Plot of Residuals", x = "Theoretical Quantiles", y = "Sample Quantiles") +
  theme_minimal()


#Plotting residual density against a normal distribution for comparison. NB this is using the residual objects created
from the primary model above
```

```
pr_resid_df <- as.data.frame(resid_pr)
sd_resid <- sd(resid_pr)
max_resid <- max(resid_pr)
min_resid <- min(resid_pr)

resid_pr_kd<- ggplot(pr_resid_df, aes(x = resid_pr)) +
  geom_density(aes(color = "Kernel density"), linewidth = 1) +
  stat_function(aes(color = "Normal density"),
                fun = function(x) dnorm(x, mean = 0, sd = sd_resid),
                linetype = "dotted", linewidth = 1) +
  theme_minimal() +
  scale_x_continuous(limits = c(min_resid, max_resid)) +
  scale_color_manual(name = "Legend", values = c("Kernel density" = "#000080", "Normal density" = "red")) +
  labs(
    title = "EYTN Numeracy - Residuals Density Plot",
    x = "Residuals",
    y = "Density"
  )

#Testing linearity and assumption that residual errors have a mean of 0
plot(pr_model, col = "red") #We want the line here to be horizontal and at 0


### Figures to export:
resid_pr_kd
qq_line

plots <- list(resid_pr_kd, qq_line)

### Export graphs to word

# Create a Word document
doc <- read_docx()

# Save each plot as an image and add to the Word document
for (i in seq_along(plots)) {
  # Export the plot to a temporary file
  file_name <- tempfile(fileext = ".png")
  ggplot2::ggsave(filename = file_name, plot = plots[[i]])

  # Add the plot to the Word document
  doc <- doc %>%
    body_add_par(value = paste("Plot", i), style = "heading 1") %>%
    body_add_img(src = file_name, width = 6, height = 4)
}

# Save the document
print(doc, target = "//Nevis/RE/Projects/In_Progress/022807.014_P2899_EEF_The_ONE/WIP/09. Impact/Graphs &
Summary Tables/Sensitivity2.docx")

##########


### 3.  Bootstrapping CIs and p-value
```

```
#+ The QQ residual plot, Shapiro-Wilk, and K-S test
#+ results suggest that the residuals are not normally distributed. We will
#+ therefore need to re-estimate CIs and p-values using bootstrapping.

set.seed(999)
boot_pr_model <- boot.pval::boot_summary(pr_model,
                                type = "norm",
                                method = NULL,
                                conf.level = 0.95)
boot_pr_model


######################




### 4.  Effect-size estimation

#Create Hedges g function.
hedges.g <- function(c, n, m, v, w){
  c/sqrt((((n - 1)*v)+((m - 1)* w))/(n + m-2))
}



#We use the output from the model above to calculate the effect size: (NB: This needs to be done for confidence intervals too!)
#We need to define the objects needed to be inputted into the Hedges G function created above (c, n, m, v, & w)

coefs <- data.frame(summary(pr_model)$coefficients) #create data frame of all coefficients from primary outcome model

c <- coefs["treatment", "Estimate"] #Extract the treatment coefficient

n <- nrow(dplyr::filter(data,
                  treatment == 1 &
                    !is.na(E_score) &
                    !is.na(B_score) &
                    !is.na(SettingRegion) &
                    !is.na(Settingtype) &
                    !is.na(SettingID) &
                    !is.na(AgeAtBase))) #Number of individuals in treatment group in the model

m <- nrow(dplyr::filter(data,
                  treatment == 0 &
                    !is.na(E_score) &
                    !is.na(B_score) &
                    !is.na(SettingRegion) &
                    !is.na(Settingtype) &
                    !is.na(SettingID) &
                    !is.na(AgeAtBase))) #Number of individuals in control group in the model
```

```
v <- var(dplyr::filter(data,
                       treatment == 1 &
                         !is.na(B_score) &
                         !is.na(SettingRegion) &
                         !is.na(Settingtype) &
                         !is.na(E_score) &
                         !is.na(SettingID) &
                         !is.na(AgeAtBase))$E_score) #Variance in outcome among treatment group

w <- var(dplyr::filter(data,
                       treatment == 0 &
                         !is.na(B_score) &
                         !is.na(SettingRegion) &
                         !is.na(Settingtype) &
                         !is.na(E_score) &
                         !is.na(SettingID) &
                         !is.na(AgeAtBase))$E_score)  #Variance in outcome among control group

primary_effect_size <- hedges.g(c, n, m, v, w)
primary_effect_size

#LOWER CI EFFECT SIZE using lower 95% CI from the bootstrapped model:

#We redefine c to equal the treatment coefficient minus the SE created above
c <- boot_pr_model$Lower.bound[2]

pr_effect_size_low <- hedges.g(c, n, m, v, w) #Re-run hedges G on lower CI treatment coefficient
pr_effect_size_low #This is the lower confidence interval for the treatment effect on the following directions subtest.



#UPPER CI EFFECT SIZE using bootrapped 95% upper CI from the above model:

#We redefine c again to equal the treatment coefficient plus the SE created above. Remember c has already been
defined to be the lower confidence interval, we therefore need to add 2 SEs to it.
c <- boot_pr_model$Upper.bound[2]

pr_effect_size_high <- hedges.g(c, n, m, v, w) #Re-run hedges G on upper CI treatment coefficient
pr_effect_size_high #This is the upper confidence interval for the treatment effect on the following directions subtest.

## P-value
p_values <- boot_pr_model$p.value[2]
p_value <- round(p_values, digits = 2)


### 5. Creating primary analysis table (to paste output into report)

##+ This part of the code is to create a formatted output table which can just be copied
##+ into the table empty table in word

## Find missing numbers for model (number of obs with outcome but missing covariates)

# Treatment
n_mis <- nrow(dplyr::filter(data,
```

```
                        treatment == 1 &
                          !is.na(E_score) &
                          (is.na(B_score) |
                             is.na(SettingRegion) |
                             is.na(Settingtype) |
                             is.na(SettingID) |
                             is.na(AgeAtBase))))

# Control
m_mis <- nrow(dplyr::filter(data,
                        treatment == 0 &
                          !is.na(E_score) &
                          (is.na(B_score) |
                             is.na(SettingRegion) |
                             is.na(Settingtype) |
                             is.na(SettingID) |
                             is.na(AgeAtBase))))

# Paste in numbers (non-missing and missing) for table
num_t <- paste0(as.character(n), " (", as.character(n_mis), ")")
num_c <- paste0(as.character(m), " (", as.character(m_mis), ")")

# Total numbers
total_n <- n + m
total_n_t_c <- paste0(as.character(total_n), " (", n, ";", m, ")")

## Generate means and CIs

# Treatment
me_t <- t.test(dplyr::filter(data,
                         treatment == 1 &
                           !is.na(E_score) &
                           !is.na(B_score) &
                           !is.na(SettingRegion) &
                           !is.na(Settingtype) &
                           !is.na(SettingID) &
                           !is.na(AgeAtBase))$E_score) #Run t-test on outcome in analytical sample

mean_t <- as.character(round(me_t$estimate, digits = 2)) #extract mean from t-test object (me), to 2dp and as character to facilitate pasting in CIs.

me_ci_t <- as.character(round(me_t$conf.int[c(1, 2)], digits = 2)) #extract CIs (95%) from t-test object (me), to 2dp and as character to facilitate pasting

mean_ci_t <- paste0(mean_t, " (", me_ci_t[1], " - ", me_ci_t[2], ")") #Paste CIs in brackets next to mean figure

# Control
me_c <- t.test(dplyr::filter(data,
                         treatment == 0 &
                           !is.na(E_score) &
                           !is.na(B_score) &
                           !is.na(SettingRegion) &
                           !is.na(Settingtype) &
                           !is.na(SettingID) &
```

```
                    !is.na(AgeAtBase))$E_score)

mean_c <- as.character(round(me_c$estimate, digits = 2))
me_ci_c <- as.character(round(me_c$conf.int[c(1, 2)], digits = 2))
mean_ci_c <- paste0(mean_c, " (", me_ci_c[1], " - ", me_ci_c[2], ")")


## Effect size with CIs
ES_CI <- paste0(as.character(round(primary_effect_size, digits = 3)),
          " (", as.character(round(pr_effect_size_low, digits = 2)),
          " -- ",
          as.character(round(pr_effect_size_high, digits = 2)),
          ")")

## Create table

primary_analysis_table <- t(as.data.frame(c(num_t, mean_ci_t, num_c, mean_ci_c, total_n_t_c, ES_CI, p_value)))
colnames(primary_analysis_table) <- c("N (intervention)", "Mean (intervention", "N (control)", "Mean (control)", "Total
(T;C)", "Hedges g (95% CIs)", "p-value")

rownames(primary_analysis_table) <- "EYTN2 Numeracy"

## Display results - TABLE 1
View(primary_analysis_table)

###########

### 6. Creating effect size estimation table. Most of the required fields have already been created

u_m <- as.character(round((mean(dplyr::filter(data, #Unadjusted difference in means
                                treatment == 1 &
                                  !is.na(B_score) &
                                  !is.na(SettingRegion) &
                                  !is.na(Settingtype) &
                                  !is.na(SettingID) &
                                  !is.na(AgeAtBase))$E_score, na.rm = TRUE)) - (mean(dplyr::filter(data,
                                                                                   treatment == 0 &
                                                                                     !is.na(B_score) &
                                                                                     !is.na(SettingRegion) &
                                                                                     !is.na(Settingtype) &
                                                                                     !is.na(SettingID) &
                                                                                     !is.na(AgeAtBase))$E_score, na.rm =
TRUE)), digits = 2))

c <- round(coefs["treatment", "Estimate"], digits = 2) #Adjusted difference in means

var_t <- round(v, digits = 2) #Variance of outcome in treatment

var_c <- round(w, digits = 2) #Variance of outcome in control

pooled_var <- round((((n - 1)*v)+((m - 1)* w))/(n + m-2), digits = 2) #pooled variance
```

```
# Create table 2
pr_es_est <- t(as.data.frame(c(u_m, c, num_t, var_t, num_c, var_c, pooled_var)))
colnames(pr_es_est) <- c("Unadj. diff in means", "Adj diff in means", "N (treatment)", "Variance (treatment)", "N
(control)", "Variance (control)", "Pooled Variance")
rownames(pr_es_est) <- "EYTN Numeracy Test"
View(pr_es_est)


#####################################################
############ Mixed Secondary Measures #############
#####################################################

# Section A: Descriptive statistics

#1. Histograms of outcomes

#Use this to check distribution of endline scores for our outcome

#We are mainly checking for ceiling effects, as flagged at baseline stage.

# In overall data
Sec1_ovr_hist <- hist(data$E_HTKS_total,
                      main="Histogram of HTKS subtest at endline (overall sample)",
                      xlab = "HTKS Scores",
                      xlim = c(0, 120),
                      breaks = seq(0, 120, 10),
                      xaxp = c(0, 120, 6))

#+In analytical data (the analytical data is defined as the data used in the multi-level regression model.
#+It includes the observations where we have complete data for all covariates included in the outcome model)
Sec1_model_hist <- hist(dplyr::filter(data,
                                      !is.na(B_Corsi_total) &
                                        !is.na(Settingtype) &
                                        !is.na(SettingRegion) &
                                        !is.na(SettingID) &
                                        !is.na(AgeAtBase))$E_HTKS_total,
                        main="Histogram of EYTN subtest at endline (analytical sample; accounting for age)",
                        xlab = "Early numeracy scores",
                        xlim = c(0, 120),
                        breaks = seq(0, 120, 10),
                        xaxp = c(0, 120, 6))

##No real floor or ceiling effects.
baseline_hist1 <- hist(data$B_Corsi_total,
                       main="Histogram of Corsi Blocks subtest at endline (overall sample)",
                       xlab = "Corsi Blocks Scores",
                       xlim = c(0, 15),
                       breaks = seq(0, 15, 3),
                       xaxp = c(0, 15, 5))

#+Some potential floor effects in the baseline control. Does warrant additional sensistivity analysis, but some
descriptive stats below.

sd <- sd(data$B_Corsi_total, na.rm = T)
```

```
# Create df which is a copy of analytical data
data_sd <- dplyr::filter(data,
                         !is.na(B_Corsi_total) &
                           !is.na(Settingtype) &
                           !is.na(SettingRegion) &
                           !is.na(SettingID) &
                           !is.na(AgeAtBase))

data_sd <- data_sd %>%
  mutate(sd_dummy = 0)

data_sd$sd_dummy <- ifelse(data_sd$B_Corsi_total < sd, 1, 0)

mean(data_sd$sd_dummy)

# 40.07% within a standard deviation of the floor.

Hmisc::describe(dplyr::filter(data,
                              !is.na(E_HTKS_total))$E_HTKS_total)
sd(dplyr::filter(data,
                 !is.na(E_HTKS_total))$E_HTKS_total)


##Analytical sample
Hmisc::describe(dplyr::filter(data,
                              !is.na(B_Corsi_total) &
                                !is.na(Settingtype) &
                                !is.na(SettingRegion) &
                                !is.na(E_HTKS_total) &
                                !is.na(SettingID) &
                                !is.na(AgeAtBase))$E_HTKS_total)
sd(dplyr::filter(data,
                 !is.na(B_Corsi_total) &
                   !is.na(Settingtype) &
                   !is.na(SettingRegion) &
                   !is.na(E_HTKS_total) &
                   !is.na(SettingID) &
                   !is.na(AgeAtBase))$E_HTKS_total)


#And by treatment and control groups:
Hmisc::describe(dplyr::filter(data,  #treatment
                              treatment == 1 &
                                !is.na(B_Corsi_total) &
                                !is.na(Settingtype) &
                                !is.na(SettingRegion) &
                                !is.na(E_HTKS_total) &
                                !is.na(SettingID) &
                                !is.na(AgeAtBase))$E_HTKS_total)

sd(dplyr::filter(data, #treatment
                 treatment == 1 &
                   !is.na(B_Corsi_total) &
```

```
                !is.na(Settingtype) &
                !is.na(SettingRegion) &
                !is.na(E_HTKS_total) &
                !is.na(SettingID) &
                !is.na(AgeAtBase))$E_HTKS_total)

Hmisc::describe(dplyr::filter(data,  #control
                        treatment == 0 &
                          !is.na(B_Corsi_total)  &
                          !is.na(Settingtype) &
                          !is.na(SettingRegion) &
                          !is.na(E_HTKS_total) &
                          !is.na(SettingID) &
                          !is.na(AgeAtBase))$E_HTKS_total)
sd(dplyr::filter(data, #control
              treatment == 0 &
                !is.na(B_Corsi_total) &
                !is.na(Settingtype) &
                !is.na(SettingRegion) &
                !is.na(E_HTKS_total) &
                !is.na(SettingID) &
                !is.na(AgeAtBase))$E_HTKS_total)

##################################################
##################################################
##################################################

# Section B: Secondary outcome model 1

### 1. Run the multi-level model.

sec1_model <- lmer(E_HTKS_total ~
                     treatment +
                     B_Corsi_total +
                     SettingRegion +
                     Settingtype +
                     AgeAtBase +
                     (1 | SettingID),
                   data=data, REML = FALSE)

summary(sec1_model) #Produce the results

performance::icc(sec1_model) #ICC

#########

### 2. Testing OLS assumptions

#Residual diagnostics: Testing normality of residuals OLS assumption
resid_pr <- resid(sec1_model) #Create object which stores the residuals of the model
plot_resid_pr <- plot(density(resid_pr)) #Kernel density plot to explore normality
shapiro.test(resid_pr) #Shapiro-wilk test for normality. !FLAG potentially K-S test better with larger sample sizes!
(p<0.05 rejects H0 that data is normally distributed)
ks.test(resid_pr, "pnorm") #K-S test (p<0.05 rejects H0 that data is normally distributed)
```

```
##Both of these tests reject H0

# QQ line
qq_line <- ggplot(data = data.frame(resid = resid_pr), aes(sample = resid)) +
  stat_qq() +
  stat_qq_line() +
  labs(title = "Q-Q Plot of Residuals", x = "Theoretical Quantiles", y = "Sample Quantiles") +
  theme_minimal()


#Plotting residual density against a normal distribution for comparison. NB this is using the residual objects created
from the secondary model above
sec1_resid_df <- as.data.frame(resid_pr)
sd_resid <- sd(resid_pr)
max_resid <- max(resid_pr)
min_resid <- min(resid_pr)

resid_sec1_kd<- ggplot(sec1_resid_df, aes(x = resid_pr)) +
  geom_density(aes(color = "Kernel density"), linewidth = 1) +
  stat_function(aes(color = "Normal density"),
                fun = function(x) dnorm(x, mean = 0, sd = sd_resid),
                linetype = "dotted", linewidth = 1) +
  theme_minimal() +
  scale_x_continuous(limits = c(min_resid, max_resid)) +
  scale_color_manual(name = "Legend", values = c("Kernel density" = "#000080", "Normal density" = "red")) +
  labs(
    title = "Executive Function Model 1 - Residuals Density Plot",
    x = "Residuals",
    y = "Density"
  )

#Testing linearity and assumption that residual errors have a mean of 0
plot(sec1_model, col = "red") #We want the line here to be horizontal and at 0

### Figures to export:
resid_sec1_kd
qq_line

plots <- list(resid_sec1_kd, qq_line)

### Export graphs to word

# Create a Word document
doc <- read_docx()

# Save each plot as an image and add to the Word document
for (i in seq_along(plots)) {
  # Export the plot to a temporary file
  file_name <- tempfile(fileext = ".png")
  ggplot2::ggsave(filename = file_name, plot = plots[[i]])

  # Add the plot to the Word document
  doc <- doc %>%
```

```
  body_add_par(value = paste("Plot", i), style = "heading 1") %>%
  body_add_img(src = file_name, width = 6, height = 4)
}

# Save the document
print(doc, target = "//Nevis/RE/Projects/In_Progress/022807.014_P2899_EEF_The_ONE/WIP/09. Impact/Graphs &
Summary Tables/SensitivityAnlaysis2_1.docx")

### 3.  Bootstrapping CIs and p-value
#+ The plots and normality test results suggest
#+ that the residuals are not normally distributed. We will
#+ therefore need to re-estimate CIs and p-values using bootstrapping.

set.seed(999)
boot_sec1_model <- boot.pval::boot_summary(sec1_model,
                                  type = "norm",
                                  method = NULL,
                                  conf.level = 0.95)
boot_sec1_model

######################




### 4.  Effect-size estimation

hedges.g <- function(c, n, m, v, w){
  c/sqrt((((n - 1)*v)+((m - 1)* w))/n + (m-2))
}

#We use the output from the model above to calculate the effect size: (NB: This needs to be done for confidence
intervals too!)
#We need to define the objects needed to be inputted into the Hedges G function created above (c, n, m, v, & w)

coefs <- data.frame(summary(sec1_model)$coefficients) #create data frame of all coefficients from secondary
outcome model

c <- coefs["treatment", "Estimate"] #Extract the treatment coefficient

n <- nrow(dplyr::filter(data,
                  treatment == 1 &
                    !is.na(E_HTKS_total) &
                    !is.na(B_Corsi_total)  &
                    !is.na(SettingRegion) &
                    !is.na(Settingtype) &
                    !is.na(SettingID) &
                    !is.na(AgeAtBase))) #Number of individuals in treatment group in the model

m <- nrow(dplyr::filter(data,
                  treatment == 0 &
                    !is.na(E_HTKS_total) &
                    !is.na(B_Corsi_total)  &
                    !is.na(SettingRegion) &
```

```
                  !is.na(Settingtype) &
                  !is.na(SettingID) &
                  !is.na(AgeAtBase))) #Number of individuals in control group in the model

v <- var(dplyr::filter(data,
                  treatment == 1 &
                  !is.na(B_Corsi_total) &
                  !is.na(SettingRegion) &
                  !is.na(Settingtype) &
                  !is.na(E_HTKS_total) &
                  !is.na(SettingID) &
                  !is.na(AgeAtBase))$E_HTKS_total) #Variance in outcome among treatment group

w <- var(dplyr::filter(data,
                  treatment == 0 &
                  !is.na(B_Corsi_total) &
                  !is.na(SettingRegion) &
                  !is.na(Settingtype) &
                  !is.na(E_HTKS_total) &
                  !is.na(SettingID) &
                  !is.na(AgeAtBase))$E_HTKS_total)  #Variance in outcome among control group

secondary1_effect_size <- hedges.g(c, n, m, v, w)
secondary1_effect_size

#LOWER CI EFFECT SIZE using lower 95% CI from the bootstrapped model:

#We redefine c to equal the treatment coefficient minus the SE created above
c <- boot_sec1_model$Lower.bound[2]

sec1_effect_size_low <- hedges.g(c, n, m, v, w) #Re-run hedges G on lower CI treatment coefficient
sec1_effect_size_low #This is the lower confidence interval for the treatment effect on the following directions subtest.



#UPPER CI EFFECT SIZE using bootrapped 95% upper CI from the above model:

#We redefine c again to equal the treatment coefficient plus the SE created above. Remember c has already been
defined to be the lower confidence interval, we therefore need to add 2 SEs to it.
c <- boot_sec1_model$Upper.bound[2]

sec1_effect_size_high <- hedges.g(c, n, m, v, w) #Re-run hedges G on upper CI treatment coefficient
sec1_effect_size_high #This is the upper confidence interval for the treatment effect on the following directions
subtest.

## P-value
p_values <- boot_sec1_model$p.value[2]
p_value <- round(p_values, digits = 2)


### 5. Creating secondary analysis table (to paste output into report)

##+ This part of the code is to create a formatted output table which can just be copied
##+ into the table empty table in word
```

```
## Find missing numbers for model (number of obs with outcome but missing covariates)

# Treatment
n_mis <- nrow(dplyr::filter(data,
                        treatment == 1 &
                          !is.na(E_HTKS_total) &
                          (is.na(B_Corsi_total) |
                             is.na(SettingRegion) |
                             is.na(Settingtype) |
                             is.na(SettingID) |
                             is.na(AgeAtBase))))

# Control
m_mis <- nrow(dplyr::filter(data,
                        treatment == 0 &
                          !is.na(E_HTKS_total) &
                          (is.na(B_Corsi_total) |
                             is.na(SettingRegion) |
                             is.na(Settingtype) |
                             is.na(SettingID) |
                             is.na(AgeAtBase))))

# Paste in numbers (non-missing and missing) for table
num_t <- paste0(as.character(n), " (", as.character(n_mis), ")")
num_c <- paste0(as.character(m), " (", as.character(m_mis), ")")

# Total numbers
total_n <- n + m
total_n_t_c <- paste0(as.character(total_n), " (", n, ";", m, ")")

## Generate means and CIs

# Treatment
me_t <- t.test(dplyr::filter(data,
                        treatment == 1 &
                          !is.na(E_HTKS_total) &
                          !is.na(B_Corsi_total)  &
                          !is.na(SettingRegion) &
                          !is.na(Settingtype) &
                          !is.na(SettingID) &
                          !is.na(AgeAtBase))$E_HTKS_total) #Run t-test on outcome in analytical sample

mean_t <- as.character(round(me_t$estimate, digits = 2)) #extract mean from t-test object (me), to 2dp and as
character to facilitate pasting in CIs.

me_ci_t <- as.character(round(me_t$conf.int[c(1, 2)], digits = 2)) #extract CIs (95%) from t-test object (me), to 2dp and
as character to facilitate pasting

mean_ci_t <- paste0(mean_t, " (", me_ci_t[1], " - ", me_ci_t[2], ")") #Paste CIs in brackets next to mean figure

# Control
me_c <- t.test(dplyr::filter(data,
                        treatment == 0 &
```

```
                    !is.na(E_HTKS_total) &
                    !is.na(B_Corsi_total) &
                    !is.na(SettingRegion) &
                    !is.na(Settingtype) &
                    !is.na(SettingID) &
                    !is.na(AgeAtBase))$E_HTKS_total)

mean_c <- as.character(round(me_c$estimate, digits = 2))
me_ci_c <- as.character(round(me_c$conf.int[c(1, 2)], digits = 2))
mean_ci_c <- paste0(mean_c, " (", me_ci_c[1], " - ", me_ci_c[2], ")")


## Effect size with CIs
ES_CI <- paste0(as.character(round(secondary1_effect_size, digits = 3)),
            " (", as.character(round(sec1_effect_size_low, digits = 2)),
            " -- ",
            as.character(round(sec1_effect_size_high, digits = 2)),
            ")")

## Create table

secondary_analysis_table_1 <- t(as.data.frame(c(num_t, mean_ci_t, num_c, mean_ci_c, total_n_t_c, ES_CI,
p_value)))
colnames(secondary_analysis_table_1) <- c("N (intervention)", "Mean (intervention", "N (control)", "Mean (control)",
"Total (T;C)", "Hedges g (95% CIs)", "p-value")

rownames(secondary_analysis_table_1) <- "HTKS Executive Function Test"

## Display results - TABLE 1
View(secondary_analysis_table_1)

############

### 6. Creating effect size estimation table. Most of the required fields have already been created

u_m <- as.character(round((mean(dplyr::filter(data, #Unadjusted difference in means
                                    treatment == 1 &
                                      !is.na(B_Corsi_total) &
                                      !is.na(SettingRegion) &
                                      !is.na(Settingtype) &
                                      !is.na(SettingID) &
                                      !is.na(AgeAtBase))$E_HTKS_total, na.rm = TRUE)) - (mean(dplyr::filter(data,
                                                                                              treatment == 0 &
                                                                                                !is.na(B_Corsi_total) &
                                                                                                !is.na(SettingRegion) &
                                                                                                !is.na(Settingtype) &
                                                                                                !is.na(SettingID) &
                                                                                                !is.na(AgeAtBase))$E_HTKS_total, na.rm
= TRUE)), digits = 2))

c <- round(coefs["treatment", "Estimate"], digits = 2) #Adjusted difference in means

var_t <- round(v, digits = 2) #Variance of outcome in treatment
```

```
var_c <- round(w, digits = 2) #Variance of outcome in control

pooled_var <- round(((((n - 1)*v)+((m - 1)* w))/(n + m - 2)), digits = 2) #pooled variance


# Create table 2
sec1_es_est <- t(as.data.frame(c(u_m, c, num_t, var_t, num_c, var_c, pooled_var)))
colnames(sec1_es_est) <- c("Unadj. diff in means", "Adj diff in means", "N (treatment)", "Variance (treatment)", "N
(control)", "Variance (control)", "Pooled Variance")
rownames(sec1_es_est) <- "HTKS Executive Function Test"
View(sec1_es_est)
################## The ONE: Missing Data Analysis ####################


######## Purpose of Script
####+ 1. Explore descriptive overview of missing endline data (primary outcome)
####+ 2. Run logit models to predict endline missingness on baseline covariate
####+ 3. Explore descriptive overview of missing baseline data (primary outcome)
####+ 4. Run logit models to predict baseline missingness on baseline covariate
####+ 5. Conduct multiple imputation and re-estimate primary model

##### Clear workspace
rm(list = ls())

# Set seed
set.seed(999)

# Load packages
library("haven")
library("dplyr")
library("ggplot2")
library("lme4")
library("sjstats")
library("Hmisc")
library("performance")
library("lmtest")
library("eeptools")
library("dplyr")
library("lmerTest")
library("jtools")
library("mice")
library("lattice")


# Load cleaned data
data <- read_dta("//Nevis/RE/Projects/In_Progress/022807.014_P2899_EEF_The_ONE/WIP/06. Data collection/02.
Endline/02. Data_NotMirrored/BaselineEndlineComplianceClean2.dta")

data <- filter(data, endlineonly != 1) #NB this is also removing missing obs for endlineonly flag.



#### 2 approaches to missingness
```

```
### Data dictionary

# E_score = Endline score
# B_score = Baseline score
# SettingRegion = Region
# Settingtype = PVI
# SettingID = Setting ID
# E_Child_Gender = Gender at endline
# E_EYPPStatus = EYPP at endline
# Age = AgeAtBase

# Secondary outcomes: HTKS-R & Corsi blocks


#### Recoding obs where variables coded as "" as missing

# Child ID
data$child_id <- ifelse(data$child_id == "", NA, data$child_id)

# Setting ID
data$SettingID <- ifelse(data$SettingID == "", NA, data$SettingID)

# Setting Name
data$Setting_Name <- ifelse(data$Setting_Name == "", NA, data$Setting_Name)

# Region
data$SettingRegion <- ifelse(data$SettingRegion == "", NA, data$SettingRegion)


# PVI status
data$Settingtype <- ifelse(data$Settingtype == "", NA, data$Settingtype)

# Gender (endline)
data$E_Child_Gender <- ifelse(data$E_Child_Gender == "", NA, data$E_Child_Gender)

# Gender (baseline)
data$B_ChildsGender <- ifelse(data$B_ChildsGender == "", NA, data$B_ChildsGender)

# EYPP status (endline)
data$E_EYPPStatus <- ifelse(data$E_EYPPStatus == "", NA, data$E_EYPPStatus)

# EYPP status (baseline)
data$B_EYPP_status <- ifelse(data$B_EYPP_status == "", NA, data$B_EYPP_status)




## Remove observations where child_id is missing @Fin to flag. A quick look at the data suggests these are non-observations (missing for all other variables)-
data <- filter(data, !is.na(child_id),
               !is.na(SettingID))


#### Create new baseline characteristic variables which combine endline and baseline measurement.
```

```
#+ Note that while we will create a new EYPP flag combining endline and baseline, this does not
#+ have implications for the main sub-group analysis as the analytical sample has no missing
#+ endline EYPP information.


#EYPP
data <- data %>%
  mutate(EYPP = E_EYPPStatus) # Create EYPP flag equal to endline EYPP

data$EYPP <- ifelse(is.na(data$EYPP), data$B_EYPP_status, data$EYPP) #Replace EYPP flag with baseline EYPP
if EYPP flag is missing

#Gender
data <- data %>%
  mutate(gender = E_Child_Gender) # Create gender flag equal to endline gender

data$gender <- ifelse(is.na(data$gender), data$B_ChildsGender, data$gender) #Replace gender flag with baseline
gender if gender flag is missing

####

### Remove endline/baseline variables for EYPP and gender so as to avoid mistakes with accidentally using these old
variables
data <- subset(data, select=-c(E_EYPPStatus, B_EYPP_status, E_Child_Gender, B_ChildsGender))

##+ Recode missing region information if we have Setting information- there is one instance where a child is
##+  recorded as being at this setting but setting region information is missing. NB this does not have
##+  implications for the analysis as the child is missing baseline data.
#data$SettingRegion <- ifelse(data$Setting_Name == "Abacus Ark (Wandsworth)", "London", data$SettingRegion)

############### Endline Missingness

#### Stage 1: Descriptive Overview

#+ This will involve tabulating balance of characteristics in treatment/control groups at baseline
#+ and comparing this to the balance in the same variables at endline.
#+




############ In non-analytical data (at baseline)

###### Child-level variables

## EYPP
EYPP_t_bas <- as.data.frame(table(filter(data, treatment == 1)$EYPP)) # treatment
prop.table(table(filter(data, treatment == 1)$EYPP))

EYPP_c_bas <- as.data.frame(table(filter(data, treatment == 0)$EYPP)) # control
prop.table(table(filter(data, treatment == 0)$EYPP))
```

```
## Gender
gender_t_bas <- as.data.frame(table(filter(data, treatment == 1)$gender)) # treatment
prop.table(table(filter(data, treatment == 1)$gender))

gender_c_bas <- as.data.frame(table(filter(data, treatment == 0)$gender)) # control
prop.table(table(filter(data, treatment == 0)$gender))

## Age at baseline
mean(filter(data, treatment == 1)$AgeAtBase, na.rm = T) # mean of treatment
sd(filter(data, treatment == 1)$AgeAtBase, na.rm = T) # sd of treatment
sum(!is.na(filter(data, treatment == 1)$AgeAtBase)) #N of missing obs in treatment

mean(filter(data, treatment == 0)$AgeAtBase, na.rm = T) # Mean of control
sd(filter(data, treatment == 0)$AgeAtBase, na.rm = T) # sd of control
sum(!is.na(filter(data, treatment == 0)$AgeAtBase)) # N of missing obs in control


##


###### Setting-level variables


#### Categorical vars

## Create dataset which only contains unique values for categorical setting-level variables

# Isolate vars
data_settings <- data %>%
  select("SettingID", "SettingRegion", "Settingtype", "treatment")

# Filter out missing values- this is fine as we're only interested in whole setting missingness, not child-level
data_settings <- filter(data_settings, !is.na(SettingRegion), !is.na(Settingtype))

# Retain on unique obs (one ob per setting)
data_settings <- unique(data_settings)



### Region

# Treatment
region_t_bas <- data.frame((table(dplyr::filter(data_settings,
                  treatment == 1)$SettingRegion)))

prop.table(table(dplyr::filter(data_settings,
                           treatment == 1)$SettingRegion)) # Percentage table


# Control
region_c_bas <- data.frame(table(dplyr::filter(data_settings,
                  treatment == 0)$SettingRegion))
```

```
prop.table(table(dplyr::filter(data_settings,
                               treatment == 0)$SettingRegion))



### Setting type

# Treatment
PVI_t_bas <- data.frame(table(dplyr::filter(data_settings,
                    treatment == 1)$Settingtype))

prop.table(table(dplyr::filter(data_settings,
                               treatment == 1)$Settingtype))


# Control
PVI_c_bas <- data.frame(table(dplyr::filter(data_settings,
                    treatment == 0)$Settingtype))

prop.table(table(dplyr::filter(data_settings,
                               treatment == 0)$Settingtype))


##########


##### Continuous setting-level vars

# Create dataset with required vars (Gender, EYPP, treatment)
data_settings_cont <- data %>%
  select("SettingID", "gender", "EYPP", "treatment")

data_settings_cont_t <- filter(data_settings_cont, treatment == 1)
data_settings_cont_c <- filter(data_settings_cont, treatment == 0)


###+ GENDER: Proportion of female children- we'll do this by creating loops which will
###+ calculate the average proportion of female students across settings

# Treatment
fem_prop_t_b_list <- c()

for(i in unique(data_settings_cont_t$SettingID)){
  d <- filter(data_settings_cont_t, SettingID == i)
  f <- sum(!is.na(filter(d, gender == "Female")$gender))
  tot <- sum(!is.na(d$gender))
  f_p <- f/tot
  fem_prop_t_b_list <- c(fem_prop_t_b_list, f_p)
}

fem_prop_t_b <- mean(fem_prop_t_b_list, na.rm = T) # Average proportion of female children in treatment settings (at
randomisation)
fem_prop_t_b
sd(fem_prop_t_b_list, na.rm = T) # SD
```

```
sum(is.na(fem_prop_t_b_list)) # Number of treatment settings with fully missing gender data

# Control
fem_prop_c_b_list <- c()

for(i in unique(data_settings_cont_c$SettingID)){
  d <- filter(data_settings_cont_c, SettingID == i)
  f <- sum(!is.na(filter(d, gender == "Female")$gender))
  tot <- sum(!is.na(d$gender))
  f_p <- f/tot
  fem_prop_c_b_list <- c(fem_prop_c_b_list, f_p)
}

fem_prop_c_b <- mean(fem_prop_c_b_list, na.rm = T) # Average proporiton of female children in control settings (at randomisation)
fem_prop_c_b
sd(fem_prop_c_b_list, na.rm = T) #SD
sum(is.na(fem_prop_c_b_list)) # Number of control settings with fully missing gender data (at randomisation)

####


###+ EYPP: Proportion of EYPP-eligible pupils at baseline

# Treatment
EYPP_prop_t_b_list <- c()

for(i in unique(data_settings_cont_t$SettingID)){
  d <- filter(data_settings_cont_t, SettingID == i)
  e <- sum(!is.na(filter(d, EYPP == "Y")$EYPP))
  tot <- sum(!is.na(d$EYPP))
  e_p <- f/tot
  EYPP_prop_t_b_list <- c(EYPP_prop_t_b_list, e_p)
}

EYPP_prop_t_b <- mean(EYPP_prop_t_b_list, na.rm = T) # Average proporiton of EYPP children in treatment settings (at randomisation)
EYPP_prop_t_b
sd(EYPP_prop_t_b_list, na.rm = T) #SD
sum(is.na(EYPP_prop_t_b_list)) # Number of treatment settings with fully missing EYPP data (at randomisation)


# Control
EYPP_prop_c_b_list <- c()

for(i in unique(data_settings_cont_c$SettingID)){
  d <- filter(data_settings_cont_c, SettingID == i)
  e <- sum(!is.na(filter(d, EYPP == "Y")$EYPP))
  tot <- sum(!is.na(d$EYPP))
  e_p <- f/tot
  EYPP_prop_c_b_list <- c(EYPP_prop_c_b_list, e_p)
}
```

```
EYPP_prop_c_b <- mean(EYPP_prop_c_b_list, na.rm = T) # Average proporiton of EYPP children in control settings (at randomisation)
EYPP_prop_c_b
sd(EYPP_prop_c_b_list, na.rm = T) # SD
sum(is.na(EYPP_prop_c_b_list)) # Number of control settings with fully missing EYPP data (at randomisation)



###

#### In analytical data (not missing for all variables included in primary model)

## Create object for analytical data
 data_ana <- filter(data, !is.na(E_score),
                    !is.na(B_score),
                    !is.na(treatment),
                    !is.na(SettingRegion),
                    !is.na(Settingtype),
                    !is.na(SettingID))



#### Child-level variables

## EYPP
EYPP_t_ana <- table(filter(data_ana, treatment == 1)$EYPP) # treatment
prop.table(table(filter(data_ana, treatment == 1)$EYPP))

EYPP_c_ana <- table(filter(data_ana, treatment == 0)$EYPP) # control
prop.table(table(filter(data_ana, treatment == 0)$EYPP))

## Gender
gender_t_ana <- table(filter(data_ana, treatment == 1)$gender) # treatment
prop.table(table(filter(data_ana, treatment == 1)$gender))

gender_c_ana <- table(filter(data_ana, treatment == 0)$gender) # control
prop.table(table(filter(data_ana, treatment == 0)$gender))

## Age at baseline
mean(filter(data_ana, treatment == 1)$AgeAtBase, na.rm = T) # mean of treatment
sd(filter(data_ana, treatment == 1)$AgeAtBase, na.rm = T) # sd of treatment
sum(!is.na(filter(data_ana, treatment == 1)$AgeAtBase)) #N of missing obs in treatment

mean(filter(data_ana, treatment == 0)$AgeAtBase, na.rm = T) # Mean of control
sd(filter(data_ana, treatment == 0)$AgeAtBase, na.rm = T) # sd of control
sum(!is.na(filter(data_ana, treatment == 0)$AgeAtBase)) # N of missing obs in control



##

###### Setting-level variables
```

```
## Create dataset which only contains unique values for continuous setting-level variables within analytical data

# Isolate vars
data_settings_ana <- data_ana %>%
  select("SettingID", "SettingRegion", "Settingtype", "treatment")

# Filter out missing values- this is fine as we're only interested in whole setting misingness, not child-level
data_settings_ana <- filter(data_settings_ana, !is.na(SettingRegion), !is.na(Settingtype))

# Retain on unique obs (one ob per setting)
data_settings_ana <- unique(data_settings_ana)



### Region

# Treatment
region_t_ana <- data.frame((table(dplyr::filter(data_settings_ana,
                                    treatment == 1)$SettingRegion)))

prop.table(table(dplyr::filter(data_settings_ana,
                         treatment == 1)$SettingRegion)) # Percentage table


# Control
region_c_ana <- data.frame(table(dplyr::filter(data_settings_ana,
                                     treatment == 0)$SettingRegion))

prop.table(table(dplyr::filter(data_settings_ana,
                         treatment == 0)$SettingRegion))



### Setting type

# Treatment
PVI_t_ana <- data.frame(table(dplyr::filter(data_settings_ana,
                                  treatment == 1)$Settingtype))

prop.table(table(dplyr::filter(data_settings_ana,
                         treatment == 1)$Settingtype))


# Control
PVI_c__ana <- data.frame(table(dplyr::filter(data_settings_ana,
                                  treatment == 0)$Settingtype))

prop.table(table(dplyr::filter(data_settings_ana,
                         treatment == 0)$Settingtype))



##### Continuous setting-level vars
```

```
# Create anaytical dataset with required vars (Gender, EYPP, treatment)
data_settings_ana_cont <- data_ana %>%
  select("SettingID", "gender", "EYPP", "treatment")

data_settings_ana_cont_t <- filter(data_settings_ana_cont, treatment == 1)
data_settings_ana_cont_c <- filter(data_settings_ana_cont, treatment == 0)


###+ GENDER: Proportion of female children- we'll do this by creating loops which will
###+ calculate the average proportion of female students across settings at analysis stage

# Treatment
fem_prop_t_a_list <- c()

for(i in unique(data_settings_ana_cont_t$SettingID)){
  d <- filter(data_settings_ana_cont_t, SettingID == i)
  f <- sum(!is.na(filter(d, gender == "Female")$gender))
  tot <- sum(!is.na(d$gender))
  f_p <- f/tot
  fem_prop_t_a_list <- c(fem_prop_t_a_list, f_p)
}

fem_prop_t_a <- mean(fem_prop_t_a_list, na.rm = T) # Average proporiton of female children in treatment settings (at
analysis)
fem_prop_t_a
sd(fem_prop_t_a_list, na.rm = T) #SD
sum(is.na(fem_prop_t_a_list)) # Number of treatment settings with fully missing gender data at analysis

# Control
fem_prop_c_a_list <- c()

for(i in unique(data_settings_ana_cont_c$SettingID)){
  d <- filter(data_settings_ana_cont_c, SettingID == i)
  f <- sum(!is.na(filter(d, gender == "Female")$gender))
  tot <- sum(!is.na(d$gender))
  f_p <- f/tot
  fem_prop_c_a_list <- c(fem_prop_c_a_list, f_p)
}

fem_prop_c_a <- mean(fem_prop_c_a_list, na.rm = T) # Average proporiton of female children in control settings (at
analysis)
fem_prop_c_a
sd(fem_prop_c_a_list, na.rm = T)
sum(is.na(fem_prop_c_a_list)) # Number of control settings with fully missing gender data (at analysis)

####


###+ EYPP: Proportion of EYPP-eligible pupils at analysis

# Treatment
EYPP_prop_t_a_list <- c()
```

```
for(i in unique(data_settings_ana_cont_t$SettingID)){
  d <- filter(data_settings_ana_cont_t, SettingID == i)
  e <- sum(!is.na(filter(d, EYPP == "Y")$EYPP))
  tot <- sum(!is.na(d$EYPP))
  e_p <- f/tot
  EYPP_prop_t_a_list <- c(EYPP_prop_t_a_list, e_p)
}

EYPP_prop_t_a <- mean(EYPP_prop_t_a_list, na.rm = T) # Average proporiton of EYPP children in treatment settings
(at analysis)
EYPP_prop_t_a
sd(EYPP_prop_t_a_list, na.rm = T) #SD
sum(is.na(EYPP_prop_t_a_list)) # Number of treatment settings with fully missing EYPP data (at analysis)


# Control
EYPP_prop_c_a_list <- c()

for(i in unique(data_settings_ana_cont_c$SettingID)){
  d <- filter(data_settings_ana_cont_c, SettingID == i)
  e <- sum(!is.na(filter(d, EYPP == "Y")$EYPP))
  tot <- sum(!is.na(d$EYPP))
  e_p <- f/tot
  EYPP_prop_c_a_list <- c(EYPP_prop_c_a_list, e_p)
}

EYPP_prop_c_a <- mean(EYPP_prop_c_a_list, na.rm = T) # Average proporiton of EYPP children in control settings
(at analysis)
EYPP_prop_c_a
sd(EYPP_prop_c_a_list, na.rm = T) #SD
sum(is.na(EYPP_prop_c_a_list)) # Number of control settings with fully missing EYPP data (at analysis)


#### Exploring make-up of characteristics in missing endline sample

## Create a binary flag denoting missingness in the primary outcome
data <- data %>%
  dplyr::mutate(missing_flag = ifelse(is.na(E_score), 1, 0))


# Create dataset of missing endline sample (non-missing for other covariates in primary model)
data_mis <- filter(data,
                   !is.na(B_score)&
                     !is.na(Settingtype) &
                     !is.na(SettingRegion)&
                     missing_flag == 1)

#### Child-level variables

## EYPP
EYPP_t_mis <- table(filter(data_mis, treatment == 1)$EYPP) # treatment
EYPP_t_mis
prop.table(table(filter(data_mis, treatment == 1)$EYPP))
```

```
EYPP_c_mis <- table(filter(data_mis, treatment == 0)$EYPP) # control
EYPP_c_mis
prop.table(table(filter(data_mis, treatment == 0)$EYPP))

## Gender
gender_t_mis <- table(filter(data_mis, treatment == 1)$gender) # treatment
gender_t_mis
prop.table(table(filter(data_mis, treatment == 1)$gender))

gender_c_mis <- table(filter(data_mis, treatment == 0)$gender) # control
gender_c_mis
prop.table(table(filter(data_mis, treatment == 0)$gender))

## Age at baseline
mean(filter(data_mis, treatment == 1)$AgeAtBase, na.rm = T) # mean of treatment
sd(filter(data_mis, treatment == 1)$AgeAtBase, na.rm = T) # sd of treatment
sum(!is.na(filter(data_mis, treatment == 1)$AgeAtBase)) #N of missing obs in treatment

mean(filter(data_mis, treatment == 0)$AgeAtBase, na.rm = T) # Mean of control
sd(filter(data_mis, treatment == 0)$AgeAtBase, na.rm = T) # sd of control
sum(!is.na(filter(data_mis, treatment == 0)$AgeAtBase)) # N of missing obs in control


### Categorical setting-level vars

## Create dataset which only contains unique values for continuous setting-level variables within missing endline data

# Isolate vars
data_settings_mis <- data_mis %>%
  select("SettingID", "SettingRegion", "Settingtype", "treatment")

# Filter out missing values- this is fine as we're only interested in whole setting misingness, not child-level
data_settings_mis <- filter(data_settings_mis, !is.na(SettingRegion), !is.na(Settingtype))

# Retain on unique obs (one ob per setting)
data_settings_mis <- unique(data_settings_mis)


### Region

# Treatment
  region_t_mis <- data.frame((table(dplyr::filter(data_settings_mis,
                                    treatment == 1)$SettingRegion)))

prop.table(table(dplyr::filter(data_settings_mis,
                         treatment == 1)$SettingRegion)) # Percentage table


# Control
region_c_mis <- data.frame(table(dplyr::filter(data_settings_mis,
                                   treatment == 0)$SettingRegion))

prop.table(table(dplyr::filter(data_settings_mis,
                         treatment == 0)$SettingRegion))
```

```
### Setting type

# Treatment
PVI_t_mis <- data.frame(table(dplyr::filter(data_settings_mis,
                                  treatment == 1)$Settingtype))

prop.table(table(dplyr::filter(data_settings_mis,
                           treatment == 1)$Settingtype))


# Control
PVI_c__mis <- data.frame(table(dplyr::filter(data_settings_mis,
                                   treatment == 0)$Settingtype))

prop.table(table(dplyr::filter(data_settings_mis,
                           treatment == 0)$Settingtype))




##### Continuous setting-level vars

# Create missing dataset with required vars (Gender, EYPP, treatment)
data_settings_mis_cont <- data_mis %>%
  select("SettingID", "gender", "EYPP", "treatment")

data_settings_mis_cont_t <- filter(data_settings_mis_cont, treatment == 1)
data_settings_mis_cont_c <- filter(data_settings_mis_cont, treatment == 0)


###+ GENDER: Proportion of female children- we'll do this by creating loops which will
###+ calculate the average proportion of female students across settings in missing data

# Treatment
fem_prop_t_a_list <- c()

for(i in unique(data_settings_mis_cont_t$SettingID)){
  d <- filter(data_settings_mis_cont_t, SettingID == i)
  f <- sum(!is.na(filter(d, gender == "Female")$gender))
  tot <- sum(!is.na(d$gender))
  f_p <- f/tot
  fem_prop_t_a_list <- c(fem_prop_t_a_list, f_p)
}

fem_prop_t_a <- mean(fem_prop_t_a_list, na.rm = T) # Average proporiton of female children in treatment settings (in
missing data)
fem_prop_t_a
sd(fem_prop_t_a_list, na.rm = T) #SD
sum(is.na(fem_prop_t_a_list)) # Number of treatment settings with fully missing gender data in missing data

# Control
```

```
fem_prop_c_a_list <- c()

for(i in unique(data_settings_mis_cont_c$SettingID)){
  d <- filter(data_settings_mis_cont_c, SettingID == i)
  f <- sum(!is.na(filter(d, gender == "Female")$gender))
  tot <- sum(!is.na(d$gender))
  f_p <- f/tot
  fem_prop_c_a_list <- c(fem_prop_c_a_list, f_p)
}

fem_prop_c_a <- mean(fem_prop_c_a_list, na.rm = T) # Average proporiton of female children in control settings (in missing data)
fem_prop_c_a
sd(fem_prop_c_a_list, na.rm = T)
sum(is.na(fem_prop_c_a_list)) # Numbe of control settings with fully missing gender data in missing data

####


###+ EYPP: Proportion of EYPP-eligible pupils in missing data

# Treatment
EYPP_prop_t_a_list <- c()

for(i in unique(data_settings_mis_cont_t$SettingID)){
  d <- filter(data_settings_mis_cont_t, SettingID == i)
  e <- sum(!is.na(filter(d, EYPP == "Y")$EYPP))
  tot <- sum(!is.na(d$EYPP))
  e_p <- f/tot
  EYPP_prop_t_a_list <- c(EYPP_prop_t_a_list, e_p)
}

EYPP_prop_t_a <- mean(EYPP_prop_t_a_list, na.rm = T) # Average proporiton of EYPP children in treatment settings in missing data
EYPP_prop_t_a
sd(EYPP_prop_t_a_list, na.rm = T) #SD
sum(is.na(EYPP_prop_t_a_list)) # Number of treatment settings with fully missing EYPP data in missing data


# Control
EYPP_prop_c_a_list <- c()

for(i in unique(data_settings_mis_cont_c$SettingID)){
  d <- filter(data_settings_mis_cont_c, SettingID == i)
  e <- sum(!is.na(filter(d, EYPP == "Y")$EYPP))
  tot <- sum(!is.na(d$EYPP))
  e_p <- f/tot
  EYPP_prop_c_a_list <- c(EYPP_prop_c_a_list, e_p)
}

EYPP_prop_c_a <- mean(EYPP_prop_c_a_list, na.rm = T) # Average proporiton of EYPP children in control settings in missing data
EYPP_prop_c_a
sd(EYPP_prop_c_a_list, na.rm = T) #SD
```

```
sum(is.na(EYPP_prop_c_a_list)) # Number of control settings with fully missing EYPP data in missing data



###########



#### Additional descriptive missing table: balance of characteristics in treatment and control (not split by t/c groups)

# Rereate a binary flag denoting missingness in the primary outcome
data <- data %>%
  dplyr::mutate(missing_flag = ifelse(is.na(E_score), 1, 0))

# Recreate missing data sub-sample (non-missing for other covariates in primary analysis)
data_mis <- filter(data,
                   !is.na(B_score)&
                     !is.na(Settingtype) &
                     !is.na(SettingRegion)&
                     missing_flag == 1)


## Child-level vars

## EYPP
table(data_mis$EYPP)
prop.table(table(data_mis$EYPP))

## Gender
table(data_mis$gender)
prop.table(table(data_mis$gender))

## Age at baseline
mean(data_mis$AgeAtBase, na.rm = T) # mean
sd(data_mis$AgeAtBase, na.rm = T) # sd

# Baselin scores
mean(data_mis$B_score)
sd(data_mis$B_score)

## Number of children by region
table(data_mis$SettingRegion)
prop.table(table(data_mis$SettingRegion))


## Number of children by setting type
table(data_mis$Settingtype)
prop.table(table(data_mis$Settingtype))

# Treatment assignment
table(data_mis$treatment)
prop.table(table(data_mis$treatment))
```

```
### Setting-level vars (categorical)

# Isolate required variables
data_settings_mis <- data_mis %>%
  select("SettingID", "SettingRegion", "Settingtype", "treatment")


# Filter out missing values- this is fine as we're only interested in whole setting misingness, not child-level
data_settings_mis <- filter(data_settings_mis, !is.na(SettingRegion), !is.na(Settingtype))

# Retain on unique obs (one ob per setting)
data_settings_mis <- unique(data_settings_mis)


## Region
table(data_settings_mis$SettingRegion)
prop.table(table(data_settings_mis$SettingRegion)) # Percentage table

## Setting type
table(data_settings_mis$Settingtype)
prop.table(table(data_settings_mis$Settingtype)) # Percentage table

#####



### Setting-level vars (continuous)

# Create missing dataset with required vars (Gender, EYPP, treatment)
data_settings_mis_cont <- data_mis %>%
  select("SettingID", "gender", "EYPP", "treatment")

#+ GENDER: Proportion of female children- we'll do this by creating loops which will
#+ calculate the average proportion of female students across settings in missing data

fem_prop_mis_list <- c()

for(i in unique(data_settings_mis_cont$SettingID)){
  d <- filter(data_settings_mis_cont, SettingID == i)
  f <- sum(!is.na(filter(d, gender == "Female")$gender))
  tot <- sum(!is.na(d$gender))
  f_p <- f/tot
  fem_prop_mis_list <- c(fem_prop_mis_list, f_p)
}

fem_prop_m <- mean(fem_prop_mis_list, na.rm = T) # Average proporiton of female children in treatment settings (in
missing data)
fem_prop_m
sd(fem_prop_mis_list, na.rm = T) #SD


#+ EYPP: Proportion of EYPP-eligible pupils in missing data
```

```
EYPP_prop_mis_list <- c()

for(i in unique(data_settings_mis_cont$SettingID)){
  d <- filter(data_settings_mis_cont, SettingID == i)
  e <- sum(!is.na(filter(d, EYPP == "Y")$EYPP))
  tot <- sum(!is.na(d$EYPP))
  e_p <- f/tot
  EYPP_prop_mis_list <- c(EYPP_prop_mis_list, e_p)
}

EYPP_prop_m <- mean(EYPP_prop_mis_list, na.rm = T) # Average proporiton of EYPP children in treatment settings in missing data
EYPP_prop_m
sd(EYPP_prop_mis_list, na.rm = T) #SD


#### 2. Modelling missingness in primary outcome at endline


# Find proportion of missingness in primary outcome
prop.table(table(data$missing_flag)) # around 9.3%

## Endline logit

#+ Modelling missingness at endline: treatment, region, baseline score, setting type, gender, EYPP, and age at baseline
logit_e <- glmer(missing_flag ~
                   treatment +
                   B_score +
                   SettingRegion +
                   Settingtype +
                   gender +
                   EYPP +
                   AgeAtBase +
                 (1 | SettingID),
                 data = data,
                 family = binomial)

summary(logit_e)
summ(logit_e)


## NB: As per the SAP, we have not included additional baseline secondary outcome scores as covariates, due to the high degree of missingness in these at baseline.

#### 3. Modelling missingness in primary outcome at baseline

# Create a binary flag denoting missingness in the primary outcome
data <- data %>%
  dplyr::mutate(missing_flag_b = ifelse(is.na(B_score), 1, 0))


# Modelling missingness at baseline: treatment, region, setting type, gender, EYPP, and age at baseline
logit_b <- glmer(missing_flag_b ~
```

```
            treatment +
            SettingRegion +
            Settingtype +
            gender +
            EYPP +
            AgeAtBase +
           (1 | SettingID),
           data = data,
           family = binomial)

summary(logit_b)
summ(logit_b)


############


####### Multiple Imputation

### [INSERT DESCRIPTION/JUSTIFICATION]- why are we using MICE?

### View the extent/distribution of missing data
#md.pattern(data)

#+ Isolate data to only include variables to be inlcuded in MI
#+ This means excluding secondary outcome baseline scores
data_imp <- select(data, "child_id", "AgeAtBase", "B_score", "E_score", "treatment", "SettingRegion", "Settingtype",
"gender", "EYPP", "SettingID")

# Remove observations with missing child ID
data_imp <- filter(data_imp, !is.na(child_id))

#+ Remove observations with both missing baseline AND endline.
#+ We only want to impute missing endline for children with baseline and vice-versa
data_imp <- data_imp %>%
  mutate(exclude = 0) #create exclude flag

data_imp$exclude <- ifelse(is.na(data_imp$E_score) & is.na(data_imp$B_score), 1, data_imp$exclude) #Assign value
of 1 if missing for both endline and baseline

data_imp <- filter(data_imp, exclude == 0) #Filter out all obs where exclude flag is equal to 1

# Remove exclude flag
data_imp <- select(data_imp, -"exclude")

### Factorise string variables. This is so that the imputation can run correctly
data_imp$EYPP <- factor(data_imp$EYPP, levels = c("N", "Y"))

data_imp$gender <- factor(data_imp$gender, levels = c("Male", "Female"))


# Convert SettingID into integer. This is so that the level-2 marker variable is correctly identified in the MI
data_imp$SettingID <- gsub("S",
                           "",
```

```
                data_imp$SettingID) # Remove S prefix

data_imp$SettingID <- as.integer(data_imp$SettingID) # Convert to integer


### Create predictor matrix
#+ This provides the information on which variables should be used to predict endline and baseline scores.
#+ It also allows us to specify which variable is the cluster variable (SettingID)

# Run empty imputation on data- this is so that we have a template to create the predictor matrix
ini <- mice(data_imp, maxit=0)

# Extract and modify the predictor matrix

pred <- ini$pred
pred #View predictor matrix

#+ Re-specify the predictor matrix for imputation so that it takes into account hierarchical nature of dataset
#+ We only want to impute endline and baseline scores. All other variables (rows in matrix) must be set to 0
#+ -2 denotes the level-2 signifier variable (SettingID)
#+ 2 denoted level-2 predictor
#+ 1 denotes level-1 predictor
#+ 0 denotes should not be used in predicting (imputing) outcome

pred["E_score", ] <- c(0, 1, 1, 0, 1, 2, 2, 1, 1, -2)  # Include all relevant predictors (rows) for endline scores
pred["B_score", ] <- c(0, 1, 0, 1, 1, 2, 2, 1, 1, -2)  # Same structure for baseline
pred["child_id", ] <- c(0,0,0,0,0,0,0,0,0,0)
pred["AgeAtBase", ] <- c(0,0,0,0,0,0,0,0,0,0)
pred["treatment", ] <- c(0,0,0,0,0,0,0,0,0,0)
pred["SettingRegion", ] <- c(0,0,0,0,0,0,0,0,0,0)
pred["Settingtype", ] <- c(0,0,0,0,0,0,0,0,0,0)
pred["gender", ] <- c(0,0,0,0,0,0,0,0,0,0)
pred["EYPP", ] <- c(0,0,0,0,0,0,0,0,0,0)
pred["SettingID", ] <- c(0,0,0,0,0,0,0,0,0,0)


### Specify the imputation methods for all variables
meth <- c("", "", "2l.lmer", "2l.lmer", "", "", "", "", "", "")  # Using "2l.lmer" for hierarchical imputation. This also specifies
that we're only imputing baseline/endline scores

# Create object indicating number of imputations
n_imp <- 20 # 20 iterations.


#### Run the multiple imputation process. Using the method vector and preditor matrix generated earlier.
# Set seed to equal the same seed set for previous analysis.
# Run 20 imputations

MI_data <- mice(data_imp,
                m=n_imp,
                meth=meth,
                pred=pred,
                maxit=20,
                seed=999)
```

```
# View imputed data
summary(MI_data)


### Run primary analysis model on pooled MI data.
pr_imp <- pool(with(MI_data, lmer(E_score ~
                          treatment +
                          B_score +
                          SettingRegion +
                          Settingtype +
                          (1 | SettingID))))

MI_results <- summary(pr_imp) # View pooled analysis results after imputation.
MI_results

### Creating pooled MI primary analysis ES


# Create long imputed dataset- this is so that we can obtain estimates for the variances
MI_data_long <- complete(MI_data, "long", complete = F)


# Create Hedges g function
hedges.g <- function(c, n, m, v, w){
  c/sqrt((((n - 1)*v)+((m - 1)* w))/(n + m-2))
}


# Assign objects for hedge.g function
c <- MI_results$estimate[2]
n <- nrow(filter(MI_data_long,
              !is.na(E_score) &
                !is.na(B_score) &
                !is.na(SettingRegion) &
                !is.na(Settingtype) &
                treatment == 1))/n_imp  # Number of children with complete data after imputation in treatment, divided by
number of imputations conducted
m <- nrow(filter(MI_data_long,
               !is.na(E_score) &
                 !is.na(B_score) &
                 !is.na(SettingRegion) &
                 !is.na(Settingtype) &
                 treatment == 0))/n_imp  # Number of children in control (as randomised)

 v <- var(filter(MI_data_long,
              !is.na(E_score) &
                !is.na(B_score) &
                !is.na(SettingRegion) &
                !is.na(Settingtype) &
                treatment == 1)$E_score, na.rm = T)  # Variance in treatment for the outcome across the MI datasets with
complete cases. Note we still need to exclude missing E_score obs as MI was only run where predictor covariates
were complete
w <- var(filter(MI_data_long,
```

```
            !is.na(E_score) &
              !is.na(B_score) &
              !is.na(SettingRegion) &
              !is.na(Settingtype) &
              treatment == 0)$E_score, na.rm = T) # Variance in control for the outcome across the MI datasets with
compete cases

# Create effect size
MI_ES <- hedges.g(c, n, m, v, w)
MI_ES

# Lower CI ES
c <- MI_results$estimate[2] - MI_results$std.error[2] # Redefine c to be 1 SE below coefficient
MI_ES_lo <- hedges.g(c, n, m, v, w)

# Upper CI ES
c <- MI_results$estimate[2] + MI_results$std.error[2] # Redefine c to be 1 SE above coefficient
MI_ES_hi <- hedges.g(c, n, m, v, w)


# Pooled mean- treatment
t.test(filter(MI_data_long, treatment == 1)$E_score)

# Pooled mean- control
t.test(filter(MI_data_long, treatment == 0)$E_score)

pooled_var<- round((((n - 1)*v)+((m - 1)* w))/(n + m-2), digits=2)
####################################
## The ONE Sensitivity Analysis ##
###### Measurement Error 1 #######
####################################

##+ Conducting sensitivity analysis of the ONE primary analysis with tests that did not follow the stopping
##+ rule correctly removed
##+ After filtering out these observations, the procedure will follow the primary analysis.

#############################################################################################

# Clear workspace
rm(list=ls())
set.seed(999) #NB this is the same seed set in the randomisation code at the start of the project so we shall retain it
here for consistency.

# Load packages
library("haven")
library("dplyr")
library("ggplot2")
library("lme4")
library("sjstats")
library("Hmisc")
library("performance")
library("lmtest")
library("eeptools")
library("lmerTest")
```

```
library("officer")
library("boot")
library("boot.pval")


# Load cleaned data
data <- read_dta("//Nevis/RE/Projects/In_Progress/022807.014_P2899_EEF_The_ONE/WIP/06. Data collection/02.
Endline/02. Data_NotMirrored/BaselineEndlineComplianceClean2.dta")
data <- filter(data, endlineonly != 1)


#Numbers of incomplete tests at baseline and endline
nrow(filter(data,
        E_EYTN_Stop == 1))

# 294 missing at endline

nrow(filter(data,
        B_EYTN_Stop == 1))

# 135 missing at baseline

### Filtering out incomplete tests
data <- dplyr::filter(data,
                      E_EYTN_Stop != 1 |
                      B_EYTN_Stop != 1)


#########################################################

# Section A: Descriptive statistics


#1. Histograms of outcomes

#Use this to check distribution of endline and baseline scores for our outcome

#We are mainly checking for ceiling and floor effects

# In overall data
EYTN_ovr_hist <- hist(data$E_score,
                      main="Histogram of EYTN subtest at endline (overall sample - excluding incomplete tests)",
                      xlab = "Early Numeracy Scores",
                      xlim = c(0, 120),
                      breaks = seq(0, 120, 10),
                      xaxp = c(0, 120, 6))

#+In analytical data (the analytical data is defined as the data used in the multi-level regression model.
#+It includes the observations where we have complete data for all covariates included in the outcome model)
EYTN_model_hist <- hist(dplyr::filter(data,
                                      !is.na(B_score) &
                                        !is.na(Settingtype) &
                                        !is.na(SettingRegion) &
                                        !is.na(SettingID))$E_score,
```

```
                main="Histogram of EYTN subtest at endline (analytical sample - excluding incomplete tests)",
                xlab = "Early numeracy scores",
                xlim = c(0, 120),
                breaks = seq(0, 120, 10),
                xaxp = c(0, 120, 6))


## No floor or ceiling effects in endline test distribution
## Now for baseline
EYTN_ovr_histB <- hist(data$B_score,
                main="Histogram of EYTN subtest at baseline (overall sample - excluding incomplete tests)",
                xlab = "Early Numeracy Scores",
                xlim = c(0, 120),
                breaks = seq(0, 120, 10),
                xaxp = c(0, 120, 6))

#+In analytical data (the analytical data is defined as the data used in the multi-level regression model.
#+It includes the observations where we have complete data for all covariates included in the outcome model)
EYTN_model_histB <- hist(dplyr::filter(data,
                               !is.na(B_score) &
                                 !is.na(Settingtype) &
                                 !is.na(SettingRegion) &
                                 !is.na(SettingID))$B_score,
                 main="Histogram of EYTN subtest at baseline (analytical sample - excluding incomplete tests)",
                 xlab = "Early numeracy scores",
                 xlim = c(0, 120),
                 breaks = seq(0, 120, 10),
                 xaxp = c(0, 120, 6))

#+ We're unlikely to include any additional sensitivity analysis for this as any floor effects are only observed at
baseline.
#+ However, will include additional analysis here to see what proportion of the observations are within one standard
deviation of the floor.

sd <- sd(data$B_score, na.rm = T)

# Create df which is a copy of analytical data
data_sd <- dplyr::filter(data,
                 !is.na(B_score) &
                   !is.na(Settingtype) &
                   !is.na(SettingRegion) &
                   !is.na(SettingID))

data_sd <- data_sd %>%
  mutate(sd_dummy = 0)

data_sd$sd_dummy <- ifelse(data_sd$B_score < sd, 1, 0)

mean(data_sd$sd_dummy)

##Here we see that 30.66% of observations are within one sd of the floor

#2. Means, SDs, Min and Max (for whole sample and analytical sample)
```

```
#Overall sample
Hmisc::describe(dplyr::filter(data,
                        !is.na(E_score))$E_score)
sd(dplyr::filter(data,
            !is.na(E_score))$E_score)


##Analytical sample
Hmisc::describe(dplyr::filter(data,
                        !is.na(B_score) &
                          !is.na(Settingtype) &
                          !is.na(SettingRegion) &
                          !is.na(E_score) &
                          !is.na(SettingID))$E_score)
sd(dplyr::filter(data,
            !is.na(B_score) &
              !is.na(Settingtype) &
              !is.na(SettingRegion) &
              !is.na(E_score) &
              !is.na(SettingID))$E_score)


#And by treatment and control groups:
Hmisc::describe(dplyr::filter(data,  #treatment
                        treatment == 1 &
                          !is.na(B_score) &
                          !is.na(Settingtype) &
                          !is.na(SettingRegion) &
                          !is.na(E_score) &
                          !is.na(SettingID))$E_score)

sd(dplyr::filter(data, #treatment
            treatment == 1 &
              !is.na(B_score) &
              !is.na(Settingtype) &
              !is.na(SettingRegion) &
              !is.na(E_score) &
              !is.na(SettingID))$E_score)

Hmisc::describe(dplyr::filter(data,  #control
                        treatment == 0 &
                          !is.na(B_score) &
                          !is.na(Settingtype) &
                          !is.na(SettingRegion) &
                          !is.na(E_score) &
                          !is.na(SettingID))$E_score)
sd(dplyr::filter(data, #control
            treatment == 0 &
              !is.na(B_score) &
              !is.na(Settingtype) &
              !is.na(SettingRegion) &
              !is.na(E_score) &
              !is.na(SettingID))$E_score)
```

```
################################################
################################################
################################################



# Section B: Primary outcome model

### 1. Run the multi-level model.

pr_model <- lmer(E_score ~
                   treatment +
                   B_score +
                   SettingRegion +
                   Settingtype +
                   (1 | SettingID),
                 data=data, REML = FALSE)

summary(pr_model) #Produce the results

performance::icc(pr_model) #ICC

#########

### 2. Testing OLS assumptions

#Residual diagnostics: Testing normality of residuals OLS assumption
resid_pr <- resid(pr_model) #Create object which stores the residuals of the model
plot_resid_pr <- plot(density(resid_pr)) #Kernel density plot to explore normality
shapiro.test(resid_pr) #Shapiro-wilk test for normality. !FLAG potentially K-S test better with larger sample sizes!
(p<0.05 rejects H0 that data is normally distributed)
ks.test(resid_pr, "pnorm") #K-S test (p<0.05 rejects H0 that data is normally distributed

##Both of these tests reject H0

# QQ line
qq_line <- ggplot(data = data.frame(resid = resid_pr), aes(sample = resid)) +
  stat_qq() +
  stat_qq_line() +
  labs(title = "Q-Q Plot of Residuals", x = "Theoretical Quantiles", y = "Sample Quantiles") +
  theme_minimal()


#Plotting residual density against a normal distribution for comparison. NB this is using the residual objects created
from the primary model above
pr_resid_df <- as.data.frame(resid_pr)
sd_resid <- sd(resid_pr)
max_resid <- max(resid_pr)
min_resid <- min(resid_pr)

resid_pr_kd<- ggplot(pr_resid_df, aes(x = resid_pr)) +
  geom_density(aes(color = "Kernel density"), linewidth = 1) +
  stat_function(aes(color = "Normal density"),
                fun = function(x) dnorm(x, mean = 0, sd = sd_resid),
```

```
            linetype = "dotted", linewidth = 1) +
  theme_minimal() +
  scale_x_continuous(limits = c(min_resid, max_resid)) +
  scale_color_manual(name = "Legend", values = c("Kernel density" = "#000080", "Normal density" = "red")) +
  labs(
    title = "Basic Concepts - Residuals Density Plot",
    x = "Residuals",
    y = "Density"
  )

#Testing linearity and assumption that residual errors have a mean of 0
plot(pr_model, col = "red") #We want the line here to be horizontal and at 0


### Figures to export:
resid_pr_kd
qq_line

plots <- list(resid_pr_kd, qq_line)

### Export graphs to word

# Create a Word document
doc <- read_docx()

# Save each plot as an image and add to the Word document
for (i in seq_along(plots)) {
  # Export the plot to a temporary file
  file_name <- tempfile(fileext = ".png")
  ggplot2::ggsave(filename = file_name, plot = plots[[i]])

  # Add the plot to the Word document
  doc <- doc %>%
    body_add_par(value = paste("Plot", i), style = "heading 1") %>%
    body_add_img(src = file_name, width = 6, height = 4)
}

# Save the document
print(doc, target = "//Nevis/RE/Projects/In_Progress/022807.014_P2899_EEF_The_ONE/WIP/09. Impact/Graphs &
Summary Tables/SensitivityAnalysis3.docx")

##########


### 3.  Bootstrapping CIs and p-value
#+ The QQ residual plot, Shapiro-Wilk, and K-S test
#+ results suggest that the residuals are not normally distributed. We will
#+ therefore need to re-estimate CIs and p-values using bootstrapping.

set.seed(999)
boot_pr_model <- boot.pval::boot_summary(pr_model,
                                         type = "norm",
                                         method = NULL,
                                         conf.level = 0.95)
```

```
boot_pr_model

#####################



### 4.  Effect-size estimation

#Create Hedges g function.
hedges.g <- function(c, n, m, v, w){
  c/sqrt((((n - 1)*v)+((m - 1)* w))/(n + m-2))
}

#We use the output from the model above to calculate the effect size: (NB: This needs to be done for confidence intervals too!)
#We need to define the objects needed to be inputted into the Hedges G function created above (c, n, m, v, & w)

coefs <- data.frame(summary(pr_model)$coefficients) #create data frame of all coefficients from primary outcome model

c <- coefs["treatment", "Estimate"] #Extract the treatment coefficient

n <- nrow(dplyr::filter(data,
                        treatment == 1 &
                          !is.na(E_score) &
                          !is.na(B_score) &
                          !is.na(SettingRegion) &
                          !is.na(Settingtype) &
                          !is.na(SettingID))) #Number of individuals in treatment group in the model

m <- nrow(dplyr::filter(data,
                        treatment == 0 &
                          !is.na(E_score) &
                          !is.na(B_score) &
                          !is.na(SettingRegion) &
                          !is.na(Settingtype) &
                          !is.na(SettingID))) #Number of individuals in control group in the model

v <- var(dplyr::filter(data,
                       treatment == 1 &
                         !is.na(B_score) &
                         !is.na(SettingRegion) &
                         !is.na(Settingtype) &
                         !is.na(E_score) &
                         !is.na(SettingID))$E_score) #Variance in outcome among treatment group

w <- var(dplyr::filter(data,
                       treatment == 0 &
                         !is.na(B_score) &
                         !is.na(SettingRegion) &
                         !is.na(Settingtype) &
                         !is.na(E_score) &
```

```
                !is.na(SettingID))$E_score)  #Variance in outcome among control group

primary_effect_size <- hedges.g(c, n, m, v, w)
primary_effect_size

#LOWER CI EFFECT SIZE using lower 95% CI from the bootstrapped model:

#We redefine c to equal the treatment coefficient minus the SE created above
c <- boot_pr_model$Lower.bound[2]

pr_effect_size_low <- hedges.g(c, n, m, v, w) #Re-run hedges G on lower CI treatment coefficient
pr_effect_size_low #This is the lower confidence interval for the treatment effect on the following directions subtest.



#UPPER CI EFFECT SIZE using bootrapped 95% upper CI from the above model:

#We redefine c again to equal the treatment coefficient plus the SE created above. Remember c has already been
defined to be the lower confidence interval, we therefore need to add 2 SEs to it.
c <- boot_pr_model$Upper.bound[2]

pr_effect_size_high <- hedges.g(c, n, m, v, w) #Re-run hedges G on upper CI treatment coefficient
pr_effect_size_high #This is the upper confidence interval for the treatment effect on the following directions subtest.

## P-value
p_values <- boot_pr_model$p.value[2]
p_value <- round(p_values, digits = 2)


### 5. Creating primary analysis table (to paste output into report)

##+ This part of the code is to create a formatted output table which can just be copied
##+ into the table empty table in word

## Find missing numbers for model (number of obs with outcome but missing covariates)

# Treatment
n_mis <- nrow(dplyr::filter(data,
                        treatment == 1 &
                          !is.na(E_score) &
                          (is.na(B_score) |
                             is.na(SettingRegion) |
                             is.na(Settingtype) |
                             is.na(SettingID))))

# Control
m_mis <- nrow(dplyr::filter(data,
                        treatment == 0 &
                          !is.na(E_score) &
                          (is.na(B_score) |
                             is.na(SettingRegion) |
                             is.na(Settingtype) |
                             is.na(SettingID))))
```

```
# Paste in numbers (non-missing and missing) for table
num_t <- paste0(as.character(n), " (", as.character(n_mis), ")")
num_c <- paste0(as.character(m), " (", as.character(m_mis), ")")

# Total numbers
total_n <- n + m
total_n_t_c <- paste0(as.character(total_n), " (", n, ";", m, ")")

## Generate means and CIs

# Treatment
me_t <- t.test(dplyr::filter(data,
                        treatment == 1 &
                          !is.na(E_score) &
                          !is.na(B_score) &
                          !is.na(SettingRegion) &
                          !is.na(Settingtype) &
                          !is.na(SettingID))$E_score) #Run t-test on outcome in analytical sample

mean_t <- as.character(round(me_t$estimate, digits = 2)) #extract mean from t-test object (me), to 2dp and as
character to facilitate pasting in CIs.

me_ci_t <- as.character(round(me_t$conf.int[c(1, 2)], digits = 2)) #extract CIs (95%) from t-test object (me), to 2dp and
as character to facilitate pasting

mean_ci_t <- paste0(mean_t, " (", me_ci_t[1], " - ", me_ci_t[2], ")") #Paste CIs in brackets next to mean figure

# Control
me_c <- t.test(dplyr::filter(data,
                        treatment == 0 &
                          !is.na(E_score) &
                          !is.na(B_score) &
                          !is.na(SettingRegion) &
                          !is.na(Settingtype) &
                          !is.na(SettingID))$E_score)

mean_c <- as.character(round(me_c$estimate, digits = 2))
me_ci_c <- as.character(round(me_c$conf.int[c(1, 2)], digits = 2))
mean_ci_c <- paste0(mean_c, " (", me_ci_c[1], " - ", me_ci_c[2], ")")


## Effect size with CIs
ES_CI <- paste0(as.character(round(primary_effect_size, digits = 3)),
            " (", as.character(round(pr_effect_size_low, digits = 2)),
            " -- ",
            as.character(round(pr_effect_size_high, digits = 2)),
            ")")

## Create table

primary_analysis_table <- t(as.data.frame(c(num_t, mean_ci_t, num_c, mean_ci_c, total_n_t_c, ES_CI, p_value)))
colnames(primary_analysis_table) <- c("N (intervention)", "Mean (intervention", "N (control)", "Mean (control)", "Total
(T;C)", "Hedges g (95% CIs)", "p-value")
```

```
rownames(primary_analysis_table) <- "EYTN2 Numeracy"

## Display results - TABLE 1
View(primary_analysis_table)

############

### 6. Creating effect size estimation table. Most of the required fields have already been created

u_m <- as.character(round((mean(dplyr::filter(data, #Unadjusted difference in means
                                      treatment == 1 &
                                        !is.na(B_score) &
                                        !is.na(SettingRegion) &
                                        !is.na(Settingtype) &
                                        !is.na(SettingID))$E_score, na.rm = TRUE)) - (mean(dplyr::filter(data,
                                                                                        treatment == 0 &
                                                                                          !is.na(B_score) &
                                                                                          !is.na(SettingRegion) &
                                                                                          !is.na(Settingtype) &
                                                                                          !is.na(SettingID))$E_score, na.rm = TRUE)),
digits = 2))

c <- round(coefs["treatment", "Estimate"], digits = 2) #Adjusted difference in means

var_t <- round(v, digits = 2) #Variance of outcome in treatment

var_c <- round(w, digits = 2) #Variance of outcome in control

pooled_var <- round(((((n - 1)*v)+((m - 1)* w))/(n + m - 2)), digits = 2) #pooled variance



# Create table 2
pr_es_est <- t(as.data.frame(c(u_m, c, num_t, var_t, num_c, var_c, pooled_var)))
colnames(pr_es_est) <- c("Unadj. diff in means", "Adj diff in means", "N (treatment)", "Variance (treatment)", "N
(control)", "Variance (control)", "Pooled Var")
rownames(pr_es_est) <- "EYTN Numeracy Test"
View(pr_es_est) #@Bhavya- copy this into Table 2.


########

### Calculation of R-squared values for MDES calculations

# Null model
null model <- lmer(E_score ~ (1 | SettingID), data = data, REML = FALSE)

# Variance components for the null model
VarCorr(null_model)

# Variance components for the full model
VarCorr(pr_model)
```

```
within_var_null <- attr(VarCorr(null_model), "sc")^2
within_var_full <- attr(VarCorr(pr_model), "sc")^2

# Level 1 R^2 (Within-cluster variance explained)
R2_level1 <- (within_var_null - within_var_full) / within_var_null
R2_level1


between_var_null <- as.numeric(VarCorr(null_model)$SettingID[1])
between_var_full <- as.numeric(VarCorr(pr_model)$SettingID[1])

# Level 2 R^2 (Between-cluster variance explained)
R2_level2 <- (between_var_null - between_var_full) / between_var_null
R2_level2
###################################
## The ONE Sensitivity Analysis ##
###### Measurement Error 2 #######
###################################

##+ Conducting sensitivity analysis of the ONE secondary analysis with tests that did not follow the stopping
##+ rule correctly removed
##+ This will be done for the mixed methods and HTKS-HTKS model as this is where the majority of the
##+ incompleteness takes place
##+ After filtering out these observations, the procedure will follow the secondary analysis.

# Clear workspace
rm(list=ls())
set.seed(999) #NB this is the same seed set in the randomisation code at the start of the project so we shall retain it
here for consistency.

# Load packages
library("haven")
library("dplyr")
library("ggplot2")
library("lme4")
library("sjstats")
library("Hmisc")
library("performance")
library("lmtest")
library("eeptools")
library("lmerTest")
library("officer")
library("boot")
library("boot.pval")


# Load cleaned data
data <- read_dta("//Nevis/RE/Projects/In_Progress/022807.014_P2899_EEF_The_ONE/WIP/06. Data collection/02.
Endline/02. Data_NotMirrored/BaselineEndlineComplianceClean2.dta")
data <- filter(data, endlineonly != 1)


#Numbers of incomplete tests at baseline and endline
nrow(filter(data,
```

```
        E_HTKS_total_invalid != 0))

# 277 missing at endline including children without any HTKS test

nrow(filter(data,
        B_HTKS_total_invalid != 0))

# 539 missing at baseline including children without any HTKS test
# To ensure the missing data counts work properly I will replace the baseline value of test validity to NA if
# baseline HTKS score is missing

data$B_HTKS_total_invalid <- ifelse(is.na(data$B_HTKS_total), NA, data$B_HTKS_total_invalid)

### Filtering out incomplete tests
data <- dplyr::filter(data,
                E_HTKSValid == 1 &
                B_HTKSValid == 1)

######################################################################
##################### Mixed Measures Model ###################
######################################################################

# Section A: Descriptive statistics

#1. Histograms of outcomes

#Use this to check distribution of endline scores for our outcome

#We are mainly checking for ceiling effects, as flagged at baseline stage.

# In overall data
HTKS_ovr_hist <- hist(data$E_HTKS_total,
                main="Histogram of HTKS subtest at endline (overall sample - removing incomplete tests)",
                xlab = "HTKS Scores",
                xlim = c(0, 120),
                breaks = seq(0, 120, 10),
                xaxp = c(0, 120, 6))

#+In analytical data (the analytical data is defined as the data used in the multi-level regression model.
#+It includes the observations where we have complete data for all covariates included in the outcome model)
Sec1_model_hist <- hist(dplyr::filter(data,
                              !is.na(B_Corsi_total) &
                                !is.na(Settingtype) &
                                !is.na(SettingRegion) &
                                !is.na(SettingID))$E_HTKS_total,
                main="Histogram of EYTN subtest at endline (analytical sample - removing incomplete tests)",
                xlab = "Early numeracy scores",
                xlim = c(0, 120),
                breaks = seq(0, 120, 10),
                xaxp = c(0, 120, 6))

##No real floor or ceiling effects.
baseline_hist1 <- hist(data$B_Corsi_total,
                main="Histogram of Corsi Blocks subtest at endline (overall sample - removing incomplete tests)",
```

```
                  xlab = "Corsi Blocks Scores",
                  xlim = c(0, 15),
                  breaks = seq(0, 15, 3),
                  xaxp = c(0, 15, 5))

#+Some potential floor effects in the baseline control. Doesn't warrant additional sensistivity analysis,
#+but some descriptive stats below.

sd <- sd(data$B_Corsi_total, na.rm = T)

# Create df which is a copy of analytical data
data_sd <- dplyr::filter(data,
                    !is.na(B_Corsi_total)  &
                      !is.na(Settingtype) &
                      !is.na(SettingRegion) &
                      !is.na(SettingID))

data_sd <- data_sd %>%
  mutate(sd_dummy = 0)

data_sd$sd_dummy <- ifelse(data_sd$B_Corsi_total < sd, 1, 0)

mean(data_sd$sd_dummy)

# 39.56% within a standard deviation of the floor.

Hmisc::describe(dplyr::filter(data,
                      !is.na(E_HTKS_total))$E_HTKS_total)
sd(dplyr::filter(data,
            !is.na(E_HTKS_total))$E_HTKS_total)


##Analytical sample
Hmisc::describe(dplyr::filter(data,
                      !is.na(B_Corsi_total) &
                        !is.na(Settingtype) &
                        !is.na(SettingRegion) &
                        !is.na(E_HTKS_total) &
                        !is.na(SettingID))$E_HTKS_total)
sd(dplyr::filter(data,
            !is.na(B_Corsi_total) &
              !is.na(Settingtype) &
              !is.na(SettingRegion) &
              !is.na(E_HTKS_total) &
              !is.na(SettingID))$E_HTKS_total)


#And by treatment and control groups:
Hmisc::describe(dplyr::filter(data,  #treatment
                      treatment == 1 &
                        !is.na(B_Corsi_total)  &
                        !is.na(Settingtype) &
                        !is.na(SettingRegion) &
                        !is.na(E_HTKS_total) &
```

```
                        !is.na(SettingID))$E_HTKS_total)

sd(dplyr::filter(data, #treatment
                 treatment == 1 &
                   !is.na(B_Corsi_total) &
                   !is.na(Settingtype) &
                   !is.na(SettingRegion) &
                   !is.na(E_HTKS_total) &
                   !is.na(SettingID))$E_HTKS_total)

Hmisc::describe(dplyr::filter(data,  #control
                          treatment == 0 &
                            !is.na(B_Corsi_total)  &
                            !is.na(Settingtype) &
                            !is.na(SettingRegion) &
                            !is.na(E_HTKS_total) &
                            !is.na(SettingID))$E_HTKS_total)
sd(dplyr::filter(data, #control
                 treatment == 0 &
                   !is.na(B_Corsi_total) &
                   !is.na(Settingtype) &
                   !is.na(SettingRegion) &
                   !is.na(E_HTKS_total) &
                   !is.na(SettingID))$E_HTKS_total)

##################################################
##################################################
##################################################

# Section B: Secondary outcome model 1

### 1. Run the multi-level model.

sec1_model <- lmer(E_HTKS_total ~
                     treatment +
                     B_Corsi_total  +
                     SettingRegion +
                     Settingtype +
                     (1 | SettingID),
                   data=data, REML = FALSE)

summary(sec1_model) #Produce the results

performance::icc(sec1_model) #ICC

#########

### 2. Testing OLS assumptions

#Residual diagnostics: Testing normality of residuals OLS assumption
resid_pr <- resid(sec1_model) #Create object which stores the residuals of the model
plot_resid_pr <- plot(density(resid_pr)) #Kernel density plot to explore normality
shapiro.test(resid_pr) #Shapiro-wilk test for normality. !FLAG potentially K-S test better with larger sample sizes!
(p<0.05 rejects H0 that data is normally distributed)
```

```
ks.test(resid_pr, "pnorm") #K-S test (p<0.05 rejects H0 that data is normally distributed)

##Both of these tests reject H0

# QQ line
qq_line <- ggplot(data = data.frame(resid = resid_pr), aes(sample = resid)) +
  stat_qq() +
  stat_qq_line() +
  labs(title = "Q-Q Plot of Residuals", x = "Theoretical Quantiles", y = "Sample Quantiles") +
  theme_minimal()


#Plotting residual density against a normal distribution for comparison. NB this is using the residual objects created
from the secondary model above
sec1_resid_df <- as.data.frame(resid_pr)
sd_resid <- sd(resid_pr)
max_resid <- max(resid_pr)
min_resid <- min(resid_pr)

resid_sec1_kd<- ggplot(sec1_resid_df, aes(x = resid_pr)) +
  geom_density(aes(color = "Kernel density"), linewidth = 1) +
  stat_function(aes(color = "Normal density"),
                fun = function(x) dnorm(x, mean = 0, sd = sd_resid),
                linetype = "dotted", linewidth = 1) +
  theme_minimal() +
  scale_x_continuous(limits = c(min_resid, max_resid)) +
  scale_color_manual(name = "Legend", values = c("Kernel density" = "#000080", "Normal density" = "red")) +
  labs(
    title = "Executive Function Model 1 - Residuals Density Plot",
    x = "Residuals",
    y = "Density"
  )

#Testing linearity and assumption that residual errors have a mean of 0
plot(sec1_model, col = "red") #We want the line here to be horizontal and at 0

### Figures to export:
resid_sec1_kd
qq_line

plots <- list(resid_sec1_kd, qq_line)

### Export graphs to word

# Create a Word document
doc <- read_docx()

# Save each plot as an image and add to the Word document
for (i in seq_along(plots)) {
  # Export the plot to a temporary file
  file_name <- tempfile(fileext = ".png")
  ggplot2::ggsave(filename = file_name, plot = plots[[i]])

  # Add the plot to the Word document
```

```
  doc <- doc %>%
    body_add_par(value = paste("Plot", i), style = "heading 1") %>%
    body_add_img(src = file_name, width = 6, height = 4)
}

# Save the document
print(doc, target = "//Nevis/RE/Projects/In_Progress/022807.014_P2899_EEF_The_ONE/WIP/09. Impact/Graphs &
Summary Tables/Sensitivity4_1.docx")

### 3.  Bootstrapping CIs and p-value
#+ The plots and normality test results suggest
#+ that the residuals are not normally distributed. We will
#+ therefore need to re-estimate CIs and p-values using bootstrapping.

set.seed(999)
boot_sec1_model <- boot.pval::boot_summary(sec1_model,
                                   type = "norm",
                                   method = NULL,
                                   conf.level = 0.95)
boot_sec1_model

######################




### 4.  Effect-size estimation

hedges.g <- function(c, n, m, v, w){
  c/sqrt((((n - 1)*v)+((m - 1)* w))/n + (m-2))
}

#We use the output from the model above to calculate the effect size: (NB: This needs to be done for confidence
intervals too!)
#We need to define the objects needed to be inputted into the Hedges G function created above (c, n, m, v, & w)

coefs <- data.frame(summary(sec1_model)$coefficients) #create data frame of all coefficients from secondary
outcome model

c <- coefs["treatment", "Estimate"] #Extract the treatment coefficient

n <- nrow(dplyr::filter(data,
                        treatment == 1 &
                          !is.na(E_HTKS_total) &
                          !is.na(B_Corsi_total)  &
                          !is.na(SettingRegion) &
                          !is.na(Settingtype) &
                          !is.na(SettingID))) #Number of individuals in treatment group in the model

m <- nrow(dplyr::filter(data,
                        treatment == 0 &
                          !is.na(E_HTKS_total) &
                          !is.na(B_Corsi_total)  &
                          !is.na(SettingRegion) &
```

```
                  !is.na(Settingtype) &
                  !is.na(SettingID))) #Number of individuals in control group in the model

v <- var(dplyr::filter(data,
                  treatment == 1 &
                   !is.na(B_Corsi_total) &
                   !is.na(SettingRegion) &
                   !is.na(Settingtype) &
                   !is.na(E_HTKS_total) &
                   !is.na(SettingID))$E_HTKS_total) #Variance in outcome among treatment group

w <- var(dplyr::filter(data,
                  treatment == 0 &
                   !is.na(B_Corsi_total) &
                   !is.na(SettingRegion) &
                   !is.na(Settingtype) &
                   !is.na(E_HTKS_total) &
                   !is.na(SettingID))$E_HTKS_total)  #Variance in outcome among control group

secondary1_effect_size <- hedges.g(c, n, m, v, w)
secondary1_effect_size

#LOWER CI EFFECT SIZE using lower 95% CI from the bootstrapped model:

#We redefine c to equal the treatment coefficient minus the SE created above
c <- boot_sec1_model$Lower.bound[2]

sec1_effect_size_low <- hedges.g(c, n, m, v, w) #Re-run hedges G on lower CI treatment coefficient
sec1_effect_size_low #This is the lower confidence interval for the treatment effect on the following directions subtest.



#UPPER CI EFFECT SIZE using bootrapped 95% upper CI from the above model:

#We redefine c again to equal the treatment coefficient plus the SE created above. Remember c has already been
defined to be the lower confidence interval, we therefore need to add 2 SEs to it.
c <- boot_sec1_model$Upper.bound[2]

sec1_effect_size_high <- hedges.g(c, n, m, v, w) #Re-run hedges G on upper CI treatment coefficient
sec1_effect_size_high #This is the upper confidence interval for the treatment effect on the following directions
subtest.

## P-value
p_values <- boot_sec1_model$p.value[2]
p_value <- round(p_values, digits = 2)


### 5. Creating secondary analysis table (to paste output into report)

##+ This part of the code is to create a formatted output table which can just be copied
##+ into the table empty table in word

## Find missing numbers for model (number of obs with outcome but missing covariates)
```

```
# Treatment
n_mis <- nrow(dplyr::filter(data,
                        treatment == 1 &
                         !is.na(E_HTKS_total) &
                         (is.na(B_Corsi_total) |
                            is.na(SettingRegion) |
                            is.na(Settingtype) |
                            is.na(SettingID))))

# Control
m_mis <- nrow(dplyr::filter(data,
                        treatment == 0 &
                         !is.na(E_HTKS_total) &
                         (is.na(B_Corsi_total) |
                            is.na(SettingRegion) |
                            is.na(Settingtype) |
                            is.na(SettingID))))

# Paste in numbers (non-missing and missing) for table
num_t <- paste0(as.character(n), " (", as.character(n_mis), ")")
num_c <- paste0(as.character(m), " (", as.character(m_mis), ")")

# Total numbers
total_n <- n + m
total_n_t_c <- paste0(as.character(total_n), " (", n, ";", m, ")")

## Generate means and CIs

# Treatment
me_t <- t.test(dplyr::filter(data,
                         treatment == 1 &
                          !is.na(E_HTKS_total) &
                          !is.na(B_Corsi_total)  &
                          !is.na(SettingRegion) &
                          !is.na(Settingtype) &
                          !is.na(SettingID))$E_HTKS_total) #Run t-test on outcome in analytical sample

mean_t <- as.character(round(me_t$estimate, digits = 2)) #extract mean from t-test object (me), to 2dp and as
character to facilitate pasting in CIs.

me_ci_t <- as.character(round(me_t$conf.int[c(1, 2)], digits = 2)) #extract CIs (95%) from t-test object (me), to 2dp and
as character to facilitate pasting

mean_ci_t <- paste0(mean_t, " (", me_ci_t[1], " - ", me_ci_t[2], ")") #Paste CIs in brackets next to mean figure

# Control
me_c <- t.test(dplyr::filter(data,
                         treatment == 0 &
                          !is.na(E_HTKS_total) &
                          !is.na(B_Corsi_total)  &
                          !is.na(SettingRegion) &
                          !is.na(Settingtype) &
                          !is.na(SettingID))$E_HTKS_total)
```

```
mean_c <- as.character(round(me_c$estimate, digits = 2))
me_ci_c <- as.character(round(me_c$conf.int[c(1, 2)], digits = 2))
mean_ci_c <- paste0(mean_c, " (", me_ci_c[1], " - ", me_ci_c[2], ")")


## Effect size with CIs
ES_CI <- paste0(as.character(round(secondary1_effect_size, digits = 3)),
                " (", as.character(round(sec1_effect_size_low, digits = 2)),
                " -- ",
                as.character(round(sec1_effect_size_high, digits = 2)),
                ")")

## Create table

secondary_analysis_table_1 <- t(as.data.frame(c(num_t, mean_ci_t, num_c, mean_ci_c, total_n_t_c, ES_CI,
p_value)))
colnames(secondary_analysis_table_1) <- c("N (intervention)", "Mean (intervention", "N (control)", "Mean (control)",
"Total (T;C)", "Hedges g (95% CIs)", "p-value")

rownames(secondary_analysis_table_1) <- "HTKS Executive Function Test"

## Display results - TABLE 1
View(secondary_analysis_table_1)

############

### 6. Creating effect size estimation table. Most of the required fields have already been created

u_m <- as.character(round((mean(dplyr::filter(data, #Unadjusted difference in means
                                      treatment == 1 &
                                        !is.na(B_Corsi_total) &
                                        !is.na(SettingRegion) &
                                        !is.na(Settingtype) &
                                        !is.na(SettingID))$E_HTKS_total, na.rm = TRUE)) - (mean(dplyr::filter(data,
                                                                                                             treatment == 0 &
                                                                                                               !is.na(B_Corsi_total) &
                                                                                                               !is.na(SettingRegion) &
                                                                                                               !is.na(Settingtype) &
                                                                                                               !is.na(SettingID))$E_HTKS_total, na.rm =
TRUE)), digits = 2))

c <- round(coefs["treatment", "Estimate"], digits = 2) #Adjusted difference in means

var_t <- round(v, digits = 2) #Variance of outcome in treatment

var_c <- round(w, digits = 2) #Variance of outcome in control

pooled_sd <- round(sqrt((((n - 1)*v)+((m - 1)* w))/(n + (m-2))), digits = 2) #pooled standard deviation



# Create table 2
sec1_es_est <- t(as.data.frame(c(u_m, c, num_t, var_t, num_c, var_c, pooled_sd)))
```

```
colnames(sec1_es_est) <- c("Unadj. diff in means", "Adj diff in means", "N (treatment)", "Variance (treatment)", "N (control)", "Variance (control)", "Pooled SD")
rownames(sec1_es_est) <- "HTKS Executive Function Test"
View(sec1_es_est)


########

### Calculation of R-squared values for MDES calculations

# Null model
null_model <- lmer(E_HTKS_total ~ (1 | SettingID), data = data, REML = FALSE)

# Variance components for the null model
VarCorr(null_model)

# Variance components for the full model
VarCorr(sec1_model)


within_var_null <- attr(VarCorr(null_model), "sc")^2
within_var_full <- attr(VarCorr(sec1_model), "sc")^2

# Level 1 R^2 (Within-cluster variance explained)
R2_level1 <- (within_var_null - within_var_full) / within_var_null
R2_level1


between_var_null <- as.numeric(VarCorr(null_model)$SettingID[1])
between_var_full <- as.numeric(VarCorr(sec1_model)$SettingID[1])

# Level 2 R^2 (Between-cluster variance explained)
R2_level2 <- (between_var_null - between_var_full) / between_var_null
R2_level2


##########################################################

##########################################################
#################### HTKS-R Model ########################
##########################################################

# Section A: Descriptive statistics

#1. Histograms of outcomes

#Use this to check distribution of endline scores for our outcome

#We are mainly checking for ceiling effects, as flagged at baseline stage.

# Already have histogram for overall data

#+In analytical data (the analytical data is defined as the data used in the multi-level regression model.
#+It includes the observations where we have complete data for all covariates included in the outcome model)
```

```
Sec2_model_hist <- hist(dplyr::filter(data,
                                      !is.na(B_HTKS_total) &
                                        !is.na(Settingtype) &
                                        !is.na(SettingRegion) &
                                        !is.na(SettingID))$E_HTKS_total,
                        main="Histogram of HTKS subtest at endline (analytical sample - accounting for measurement error)",
                        xlab = "HTKS scores",
                        xlim = c(0, 120),
                        breaks = seq(0, 120, 10),
                        xaxp = c(0, 120, 6))

##No real floor or ceiling effects.

Hmisc::describe(dplyr::filter(data,
                              !is.na(E_HTKS_total))$E_HTKS_total)
sd(dplyr::filter(data,
                 !is.na(E_HTKS_total))$E_HTKS_total)


##Analytical sample
Hmisc::describe(dplyr::filter(data,
                              !is.na(B_HTKS_total) &
                                !is.na(Settingtype) &
                                !is.na(SettingRegion) &
                                !is.na(E_HTKS_total) &
                                !is.na(SettingID))$E_HTKS_total)
sd(dplyr::filter(data,
                 !is.na(B_HTKS_total) &
                   !is.na(Settingtype) &
                   !is.na(SettingRegion) &
                   !is.na(E_HTKS_total) &
                   !is.na(SettingID))$E_HTKS_total)


#And by treatment and control groups:
Hmisc::describe(dplyr::filter(data,  #treatment
                              treatment == 1 &
                                !is.na(B_HTKS_total) &
                                !is.na(Settingtype) &
                                !is.na(SettingRegion) &
                                !is.na(E_HTKS_total) &
                                !is.na(SettingID))$E_HTKS_total)

sd(dplyr::filter(data, #treatment
                 treatment == 1 &
                   !is.na(B_HTKS_total) &
                   !is.na(Settingtype) &
                   !is.na(SettingRegion) &
                   !is.na(E_HTKS_total) &
                   !is.na(SettingID))$E_HTKS_total)

Hmisc::describe(dplyr::filter(data,  #control
                              treatment == 0 &
                                !is.na(B_HTKS_total) &
```

```
                        !is.na(Settingtype) &
                        !is.na(SettingRegion) &
                        !is.na(E_HTKS_total) &
                        !is.na(SettingID))$E_HTKS_total)
sd(dplyr::filter(data, #control
                 treatment == 0 &
                   !is.na(B_HTKS_total) &
                   !is.na(Settingtype) &
                   !is.na(SettingRegion) &
                   !is.na(E_HTKS_total) &
                   !is.na(SettingID))$E_HTKS_total)

##################################################
##################################################
##################################################

# Section B: Secondary outcome model 1

### 1. Run the multi-level model.

sec2_model <- lmer(E_HTKS_total ~
                     treatment +
                     B_HTKS_total +
                     SettingRegion +
                     Settingtype +
                     (1 | SettingID),
                   data=data, REML = FALSE)

summary(sec2_model) #Produce the results

performance::icc(sec2_model) #ICC

#########

### 2. Testing OLS assumptions

#Residual diagnostics: Testing normality of residuals OLS assumption
resid_sec2 <- resid(sec2_model) #Create object which stores the residuals of the model
plot_resid_sec2 <- plot(density(resid_sec2)) #Kernel density plot to explore normality
shapiro.test(resid_sec2) #Shapiro-wilk test for normality. !FLAG potentially K-S test better with larger sample sizes!
(p<0.05 rejects H0 that data is normally distributed)
ks.test(resid_sec2, "pnorm") #K-S test (p<0.05 rejects H0 that data is normally distributed)

##Both of these tests reject H0

# QQ line
qq_line2 <- ggplot(data = data.frame(resid = resid_sec2), aes(sample = resid)) +
  stat_qq() +
  stat_qq_line() +
  labs(title = "Q-Q Plot of Residuals", x = "Theoretical Quantiles", y = "Sample Quantiles") +
  theme_minimal()
```

```
#Plotting residual density against a normal distribution for comparison. NB this is using the residual objects created from the secondary model above
sec2_resid_df <- as.data.frame(resid_sec2)
sd_resid_sec2 <- sd(resid_sec2)
max_resid_sec2 <- max(resid_sec2)
min_resid_sec2 <- min(resid_sec2)

resid_sec2_kd<- ggplot(sec2_resid_df, aes(x = resid_sec2)) +
  geom_density(aes(color = "Kernel density"), linewidth = 1) +
  stat_function(aes(color = "Normal density"),
                fun = function(x) dnorm(x, mean = 0, sd = sd_resid_sec2),
                linetype = "dotted", linewidth = 1) +
  theme_minimal() +
  scale_x_continuous(limits = c(min_resid_sec2, max_resid_sec2)) +
  scale_color_manual(name = "Legend", values = c("Kernel density" = "#000080", "Normal density" = "red")) +
  labs(
    title = "Executive Function Model 2 - Residuals Density Plot",
    x = "Residuals",
    y = "Density"
  )

#Testing linearity and assumption that residual errors have a mean of 0
plot(sec2_model, col = "red") #We want the line here to be horizontal and at 0

### Figures to export:
resid_sec2_kd
qq_line2

plots <- list(resid_sec2_kd, qq_line2)

### Export graphs to word

# Create a Word document
doc <- read_docx()

# Save each plot as an image and add to the Word document
for (i in seq_along(plots)) {
  # Export the plot to a temporary file
  file_name <- tempfile(fileext = ".png")
  ggplot2::ggsave(filename = file_name, plot = plots[[i]])

  # Add the plot to the Word document
  doc <- doc %>%
    body_add_par(value = paste("Plot", i), style = "heading 1") %>%
    body_add_img(src = file_name, width = 6, height = 4)
}

# Save the document
print(doc, target = "//Nevis/RE/Projects/In_Progress/022807.014_P2899_EEF_The_ONE/WIP/09. Impact/Graphs & Summary Tables/SensitivityAnalysis4_2.docx")

### 3.  Bootstrapping CIs and p-value
#+ The plots and normality test results suggest
#+ that the residuals are not normally distributed. We will
```

```
#+ therefore need to re-estimate CIs and p-values using bootstrapping.

set.seed(999)
boot_sec2_model <- boot.pval::boot_summary(sec2_model,
                                  type = "norm",
                                  method = NULL,
                                  conf.level = 0.95)
boot_sec2_model

######################




### 4.  Effect-size estimation

coefs2 <- data.frame(summary(sec2_model)$coefficients) #create data frame of all coefficients from secondary
outcome model

c <- coefs2["treatment", "Estimate"] #Extract the treatment coefficient

n <- nrow(dplyr::filter(data,
                  treatment == 1 &
                    !is.na(E_HTKS_total) &
                    !is.na(B_HTKS_total) &
                    !is.na(SettingRegion) &
                    !is.na(Settingtype) &
                    !is.na(SettingID))) #Number of individuals in treatment group in the model

m <- nrow(dplyr::filter(data,
                  treatment == 0 &
                    !is.na(E_HTKS_total) &
                    !is.na(B_HTKS_total) &
                    !is.na(SettingRegion) &
                    !is.na(Settingtype) &
                    !is.na(SettingID))) #Number of individuals in control group in the model

v <- var(dplyr::filter(data,
                 treatment == 1 &
                   !is.na(B_HTKS_total) &
                   !is.na(SettingRegion) &
                   !is.na(Settingtype) &
                   !is.na(E_HTKS_total) &
                   !is.na(SettingID))$E_HTKS_total) #Variance in outcome among treatment group

w <- var(dplyr::filter(data,
                 treatment == 0 &
                   !is.na(B_HTKS_total) &
                   !is.na(SettingRegion) &
                   !is.na(Settingtype) &
                   !is.na(E_HTKS_total) &
                   !is.na(SettingID))$E_HTKS_total)  #Variance in outcome among control group

secondary2_effect_size <- hedges.g(c, n, m, v, w)
```

```
secondary2_effect_size

#LOWER CI EFFECT SIZE using lower 95% CI from the bootstrapped model:

#We redefine c to equal the treatment coefficient minus the SE created above
c <- boot_sec2_model$Lower.bound[2]

sec2_effect_size_low <- hedges.g(c, n, m, v, w) #Re-run hedges G on lower CI treatment coefficient
sec2_effect_size_low #This is the lower confidence interval for the treatment effect on the following directions subtest.



#UPPER CI EFFECT SIZE using bootrapped 95% upper CI from the above model:

#We redefine c again to equal the treatment coefficient plus the SE created above. Remember c has already been
defined to be the lower confidence interval, we therefore need to add 2 SEs to it.
c <- boot_sec2_model$Upper.bound[2]

sec2_effect_size_high <- hedges.g(c, n, m, v, w) #Re-run hedges G on upper CI treatment coefficient
sec2_effect_size_high #This is the upper confidence interval for the treatment effect on the following directions
subtest.

## P-value
p_values <- boot_sec2_model$p.value[2]
p_value <- round(p_values, digits = 2)


### 5. Creating secondary analysis table (to paste output into report)

##+ This part of the code is to create a formatted output table which can just be copied
##+ into the table empty table in word

## Find missing numbers for model (number of obs with outcome but missing covariates)

# Treatment
n_mis <- nrow(dplyr::filter(data,
                            treatment == 1 &
                              !is.na(E_HTKS_total) &
                              (is.na(B_HTKS_total) |
                                 is.na(SettingRegion) |
                                 is.na(Settingtype) |
                                 is.na(SettingID))))

# Control
m_mis <- nrow(dplyr::filter(data,
                            treatment == 0 &
                              !is.na(E_HTKS_total) &
                              (is.na(B_HTKS_total) |
                                 is.na(SettingRegion) |
                                 is.na(Settingtype) |
                                 is.na(SettingID))))

# Paste in numbers (non-missing and missing) for table
num_t <- paste0(as.character(n), " (", as.character(n_mis), ")")
```

```
num_c <- paste0(as.character(m), " (", as.character(m_mis), ")")

# Total numbers
total_n <- n + m
total_n_t_c <- paste0(as.character(total_n), " (", n, ";", m, ")")

## Generate means and CIs

# Treatment
me_t <- t.test(dplyr::filter(data,
                        treatment == 1 &
                          !is.na(E_HTKS_total) &
                          !is.na(B_HTKS_total) &
                          !is.na(SettingRegion) &
                          !is.na(Settingtype) &
                          !is.na(SettingID))$E_HTKS_total) #Run t-test on outcome in analytical sample

mean_t <- as.character(round(me_t$estimate, digits = 2)) #extract mean from t-test object (me), to 2dp and as character to facilitate pasting in CIs.

me_ci_t <- as.character(round(me_t$conf.int[c(1, 2)], digits = 2)) #extract CIs (95%) from t-test object (me), to 2dp and as character to facilitate pasting

mean_ci_t <- paste0(mean_t, " (", me_ci_t[1], " - ", me_ci_t[2], ")") #Paste CIs in brackets next to mean figure

# Control
me_c <- t.test(dplyr::filter(data,
                        treatment == 0 &
                          !is.na(E_HTKS_total) &
                          !is.na(B_HTKS_total) &
                          !is.na(SettingRegion) &
                          !is.na(Settingtype) &
                          !is.na(SettingID))$E_HTKS_total)

mean_c <- as.character(round(me_c$estimate, digits = 2))
me_ci_c <- as.character(round(me_c$conf.int[c(1, 2)], digits = 2))
mean_ci_c <- paste0(mean_c, " (", me_ci_c[1], " - ", me_ci_c[2], ")")


## Effect size with CIs
ES_CI <- paste0(as.character(round(secondary2_effect_size, digits = 3)),
            " (", as.character(round(sec2_effect_size_low, digits = 2)),
            " -- ",
            as.character(round(sec2_effect_size_high, digits = 2)),
            ")")

## Create table

secondary_analysis_table_2 <- t(as.data.frame(c(num_t, mean_ci_t, num_c, mean_ci_c, total_n_t_c, ES_CI,
p_value)))
colnames(secondary_analysis_table_2) <- c("N (intervention)", "Mean (intervention", "N (control)", "Mean (control)",
"Total (T;C)", "Hedges g (95% CIs)", "p-value")

rownames(secondary_analysis_table_2) <- "HTKS Executive Function Test"
```

```
## Display results - TABLE 1
View(secondary_analysis_table_2)

###########

### 6. Creating effect size estimation table. Most of the required fields have already been created

u_m <- as.character(round((mean(dplyr::filter(data, #Unadjusted difference in means
                                      treatment == 1 &
                                        !is.na(B_HTKS_total) &
                                        !is.na(SettingRegion) &
                                        !is.na(Settingtype) &
                                        !is.na(SettingID))$E_HTKS_total, na.rm = TRUE)) - (mean(dplyr::filter(data,
                                                                                                             treatment == 0 &
                                                                                                               !is.na(B_HTKS_total) &
                                                                                                               !is.na(SettingRegion) &
                                                                                                               !is.na(Settingtype) &
                                                                                                               !is.na(SettingID))$E_HTKS_total, na.rm =
TRUE)), digits = 2))

c <- round(coefs2["treatment", "Estimate"], digits = 2) #Adjusted difference in means

var_t <- round(v, digits = 2) #Variance of outcome in treatment

var_c <- round(w, digits = 2) #Variance of outcome in control

pooled_sd <- round(sqrt((((n - 1)*v)+((m - 1)* w))/(n + (m-2))), digits = 2) #pooled standard deviation



# Create table 2
sec2_es_est <- t(as.data.frame(c(u_m, c, num_t, var_t, num_c, var_c, pooled_sd)))
colnames(sec2_es_est) <- c("Unadj. diff in means", "Adj diff in means", "N (treatment)", "Variance (treatment)", "N
(control)", "Variance (control)", "Pooled SD")
rownames(sec2_es_est) <- "HTKS Executive Function Test"
View(sec2_es_est)


#########################################################
###################################
## The ONE Sensitivity Analysis ##
######### Endline Only ###########
###################################

##+ Conducting sensitivity analysis of the ONE primary results while excluding the baseline measure of numeracy
##+ ability from the analysis. This accounts for data collection failure at baseline.
##+ The procedure will follow that of the primary analysis.

# Clear workspace
rm(list=ls())
set.seed(999) #NB this is the same seed set in the randomisation code at the start of the project so we shall retain it
here for consistency.
```

```
# Load packages
library("haven")
library("dplyr")
library("ggplot2")
library("lme4")
library("sjstats")
library("Hmisc")
library("performance")
library("lmtest")
library("eeptools")
library("lmerTest")
library("officer")
library("boot")
library("boot.pval")


# Load cleaned data
data <- read_dta("//Nevis/RE/Projects/In_Progress/022807.014_P2899_EEF_The_ONE/WIP/06. Data collection/02.
Endline/02. Data_NotMirrored/BaselineEndlineComplianceClean2.dta")
data <- filter(data, endlineonly != 1)


# Section A: Descriptive statistics


#1. Histograms of outcomes

#Use this to check distribution of endline and baseline scores for our outcome

#We are mainly checking for ceiling and floor effects

# Overall data histogram already available in folder

#+In analytical data (the analytical data is defined as the data used in the multi-level regression model.
#+It includes the observations where we have complete data for all covariates included in the outcome model)
EYTN_model_hist <- hist(dplyr::filter(data,
                                      !is.na(Settingtype) &
                                      !is.na(SettingRegion) &
                                      !is.na(SettingID))$E_score,
                        main="Histogram of EYTN subtest at endline (endline only model sample)",
                        xlab = "Early numeracy scores",
                        xlim = c(0, 120),
                        breaks = seq(0, 120, 10),
                        xaxp = c(0, 120, 6))

#2. Means, SDs, Min and Max (for whole sample and analytical sample)

#Overall sample
Hmisc::describe(dplyr::filter(data,
                              !is.na(E_score))$E_score)
sd(dplyr::filter(data,
                 !is.na(E_score))$E_score)
```

```
##Analytical sample
Hmisc::describe(dplyr::filter(data,
                              !is.na(Settingtype) &
                              !is.na(SettingRegion) &
                              !is.na(E_score) &
                              !is.na(SettingID))$E_score)
sd(dplyr::filter(data,
                 !is.na(Settingtype) &
                 !is.na(SettingRegion) &
                 !is.na(E_score) &
                 !is.na(SettingID))$E_score)
##+As I suspected, the overall and analytical sample are the same in this model--the discrepancy in the main primary analysis comes from
##+children having results at baseline but not endline

#And by treatment and control groups:
Hmisc::describe(dplyr::filter(data,  #treatment
                              treatment == 1 &
                                !is.na(Settingtype) &
                                !is.na(SettingRegion) &
                                !is.na(E_score) &
                                !is.na(SettingID))$E_score)

sd(dplyr::filter(data, #treatment
                 treatment == 1 &
                   !is.na(Settingtype) &
                   !is.na(SettingRegion) &
                   !is.na(E_score) &
                   !is.na(SettingID))$E_score)

Hmisc::describe(dplyr::filter(data,  #control
                              treatment == 0 &
                                !is.na(Settingtype) &
                                !is.na(SettingRegion) &
                                !is.na(E_score) &
                                !is.na(SettingID))$E_score)
sd(dplyr::filter(data, #control
                 treatment == 0 &
                   !is.na(Settingtype) &
                   !is.na(SettingRegion) &
                   !is.na(E_score) &
                   !is.na(SettingID))$E_score)

##Seems unlikely that there will be any significant effects in this model

##################################################
##################################################
##################################################

# Section B: Primary outcome model

### 1. Run the multi-level model.

pr_model <- lmer(E_score ~
```

```
              treatment +
              SettingRegion +
              Settingtype +
              (1 | SettingID),
            data=data, REML = FALSE)

summary(pr_model) #Produce the results

performance::icc(pr_model) #ICC

#########

### 2. Test OLS model assumptions
resid_pr <- resid(pr_model)
shapiro.test(resid_pr) #Shapiro-wilk test for normality. !FLAG potentially K-S test better with larger sample sizes!
(p<0.05 rejects H0 that data is normally distributed)
ks.test(resid_pr, "pnorm") #K-S test (p<0.05 rejects H0 that data is normally distributed


## These tests both reject normality
pr_resid_df <- as.data.frame(resid_pr)

qq_line <- ggplot(data = data.frame(resid = resid_pr), aes(sample = resid)) +
  stat_qq() +
  stat_qq_line() +
  labs(title = "Q-Q Plot of Residuals", x = "Theoretical Quantiles", y = "Sample Quantiles") +
  theme_minimal()

sd_resid <- sd(resid_pr)
max_resid <- max(resid_pr)
min_resid <- min(resid_pr)

resid_pr_kd<- ggplot(pr_resid_df, aes(x = resid_pr)) +
  geom_density(aes(color = "Kernel density"), linewidth = 1) +
  stat_function(aes(color = "Normal density"),
                fun = function(x) dnorm(x, mean = 0, sd = sd_resid),
                linetype = "dotted", linewidth = 1) +
  theme_minimal() +
  scale_x_continuous(limits = c(min_resid, max_resid)) +
  scale_color_manual(name = "Legend", values = c("Kernel density" = "#000080", "Normal density" = "red")) +
  labs(
    title = "Basic Concepts - Residuals Density Plot",
    x = "Residuals",
    y = "Density"
  )

qq_line
resid_pr_kd

plots <- list(resid_pr_kd, qq_line)

### Export graphs to word
# Create a Word document
doc <- read_docx()
```

```
# Save each plot as an image and add to the Word document
for (i in seq_along(plots)) {
  # Export the plot to a temporary file
  file_name <- tempfile(fileext = ".png")
  ggplot2::ggsave(filename = file_name, plot = plots[[i]])

  # Add the plot to the Word document
  doc <- doc %>%
    body_add_par(value = paste("Plot", i), style = "heading 1") %>%
    body_add_img(src = file_name, width = 6, height = 4)
}

# Save the document
print(doc, target = "//Nevis/RE/Projects/In_Progress/022807.014_P2899_EEF_The_ONE/WIP/09. Impact/Graphs &
Summary Tables/SensitivityAnalysis5.docx")

##########

### 3.  Bootstrapping CIs and p-value
#+ The QQ residual plot, Shapiro-Wilk, and K-S test
#+ results suggest that the residuals are not normally distributed. We will
#+ therefore need to re-estimate CIs and p-values using bootstrapping.

set.seed(999)
boot_pr_model <- boot.pval::boot_summary(pr_model,
                                type = "norm",
                                method = NULL,
                                conf.level = 0.95)
boot_pr_model


######################

### 4.  Effect-size estimation

#Create Hedges g function.
hedges.g <- function(c, n, m, v, w){
  c/sqrt((((n - 1)*v)+((m - 1)* w))/(n + m-2))
}

coefs <- data.frame(summary(pr_model)$coefficients) #create data frame of all coefficients from primary outcome
model

c <- coefs["treatment", "Estimate"] #Extract the treatment coefficient

n <- nrow(dplyr::filter(data,
                        treatment == 1 &
                          !is.na(E_score) &
                          !is.na(SettingRegion) &
                          !is.na(Settingtype) &
                          !is.na(SettingID))) #Number of individuals in treatment group in the model

m <- nrow(dplyr::filter(data,
```

```
                 treatment == 0 &
                  !is.na(E_score) &
                  !is.na(SettingRegion) &
                  !is.na(Settingtype) &
                  !is.na(SettingID))) #Number of individuals in control group in the model

v <- var(dplyr::filter(data,
                  treatment == 1 &
                  !is.na(SettingRegion) &
                  !is.na(Settingtype) &
                  !is.na(E_score) &
                  !is.na(SettingID))$E_score) #Variance in outcome among treatment group

w <- var(dplyr::filter(data,
                  treatment == 0 &
                  !is.na(SettingRegion) &
                  !is.na(Settingtype) &
                  !is.na(E_score) &
                  !is.na(SettingID))$E_score)  #Variance in outcome among control group

primary_effect_size <- hedges.g(c, n, m, v, w)
primary_effect_size

#LOWER CI EFFECT SIZE using lower 95% CI from the bootstrapped model:

#We redefine c to equal the treatment coefficient minus the SE created above
c <- boot_pr_model$Lower.bound[2]

pr_effect_size_low <- hedges.g(c, n, m, v, w) #Re-run hedges G on lower CI treatment coefficient
pr_effect_size_low #This is the lower confidence interval for the treatment effect on the following directions subtest.

#UPPER CI EFFECT SIZE using bootrapped 95% upper CI from the above model:

#We redefine c again to equal the treatment coefficient plus the SE created above. Remember c has already been
defined to be the lower confidence interval, we therefore need to add 2 SEs to it.
c <- boot_pr_model$Upper.bound[2]

pr_effect_size_high <- hedges.g(c, n, m, v, w) #Re-run hedges G on upper CI treatment coefficient
pr_effect_size_high #This is the upper confidence interval for the treatment effect on the following directions subtest.

## P-value
p_values <- boot_pr_model$p.value[2]
p_value <- round(p_values, digits = 2)

####################

### 5. Creating primary analysis table (to paste output into report)

##+ This part of the code is to create a formatted output table which can just be copied
##+ into the table empty table in word

## Find missing numbers for model (number of obs with outcome but missing covariates)

n_mis <- nrow(filter(data,
```

```
                treatment == 1 &
                  !is.na(E_score) &
                  (is.na(SettingRegion) |
                     is.na(Settingtype) |
                     is.na(SettingID))))

m_mis <- nrow(filter(data,
                treatment == 0 &
                  is.na(E_score) &
                  (is.na(SettingRegion) |
                     is.na(Settingtype) |
                     is.na(SettingID))))

# Paste in numbers (non-missing and missing) for table
num_t <- paste0(as.character(n), " (", as.character(n_mis), ")")
num_c <- paste0(as.character(m), " (", as.character(m_mis), ")")

# Total numbers
total_n <- n + m
total_n_t_c <- paste0(as.character(total_n), " (", n, ";", m, ")")

## Generate means and CIs

# Treatment
me_t <- t.test(dplyr::filter(data,
                        treatment == 1 &
                          !is.na(E_score) &
                          !is.na(SettingRegion) &
                          !is.na(Settingtype) &
                          !is.na(SettingID))$E_score) #Run t-test on outcome in analytical sample

mean_t <- as.character(round(me_t$estimate, digits = 2)) #extract mean from t-test object (me), to 2dp and as
character to facilitate pasting in CIs.

me_ci_t <- as.character(round(me_t$conf.int[c(1, 2)], digits = 2)) #extract CIs (95%) from t-test object (me), to 2dp and
as character to facilitate pasting

mean_ci_t <- paste0(mean_t, " (", me_ci_t[1], " - ", me_ci_t[2], ")") #Paste CIs in brackets next to mean figure

# Control
me_c <- t.test(dplyr::filter(data,
                        treatment == 0 &
                          !is.na(E_score) &
                          !is.na(SettingRegion) &
                          !is.na(Settingtype) &
                          !is.na(SettingID))$E_score)

mean_c <- as.character(round(me_c$estimate, digits = 2))
me_ci_c <- as.character(round(me_c$conf.int[c(1, 2)], digits = 2))
mean_ci_c <- paste0(mean_c, " (", me_ci_c[1], " - ", me_ci_c[2], ")")


## Effect size with CIs
ES_CI <- paste0(as.character(round(primary_effect_size, digits = 3)),
```

```
        " (", as.character(round(pr_effect_size_low, digits = 2)),
        " -- ",
        as.character(round(pr_effect_size_high, digits = 2)),
        ")")


primary_analysis_table <- t(as.data.frame(c(num_t, mean_ci_t, num_c, mean_ci_c, total_n_t_c, ES_CI, p_value)))
colnames(primary_analysis_table) <- c("N (intervention)", "Mean (intervention", "N (control)", "Mean (control)", "Total
(T;C)", "Hedges g (95% CIs)", "p-value")

rownames(primary_analysis_table) <- "EYTN2 Numeracy"

## Display results - TABLE 1
View(primary_analysis_table)

##Effect Size Estimate Table
u_m <- as.character(round((mean(dplyr::filter(data, #Unadjusted difference in means
                                  treatment == 1 &
                                    !is.na(SettingRegion) &
                                    !is.na(Settingtype) &
                                    !is.na(SettingID))$E_score, na.rm = TRUE)) - (mean(dplyr::filter(data,
                                                                                    treatment == 0 &
                                                                                      !is.na(SettingRegion) &
                                                                                      !is.na(Settingtype) &
                                                                                      !is.na(SettingID))$E_score, na.rm = TRUE)),
digits = 2))

c <- round(coefs["treatment", "Estimate"], digits = 2) #Adjusted difference in means

var_t <- round(v, digits = 2) #Variance of outcome in treatment

var_c <- round(w, digits = 2) #Variance of outcome in control

pooled_var <- round(((((n - 1)*v)+((m - 1)* w))/(n + m - 2)), digits = 2) #pooled variance



# Create table 2
pr_es_est <- t(as.data.frame(c(u_m, c, num_t, var_t, num_c, var_c, pooled_var)))
colnames(pr_es_est) <- c("Unadj. diff in means", "Adj diff in means", "N (treatment)", "Variance (treatment)", "N
(control)", "Variance (control)", "Pooled Variance")
rownames(pr_es_est) <- "EYTN Numeracy Test"
View(pr_es_est)


###################### The ONE Analysis: EYPP Subgroup ##################

## [PURPOSE OF SCRIPT]:

####+  This script conducts the EYPP subgroup analysis for the ONE
####+  This conducts the primary analysis on the EYPP subgroup with the outcome of interest being the results of the
####+  EYTN test for early numeracy
####+  This will follow the same procedure as the primary analysis
```

####+  The primary model will then be adjusted to include EYPP eligibility as a factor and analysis will be conducted in the whole sample

############################

```
# Clear workspace
rm(list=ls())
set.seed(999) #NB this is the same seed set in the randomisation code at the start of the project so we shall retain it here for consistency.

# Load packages
library("haven")
library("dplyr")
library("ggplot2")
library("lme4")
library("sjstats")
library("Hmisc")
library("performance")
library("lmtest")
library("eeptools")
library("lmerTest")
library("officer")
library("boot")
library("boot.pval")
library("PowerUpR")

# Load cleaned data
data <- read_dta("//Nevis/RE/Projects/In_Progress/022807.014_P2899_EEF_The_ONE/WIP/06. Data collection/02. Endline/02. Data_NotMirrored/BaselineEndlineComplianceClean2.dta")

data <- filter(data, endlineonly != 1)

#descriptive statistics of EYPP status
Hmisc::describe(data$E_EYPPStatus)

#6 missing-- none of these would have been in the analytical sample
#311/1953 are EYPP eligible

data.class(data$E_EYPPStatus)
#character variable--I'll convert to factor. Only two distinct values so no further cleaning needed
data$E_EYPPStatus <- as.factor(data$E_EYPPStatus)

#creating new dataset for the subgroup analysis
data_EYPP <- filter(data, E_EYPPStatus=="Y")


############################################################
############### Analysis of EYPP Subgroup ################
############################################################

# Section A: Descriptive statistics


#1. Histograms of outcomes
```

```
#Use this to check distribution of endline and baseline scores for our outcome

#We are mainly checking for ceiling and floor effects

# In overall data
EYTN_ovr_hist <- hist(data_EYPP$E_score,
                     main="Histogram of EYPP Pupils' EYTN subtest at endline (overall sample)",
                     xlab = "Early Numeracy Scores",
                     xlim = c(0, 120),
                     breaks = seq(0, 120, 10),
                     xaxp = c(0, 120, 6))

#+In analytical data (the analytical data is defined as the data used in the multi-level regression model.
#+It includes the observations where we have complete data for all covariates included in the outcome model)
EYTN_model_hist <- hist(dplyr::filter(data_EYPP,
                                      !is.na(B_score) &
                                        !is.na(Settingtype) &
                                        !is.na(SettingRegion) &
                                        !is.na(SettingID))$E_score,
                        main="Histogram of EYPP Pupils' EYTN subtest at endline (analytical sample)",
                        xlab = "Early numeracy scores",
                        xlim = c(0, 120),
                        breaks = seq(0, 120, 10),
                        xaxp = c(0, 120, 6))


## No floor or ceiling effects in endline test distribution
## Now for baseline
EYTN_ovr_histB <- hist(data_EYPP$B_score,
                       main="Histogram of EYPP EYTN subtest at baseline (overall sample)",
                       xlab = "Early Numeracy Scores",
                       xlim = c(0, 120),
                       breaks = seq(0, 120, 10),
                       xaxp = c(0, 120, 6))

#+In analytical data (the analytical data is defined as the data used in the multi-level regression model.
#+It includes the observations where we have complete data for all covariates included in the outcome model)
EYTN_model_histB <- hist(dplyr::filter(data_EYPP,
                                       !is.na(B_score) &
                                         !is.na(Settingtype) &
                                         !is.na(SettingRegion) &
                                         !is.na(SettingID))$B_score,
                         main="Histogram of EYPP EYTN subtest at baseline (analytical sample)",
                         xlab = "Early numeracy scores",
                         xlim = c(0, 120),
                         breaks = seq(0, 120, 10),
                         xaxp = c(0, 120, 6))

##Looks like floor effects at baseline. Not a major concern but will include some stats re standard dev from the floor
sd <- sd(data_EYPP$B_score, na.rm = T)

# Create df which is a copy of analytical data
data_sd <- dplyr::filter(data_EYPP,
```

```
                    !is.na(B_score) &
                      !is.na(Settingtype) &
                      !is.na(SettingRegion) &
                      !is.na(SettingID))

data_sd <- data_sd %>%
  mutate(sd_dummy = 0)

data_sd$sd_dummy <- ifelse(data_sd$B_score < sd, 1, 0)

mean(data_sd$sd_dummy)

## 32.32% of observations are within 1 sd of the floor

#2. Means, SDs, Min and Max (for whole sample and analytical sample)

#Overall sample
Hmisc::describe(dplyr::filter(data_EYPP,
                        !is.na(E_score))$E_score)
sd(dplyr::filter(data_EYPP,
             !is.na(E_score))$E_score)


##Analytical sample
Hmisc::describe(dplyr::filter(data_EYPP,
                        !is.na(B_score) &
                          !is.na(Settingtype) &
                          !is.na(SettingRegion) &
                          !is.na(E_score) &
                          !is.na(SettingID))$E_score)
sd(dplyr::filter(data_EYPP,
             !is.na(B_score) &
               !is.na(Settingtype) &
               !is.na(SettingRegion) &
               !is.na(E_score) &
               !is.na(SettingID))$E_score)


#And by treatment and control groups:
Hmisc::describe(dplyr::filter(data_EYPP,  #treatment
                        treatment == 1 &
                          !is.na(B_score) &
                          !is.na(Settingtype) &
                          !is.na(SettingRegion) &
                          !is.na(E_score) &
                          !is.na(SettingID))$E_score)

sd(dplyr::filter(data_EYPP, #treatment
             treatment == 1 &
               !is.na(B_score) &
               !is.na(Settingtype) &
               !is.na(SettingRegion) &
               !is.na(E_score) &
               !is.na(SettingID))$E_score)
```

```
Hmisc::describe(dplyr::filter(data_EYPP,  #control
                              treatment == 0 &
                                !is.na(B_score) &
                                !is.na(Settingtype) &
                                !is.na(SettingRegion) &
                                !is.na(E_score) &
                                !is.na(SettingID))$E_score)
sd(dplyr::filter(data_EYPP, #control
                 treatment == 0 &
                   !is.na(B_score) &
                   !is.na(Settingtype) &
                   !is.na(SettingRegion) &
                   !is.na(E_score) &
                   !is.na(SettingID))$E_score)

##################################################
##################################################
##################################################

# Section B: Primary outcome model

### 1. Run the multi-level model.

pr_model1 <- lmer(E_score ~
                    treatment +
                    B_score +
                    SettingRegion +
                    Settingtype +
                    (1 | SettingID),
                  data=data_EYPP, REML = FALSE)

summary(pr_model1) #Produce the results

performance::icc(pr_model1, ci = 0.95) #ICC

#########

### 2. Testing OLS assumptions

#Residual diagnostics: Testing normality of residuals OLS assumption
resid_pr <- resid(pr_model1) #Create object which stores the residuals of the model
plot_resid_pr <- plot(density(resid_pr)) #Kernel density plot to explore normality
shapiro.test(resid_pr) #Shapiro-wilk test for normality. !FLAG potentially K-S test better with larger sample sizes!
(p<0.05 rejects H0 that data is normally distributed)
ks.test(resid_pr, "pnorm") #K-S test (p<0.05 rejects H0 that data is normally distributed

##KS rejects H0 but SW does not.

# QQ line
qq_line <- ggplot(data = data.frame(resid = resid_pr), aes(sample = resid)) +
  stat_qq() +
  stat_qq_line() +
  labs(title = "Q-Q Plot of Residuals", x = "Theoretical Quantiles", y = "Sample Quantiles") +
```

```
  theme_minimal()

pr_resid_df <- as.data.frame(resid_pr)
sd_resid <- sd(resid_pr)
max_resid <- max(resid_pr)
min_resid <- min(resid_pr)

resid_pr_kd<- ggplot(pr_resid_df, aes(x = resid_pr)) +
  geom_density(aes(color = "Kernel density"), linewidth = 1) +
  stat_function(aes(color = "Normal density"),
                fun = function(x) dnorm(x, mean = 0, sd = sd_resid),
                linetype = "dotted", linewidth = 1) +
  theme_minimal() +
  scale_x_continuous(limits = c(min_resid, max_resid)) +
  scale_color_manual(name = "Legend", values = c("Kernel density" = "#000080", "Normal density" = "red")) +
  labs(
    title = "Early Numeracy - Residuals Density Plot",
    x = "Residuals",
    y = "Density"
  )

#Testing linearity and assumption that residual errors have a mean of 0
plot(pr_model1, col = "red") #We want the line here to be horizontal and at 0


### Figures to export:
resid_pr_kd
qq_line

plots <- list(resid_pr_kd, qq_line)

### Export graphs to word

# Create a Word document
doc <- read_docx()

# Save each plot as an image and add to the Word document
for (i in seq_along(plots)) {
  # Export the plot to a temporary file
  file_name <- tempfile(fileext = ".png")
  ggplot2::ggsave(filename = file_name, plot = plots[[i]])

  # Add the plot to the Word document
  doc <- doc %>%
    body_add_par(value = paste("Plot", i), style = "heading 1") %>%
    body_add_img(src = file_name, width = 6, height = 4)
}

# Save the document
print(doc, target = "//Nevis/RE/Projects/In_Progress/022807.014_P2899_EEF_The_ONE/WIP/09. Impact/Graphs &
Summary Tables/EYPPsample_primary_outcome.docx")

##########
```

```
### 3.  Bootstrapping CIs and p-value
#+ The QQ residual plot, Shapiro-Wilk, and K-S test
#+ results suggest that the residuals are not normally distributed. We will
#+ therefore need to re-estimate CIs and p-values using bootstrapping.

set.seed(999)
boot_pr_model1 <- boot.pval::boot_summary(pr_model1,
                                  type = "norm",
                                  method = NULL,
                                  conf.level = 0.95)
boot_pr_model1

######################

### 4.  Effect-size estimation

#Create Hedges g function.
hedges.g <- function(c, n, m, v, w){
  c/sqrt((((n - 1)*v)+((m - 1)* w))/(n + m-2))
}


#We use the output from the model above to calculate the effect size: (NB: This needs to be done for confidence intervals too!)
#We need to define the objects needed to be inputted into the Hedges G function created above (c, n, m, v, & w)

coefs <- data.frame(summary(pr_model1)$coefficients) #create data frame of all coefficients from primary outcome model

c <- coefs["treatment", "Estimate"] #Extract the treatment coefficient

n <- nrow(dplyr::filter(data_EYPP,
                        treatment == 1 &
                          !is.na(E_score) &
                          !is.na(B_score) &
                          !is.na(SettingRegion) &
                          !is.na(Settingtype) &
                          !is.na(SettingID))) #Number of individuals in treatment group in the model

m <- nrow(dplyr::filter(data_EYPP,
                        treatment == 0 &
                          !is.na(E_score) &
                          !is.na(B_score) &
                          !is.na(SettingRegion) &
                          !is.na(Settingtype) &
                          !is.na(SettingID))) #Number of individuals in control group in the model

 v <- var(dplyr::filter(data_EYPP,
                        treatment == 1 &
                         !is.na(B_score) &
                         !is.na(SettingRegion) &
                         !is.na(Settingtype) &
                         !is.na(E_score) &
                         !is.na(SettingID))$E_score) #Variance in outcome among treatment group
```

```
 w <- var(dplyr::filter(data_EYPP,
                        treatment == 0 &
                         !is.na(B_score) &
                         !is.na(SettingRegion) &
                         !is.na(Settingtype) &
                         !is.na(E_score) &
                         !is.na(SettingID))$E_score)  #Variance in outcome among control group

primary_effect_size <- hedges.g(c, n, m, v, w)
primary_effect_size

#LOWER CI EFFECT SIZE using lower 95% CI from the bootstrapped model:

#We redefine c to equal the treatment coefficient minus the SE created above
c <- boot_pr_model1$Lower.bound[2]

pr_effect_size_low1 <- hedges.g(c, n, m, v, w) #Re-run hedges G on lower CI treatment coefficient
pr_effect_size_low1 #This is the lower confidence interval for the treatment effect on the following directions subtest.



#UPPER CI EFFECT SIZE using bootrapped 95% upper CI from the above model:

#We redefine c again to equal the treatment coefficient plus the SE created above. Remember c has already been
defined to be the lower confidence interval, we therefore need to add 2 SEs to it.
c <- boot_pr_model1$Upper.bound[2]

pr_effect_size_high1 <- hedges.g(c, n, m, v, w) #Re-run hedges G on upper CI treatment coefficient
pr_effect_size_high1 #This is the upper confidence interval for the treatment effect on the following directions subtest.

## P-value
p_values <- boot_pr_model1$p.value[2]
p_value <- round(p_values, digits = 2)


### 5. Creating primary analysis table (to paste output into report)

##+ This part of the code is to create a formatted output table which can just be copied
##+ into the table empty table in word

## Find missing numbers for model (number of obs with outcome but missing covariates)

# Treatment
n_mis <- nrow(dplyr::filter(data_EYPP,
                            treatment == 1 &
                              !is.na(E_score) &
                              (is.na(B_score) |
                                 is.na(SettingRegion) |
                                 is.na(Settingtype) |
                                 is.na(SettingID))))

# Control
m_mis <- nrow(dplyr::filter(data_EYPP,
```

```
                    treatment == 0 &
                      !is.na(E_score) &
                      (is.na(B_score) |
                         is.na(SettingRegion) |
                         is.na(Settingtype) |
                         is.na(SettingID))))

# Paste in numbers (non-missing and missing) for table
num_t <- paste0(as.character(n), " (", as.character(n_mis), ")")
num_c <- paste0(as.character(m), " (", as.character(m_mis), ")")

# Total numbers
total_n <- n + m
total_n_t_c <- paste0(as.character(total_n), " (", n, ";", m, ")")

## Generate means and CIs

# Treatment
me_t <- t.test(dplyr::filter(data_EYPP,
                    treatment == 1 &
                      !is.na(E_score) &
                      !is.na(B_score) &
                      !is.na(SettingRegion) &
                      !is.na(Settingtype) &
                      !is.na(SettingID))$E_score) #Run t-test on outcome in analytical sample

mean_t <- as.character(round(me_t$estimate, digits = 2)) #extract mean from t-test object (me), to 2dp and as character to facilitate pasting in CIs.

me_ci_t <- as.character(round(me_t$conf.int[c(1, 2)], digits = 2)) #extract CIs (95%) from t-test object (me), to 2dp and as character to facilitate pasting

mean_ci_t <- paste0(mean_t, " (", me_ci_t[1], " - ", me_ci_t[2], ")") #Paste CIs in brackets next to mean figure

# Control
me_c <- t.test(dplyr::filter(data_EYPP,
                    treatment == 0 &
                      !is.na(E_score) &
                      !is.na(B_score) &
                      !is.na(SettingRegion) &
                      !is.na(Settingtype) &
                      !is.na(SettingID))$E_score)

mean_c <- as.character(round(me_c$estimate, digits = 2))
me_ci_c <- as.character(round(me_c$conf.int[c(1, 2)], digits = 2))
mean_ci_c <- paste0(mean_c, " (", me_ci_c[1], " - ", me_ci_c[2], ")")


## Effect size with CIs
ES_CI <- paste0(as.character(round(primary_effect_size, digits = 3)),
            " (", as.character(round(pr_effect_size_low1, digits = 2)),
            " -- ",
            as.character(round(pr_effect_size_high1, digits = 2)),
            ")")
```

```
## Create table

primary_analysis_table <- t(as.data.frame(c(num_t, mean_ci_t, num_c, mean_ci_c, total_n_t_c, ES_CI, p_value)))
colnames(primary_analysis_table) <- c("N (intervention)", "Mean (intervention", "N (control)", "Mean (control)", "Total
(T;C)", "Hedges g (95% CIs)", "p-value")

rownames(primary_analysis_table) <- "EYTN2 Numeracy"

## Display results - TABLE 1
View(primary_analysis_table)

############

### 6. Creating effect size estimation table. Most of the required fields have already been created

u_m <- as.character(round((mean(dplyr::filter(data_EYPP, #Unadjusted difference in means
                                      treatment == 1 &
                                        !is.na(B_score) &
                                        !is.na(SettingRegion) &
                                        !is.na(Settingtype) &
                                        !is.na(SettingID))$E_score, na.rm = TRUE)) - (mean(dplyr::filter(data_EYPP,
                                                                                        treatment == 0 &
                                                                                          !is.na(B_score) &
                                                                                          !is.na(SettingRegion) &
                                                                                          !is.na(Settingtype) &
                                                                                          !is.na(SettingID))$E_score, na.rm = TRUE)),
digits = 2))

c <- round(coefs["treatment", "Estimate"], digits = 2) #Adjusted difference in means

var_t <- round(v, digits = 2) #Variance of outcome in treatment

var_c <- round(w, digits = 2) #Variance of outcome in control

pooled_var <- round((((n - 1)*v)+((m - 1)* w))/(n + m-2), digits = 2) #pooled standard deviation



# Create table 2
pr_es_est <- t(as.data.frame(c(u_m, c, num_t, var_t, num_c, var_c, pooled_var)))
colnames(pr_es_est) <- c("Unadj. diff in means", "Adj diff in means", "N (treatment)", "Variance (treatment)", "N
(control)", "Variance (control)", "Pooled Variance")
rownames(pr_es_est) <- "EYTN Numeracy Test"
View(pr_es_est)


#Caclulation of R-squared values for MDES calculations for EYPP subgroup

# Null model
null_model <- lmer(E_score ~ treatment + (1 | SettingID), data = data_EYPP, REML = FALSE)

# Variance components for the null model
VarCorr(null_model)
```

```
# Variance components for the full model
VarCorr(pr_model1)


within_var_null <- attr(VarCorr(null_model), "sc")^2
within_var_full <- attr(VarCorr(pr_model1), "sc")^2

# Level 1 R^2 (Within-cluster variance explained)
R2_level1 <- (within_var_null - within_var_full) / within_var_null
R2_level1


between_var_null <- as.numeric(VarCorr(null_model)$SettingID[1])
between_var_full <- as.numeric(VarCorr(pr_model1)$SettingID[1])

# Level 2 R^2 (Between-cluster variance explained)
R2_level2 <- (between_var_null - between_var_full) / between_var_null
R2_level2

settingcounts <- data_EYPP %>%
  filter(!is.na(E_score) &
           !is.na(B_score) &
           !is.na(SettingRegion) &
           !is.na(Settingtype) &
           !is.na(SettingID)) %>%
  group_by(SettingID) %>%
  dplyr::summarise(, n = n(), status = mean(treatment))

settingaverage <- mean(settingcounts$n)

J = nrow(settingcounts)

## mdes
mdes.cra2(power = 0.8,
          alpha = 0.05,
          two.tailed = TRUE,
          rho2 = 0.064,
          p = 0.5,
          r21 = R2_level1,
          r22 = R2_level2,
          n = settingaverage,
          J = J)

#############################################################################
############### Primary Analysis accounting for EYPP Status ###############
#############################################################################

##Overall stats already exist. Will only reproduce code for the analytical sample.

#+In analytical data (the analytical data is defined as the data used in the multi-level regression model.
#+It includes the observations where we have complete data for all covariates included in the outcome model)
EYTN_model_hist <- hist(dplyr::filter(data,
                                      !is.na(B_score) &
```

```
                        !is.na(E_EYPPStatus) &
                        !is.na(Settingtype) &
                        !is.na(SettingRegion) &
                        !is.na(SettingID))$E_score,
                main="Histogram of EYTN subtest at endline (analytical sample)",
                xlab = "Early numeracy scores",
                xlim = c(0, 120),
                breaks = seq(0, 120, 10),
                xaxp = c(0, 120, 6))

EYTN_model_histB <- hist(dplyr::filter(data,
                        !is.na(B_score) &
                         !is.na(E_EYPPStatus) &
                         !is.na(Settingtype) &
                         !is.na(SettingRegion) &
                         !is.na(SettingID))$B_score,
                main="Histogram of EYTN subtest at baseline (analytical sample)",
                xlab = "Early numeracy scores",
                xlim = c(0, 120),
                breaks = seq(0, 120, 10),
                xaxp = c(0, 120, 6))

## Will add analysis of sds from the floor here
sd <- sd(data$B_score, na.rm = T)

# Create df which is a copy of analytical data
data_sd <- dplyr::filter(data,
                !is.na(B_score) &
                 !is.na(E_EYPPStatus) &
                 !is.na(Settingtype) &
                 !is.na(SettingRegion) &
                 !is.na(SettingID))

data_sd <- data_sd %>%
  mutate(sd_dummy = 0)

data_sd$sd_dummy <- ifelse(data_sd$B_score < sd, 1, 0)

mean(data_sd$sd_dummy)

##Here we see that 30.57% of observations are within one sd of the floor

#2. Means, SDs, Min and Max (for analytical sample)
##Analytical sample
Hmisc::describe(dplyr::filter(data,
                    !is.na(B_score) &
                     !is.na(E_EYPPStatus) &
                     !is.na(Settingtype) &
                     !is.na(SettingRegion) &
                     !is.na(E_score) &
                     !is.na(SettingID))$E_score)
sd(dplyr::filter(data,
            !is.na(B_score) &
              !is.na(E_EYPPStatus) &
```

```
                !is.na(Settingtype) &
                !is.na(SettingRegion) &
                !is.na(E_score) &
                !is.na(SettingID))$E_score)

##The same as the primary analysis sample--introducing EYPP status doesn't result in any further missingness

#And by treatment and control groups:
Hmisc::describe(dplyr::filter(data,  #treatment
                          treatment == 1 &
                            !is.na(B_score) &
                            !is.na(E_EYPPStatus) &
                            !is.na(Settingtype) &
                            !is.na(SettingRegion) &
                            !is.na(E_score) &
                            !is.na(SettingID))$E_score)

sd(dplyr::filter(data, #treatment
              treatment == 1 &
                !is.na(B_score) &
                !is.na(E_EYPPStatus) &
                !is.na(Settingtype) &
                !is.na(SettingRegion) &
                !is.na(E_score) &
                !is.na(SettingID))$E_score)

Hmisc::describe(dplyr::filter(data,  #control
                          treatment == 0 &
                            !is.na(B_score) &
                            !is.na(E_EYPPStatus) &
                            !is.na(Settingtype) &
                            !is.na(SettingRegion) &
                            !is.na(E_score) &
                            !is.na(SettingID))$E_score)
sd(dplyr::filter(data, #control
              treatment == 0 &
                !is.na(B_score) &
                !is.na(E_EYPPStatus) &
                !is.na(Settingtype) &
                !is.na(SettingRegion) &
                !is.na(E_score) &
                !is.na(SettingID))$E_score)

################################################
################################################
################################################

# Section B: Primary outcome model

### 1. Run the multi-level model.

pr_model2 <- lmer(E_score ~
                treatment +
                B_score +
```

```
              E_EYPPStatus +
              treatment:E_EYPPStatus +
              SettingRegion +
              Settingtype +
              (1 | SettingID),
            data=data, REML = FALSE)

summary(pr_model2) #Produce the results

performance::icc(pr_model2) #ICC

#########

### 2. Testing OLS assumptions

#Residual diagnostics: Testing normality of residuals OLS assumption
resid_pr <- resid(pr_model2) #Create object which stores the residuals of the model
plot_resid_pr <- plot(density(resid_pr)) #Kernel density plot to explore normality
shapiro.test(resid_pr) #Shapiro-wilk test for normality. !FLAG potentially K-S test better with larger sample sizes!
(p<0.05 rejects H0 that data is normally distributed)
ks.test(resid_pr, "pnorm") #K-S test (p<0.05 rejects H0 that data is normally distributed

##Both reject normality

# QQ line
qq_line <- ggplot(data = data.frame(resid = resid_pr), aes(sample = resid)) +
  stat_qq() +
  stat_qq_line() +
  labs(title = "Q-Q Plot of Residuals", x = "Theoretical Quantiles", y = "Sample Quantiles") +
  theme_minimal()


#Plotting residual density against a normal distribution for comparison. NB this is using the residual objects created
from the primary model above
pr_resid_df <- as.data.frame(resid_pr)
sd_resid <- sd(resid_pr)
max_resid <- max(resid_pr)
min_resid <- min(resid_pr)

resid_pr_kd<- ggplot(pr_resid_df, aes(x = resid_pr)) +
  geom_density(aes(color = "Kernel density"), linewidth = 1) +
  stat_function(aes(color = "Normal density"),
                fun = function(x) dnorm(x, mean = 0, sd = sd_resid),
                linetype = "dotted", linewidth = 1) +
  theme_minimal() +
  scale_x_continuous(limits = c(min_resid, max_resid)) +
  scale_color_manual(name = "Legend", values = c("Kernel density" = "#000080", "Normal density" = "red")) +
  labs(
    title = "Early Numeracy - Residuals Density Plot",
    x = "Residuals",
    y = "Density"
  )

#Testing linearity and assumption that residual errors have a mean of 0
```

```
plot(pr_model2, col = "red") #We want the line here to be horizontal and at 0


### Figures to export:
resid_pr_kd
qq_line

plots <- list(resid_pr_kd, qq_line)

### Export graphs to word

# Create a Word document
doc <- read_docx()

# Save each plot as an image and add to the Word document
for (i in seq_along(plots)) {
  # Export the plot to a temporary file
  file_name <- tempfile(fileext = ".png")
  ggplot2::ggsave(filename = file_name, plot = plots[[i]])

  # Add the plot to the Word document
  doc <- doc %>%
    body_add_par(value = paste("Plot", i), style = "heading 1") %>%
    body_add_img(src = file_name, width = 6, height = 4)
}

print(doc, target = "//Nevis/RE/Projects/In_Progress/022807.014_P2899_EEF_The_ONE/WIP/09. Impact/Graphs &
Summary Tables/EYPPanalysis_outcome.docx")


##########


### 3.  Bootstrapping CIs and p-value
#+ The QQ residual plot, Shapiro-Wilk, and K-S test
#+ results suggest that the residuals are not normally distributed. We will
#+ therefore need to re-estimate CIs and p-values using bootstrapping.

set.seed(999)
boot_pr_model2 <- boot.pval::boot_summary(pr_model2,
                                  type = "norm",
                                  method = NULL,
                                  conf.level = 0.95)
boot_pr_model2

######################
##+ As we have introduced an interaction between the treatment effect and EYPP effect, we'll need to redefine
##+ the effect size calculation
hedges.g <- function(c, i, n, m, v, w){
  (c + i)/sqrt((((n - 1)*v)+((m - 1)* w))/n + (m-2))
}


### 4.  Effect-size estimation
```

```
coefs <- data.frame(summary(pr_model2)$coefficients) #create data frame of all coefficients from primary outcome model

c <- coefs["treatment", "Estimate"] #Extract the treatment coefficient

i <- coefs["treatment:E_EYPPStatusY", "Estimate"]

n <- nrow(dplyr::filter(data,
                        treatment == 1 &
                          !is.na(E_score) &
                          !is.na(B_score) &
                          !is.na(E_EYPPStatus) &
                          !is.na(SettingRegion) &
                          !is.na(Settingtype) &
                          !is.na(SettingID))) #Number of individuals in treatment group in the model

m <- nrow(dplyr::filter(data,
                        treatment == 0 &
                          !is.na(E_score) &
                          !is.na(B_score) &
                          !is.na(E_EYPPStatus) &
                          !is.na(SettingRegion) &
                          !is.na(Settingtype) &
                          !is.na(SettingID))) #Number of individuals in control group in the model

v <- var(dplyr::filter(data,
                       treatment == 1 &
                         !is.na(B_score) &
                         !is.na(E_EYPPStatus) &
                         !is.na(SettingRegion) &
                         !is.na(Settingtype) &
                         !is.na(E_score) &
                         !is.na(SettingID))$E_score) #Variance in outcome among treatment group

w <- var(dplyr::filter(data,
                       treatment == 0 &
                         !is.na(B_score) &
                         !is.na(E_EYPPStatus) &
                         !is.na(SettingRegion) &
                         !is.na(Settingtype) &
                         !is.na(E_score) &
                         !is.na(SettingID))$E_score)  #Variance in outcome among control group

primary_effect_size2 <- hedges.g(c, i, n, m, v, w)
primary_effect_size2

#LOWER CI EFFECT SIZE using lower 95% CI from the bootstrapped model:

#We redefine c to equal the treatment coefficient minus the SE created above
c <- boot_pr_model2$Lower.bound[2]

i <- boot_pr_model2$Lower.bound[9]
```

```
pr_effect_size_low2 <- hedges.g(c, i, n, m, v, w) #Re-run hedges G on lower CI treatment coefficient
pr_effect_size_low2 #This is the lower confidence interval for the treatment effect on the following directions subtest.



#UPPER CI EFFECT SIZE using bootrapped 95% upper CI from the above model:

#We redefine c again to equal the treatment coefficient plus the SE created above. Remember c has already been defined to be the lower confidence interval, we therefore need to add 2 SEs to it.
c <- boot_pr_model2$Upper.bound[2]
i <- boot_pr_model2$Upper.bound[9]

pr_effect_size_high2 <- hedges.g(c, i, n, m, v, w) #Re-run hedges G on upper CI treatment coefficient
pr_effect_size_high2 #This is the upper confidence interval for the treatment effect on the following directions subtest.

## P-value
p_values <- boot_pr_model2$p.value[2]
p_value <- round(p_values, digits = 2)


### 5. Creating primary analysis table (to paste output into report)

##+ This part of the code is to create a formatted output table which can just be copied
##+ into the table empty table in word

## Find missing numbers for model (number of obs with outcome but missing covariates)

# Treatment
n_mis <- nrow(dplyr::filter(data,
                            treatment == 1 &
                              !is.na(E_score) &
                              (is.na(B_score) |
                                 is.na(E_EYPPStatus) |
                                 is.na(SettingRegion) |
                                 is.na(Settingtype) |
                                 is.na(SettingID))))

# Control
m_mis <- nrow(dplyr::filter(data,
                            treatment == 0 &
                              !is.na(E_score) &
                              (is.na(B_score) |
                                 is.na(E_EYPPStatus) |
                                 is.na(SettingRegion) |
                                 is.na(Settingtype) |
                                 is.na(SettingID))))

# Paste in numbers (non-missing and missing) for table
num_t <- paste0(as.character(n), " (", as.character(n_mis), ")")
num_c <- paste0(as.character(m), " (", as.character(m_mis), ")")

# Total numbers
total_n <- n + m
total_n_t_c <- paste0(as.character(total_n), " (", n, ";", m, ")")
```

```
## Generate means and CIs

# Treatment
me_t <- t.test(dplyr::filter(data,
                        treatment == 1 &
                          !is.na(E_score) &
                          !is.na(B_score) &
                          !is.na(E_EYPPStatus) &
                          !is.na(SettingRegion) &
                          !is.na(Settingtype) &
                          !is.na(SettingID))$E_score) #Run t-test on outcome in analytical sample

mean_t <- as.character(round(me_t$estimate, digits = 2)) #extract mean from t-test object (me), to 2dp and as character to facilitate pasting in CIs.

me_ci_t <- as.character(round(me_t$conf.int[c(1, 2)], digits = 2)) #extract CIs (95%) from t-test object (me), to 2dp and as character to facilitate pasting

mean_ci_t <- paste0(mean_t, " (", me_ci_t[1], " - ", me_ci_t[2], ")") #Paste CIs in brackets next to mean figure

# Control
me_c <- t.test(dplyr::filter(data,
                        treatment == 0 &
                          !is.na(E_score) &
                          !is.na(B_score) &
                          !is.na(E_EYPPStatus) &
                          !is.na(SettingRegion) &
                          !is.na(Settingtype) &
                          !is.na(SettingID))$E_score)

mean_c <- as.character(round(me_c$estimate, digits = 2))
me_ci_c <- as.character(round(me_c$conf.int[c(1, 2)], digits = 2))
mean_ci_c <- paste0(mean_c, " (", me_ci_c[1], " - ", me_ci_c[2], ")")


## Effect size with CIs
ES_CI <- paste0(as.character(round(primary_effect_size2, digits = 3)),
            " (", as.character(round(pr_effect_size_low2, digits = 2)),
            " -- ",
            as.character(round(pr_effect_size_high2, digits = 2)),
            ")")

## Create table

primary_analysis_table2 <- t(as.data.frame(c(num_t, mean_ci_t, num_c, mean_ci_c, total_n_t_c, ES_CI, p_value)))
colnames(primary_analysis_table2) <- c("N (intervention)", "Mean (intervention", "N (control)", "Mean (control)", "Total (T;C)", "Hedges g (95% CIs)", "p-value")

rownames(primary_analysis_table2) <- "EYTN2 Numeracy"

## Display results - TABLE 1
View(primary_analysis_table2)
```

```
###########

### 6. Creating effect size estimation table. Most of the required fields have already been created

u_m <- as.character(round((mean(dplyr::filter(data,
                                    treatment == 1 &
                                      !is.na(B_score) &
                                      !is.na(E_EYPPStatus) &
                                      !is.na(SettingRegion) &
                                      !is.na(Settingtype) &
                                      !is.na(SettingID))$E_score, na.rm = TRUE)) - (mean(dplyr::filter(data,
                                                                                                      treatment == 0 &
                                                                                                        !is.na(B_score) &
                                                                                                        !is.na(E_EYPPStatus) &
                                                                                                        !is.na(SettingRegion) &
                                                                                                        !is.na(Settingtype) &
                                                                                                        !is.na(SettingID))$E_score, na.rm = TRUE)),
digits = 2))

c <- round(coefs["treatment", "Estimate"], digits = 2)
i <- round(coefs["treatment:E_EYPPStatusY", "Estimate"], digits = 2)

var_t <- round(v, digits = 2) #Variance of outcome in treatment

var_c <- round(w, digits = 2) #Variance of outcome in control

pooled_var <- round(((((n - 1)*v)+((m - 1)* w))/(n + m - 2)), digits = 2) #pooled standard deviation



# Create table 2
pr_es_est2 <- t(as.data.frame(c(u_m, c, i, num_t, var_t, num_c, var_c, pooled_var)))
colnames(pr_es_est2) <- c("Unadj. diff in means", "Adj diff in means in non-EYPP subgroup", "Adj diff in means in
EYPP subgroup", "N (treatment)", "Variance (treatment)", "N (control)", "Variance (control)", "Pooled Variance")
rownames(pr_es_est2) <- "EYTN Numeracy Test"
View(pr_es_est2)
#####################  The ONE Analysis: Sensitivity Analysis ##################
######################### Excluding Maths Champions Settings ####################

## [PURPOSE OF SCRIPT]:

####+  This script conducts a sensitivity analysis for the ONE.
####+  We are exploring results of the impact of the ONE on the results of the EYTN test for early numeracy
####+  excluding the data from tests for the settings which participated in  the maths champions program.

####+    For this outcome, we need to create a dataset excluding these settings
####+    The remainder of the analysis follows the same procedure as the primary data ana;usis

#############################
#############################

# Clear workspace
rm(list=ls())
```

```
set.seed(999) #NB this is the same seed set in the randomisation code at the start of the project so we shall retain it here for consistency.

# Load packages
library("haven")
library("dplyr")
library("ggplot2")
library("lme4")
library("sjstats")
library("Hmisc")
library("performance")
library("lmtest")
library("eeptools")
library("lmerTest")
library("officer")
library("boot")
library("boot.pval")


# Load cleaned data
data <- read_dta("//Nevis/RE/Projects/In_Progress/022807.014_P2899_EEF_The_ONE/WIP/06. Data collection/02. Endline/02. Data_NotMirrored/BaselineEndlineComplianceClean2.dta")

data <- filter(data, mathschampions == 0)
data <- filter(data, endlineonly != 1)

#######################################################

# Section A: Descriptive statistics


#1. Histograms of outcomes

#Use this to check distribution of endline and baseline scores for our outcome

#We are mainly checking for ceiling and floor effects

# In overall data
EYTN_ovr_hist <- hist(data$E_score,
                      main="Histogram of EYTN subtest at endline (overall sample)",
                      xlab = "Early Numeracy Scores",
                      xlim = c(0, 120),
                      breaks = seq(0, 120, 10),
                      xaxp = c(0, 120, 6))

#+In analytical data (the analytical data is defined as the data used in the multi-level regression model.
#+It includes the observations where we have complete data for all covariates included in the outcome model)
EYTN_model_hist <- hist(dplyr::filter(data,
                                      !is.na(B_score) &
                                        !is.na(Settingtype) &
                                        !is.na(SettingRegion) &
                                        !is.na(SettingID))$E_score,
                        main="Histogram of EYTN subtest at endline (analytical sample)",
                        xlab = "Early numeracy scores",
```

```
                xlim = c(0, 120),
                breaks = seq(0, 120, 10),
                xaxp = c(0, 120, 6))


## No floor or ceiling effects in endline test distribution
## Now for baseline
EYTN_ovr_histB <- hist(data$B_score,
                main="Histogram of EYTN subtest at baseline (overall sample)",
                xlab = "Early Numeracy Scores",
                xlim = c(0, 120),
                breaks = seq(0, 120, 10),
                xaxp = c(0, 120, 6))

#+In analytical data (the analytical data is defined as the data used in the multi-level regression model.
#+It includes the observations where we have complete data for all covariates included in the outcome model)
EYTN_model_histB <- hist(dplyr::filter(data,
                               !is.na(B_score) &
                                 !is.na(Settingtype) &
                                 !is.na(SettingRegion) &
                                 !is.na(SettingID))$B_score,
                 main="Histogram of EYTN subtest at baseline (analytical sample)",
                 xlab = "Early numeracy scores",
                 xlim = c(0, 120),
                 breaks = seq(0, 120, 10),
                 xaxp = c(0, 120, 6))

#+ We're unlikely to include any additional sensitivity analysis for this as any floor effects are only observed at
baseline.
#+ However, will include additional analysis here to see what proportion of the observations are within one standard
deviation of the floor.

sd <- sd(data$B_score, na.rm = T)

# Create df which is a copy of analytical data
data_sd <- dplyr::filter(data,
                  !is.na(B_score) &
                    !is.na(Settingtype) &
                    !is.na(SettingRegion) &
                    !is.na(SettingID))

data_sd <- data_sd %>%
  mutate(sd_dummy = 0)

data_sd$sd_dummy <- ifelse(data_sd$B_score < sd, 1, 0)

mean(data_sd$sd_dummy)

##Here we see that 31.00% of observations are within one sd of the floor

#2. Means, SDs, Min and Max (for whole sample and analytical sample)

#Overall sample
Hmisc::describe(dplyr::filter(data,
```

```
                        !is.na(E_score))$E_score)
sd(dplyr::filter(data,
              !is.na(E_score))$E_score)


##Analytical sample
Hmisc::describe(dplyr::filter(data,
                        !is.na(B_score) &
                          !is.na(Settingtype) &
                          !is.na(SettingRegion) &
                          !is.na(E_score) &
                          !is.na(SettingID))$E_score)
sd(dplyr::filter(data,
              !is.na(B_score) &
                !is.na(Settingtype) &
                !is.na(SettingRegion) &
                !is.na(E_score) &
                !is.na(SettingID))$E_score)


#And by treatment and control groups:
Hmisc::describe(dplyr::filter(data,  #treatment
                        treatment == 1 &
                          !is.na(B_score) &
                          !is.na(Settingtype) &
                          !is.na(SettingRegion) &
                          !is.na(E_score) &
                          !is.na(SettingID))$E_score)

sd(dplyr::filter(data, #treatment
              treatment == 1 &
                !is.na(B_score) &
                !is.na(Settingtype) &
                !is.na(SettingRegion) &
                !is.na(E_score) &
                !is.na(SettingID))$E_score)

Hmisc::describe(dplyr::filter(data,  #control
                        treatment == 0 &
                          !is.na(B_score) &
                          !is.na(Settingtype) &
                          !is.na(SettingRegion) &
                          !is.na(E_score) &
                          !is.na(SettingID))$E_score)
sd(dplyr::filter(data, #control
              treatment == 0 &
                !is.na(B_score) &
                !is.na(Settingtype) &
                !is.na(SettingRegion) &
                !is.na(E_score) &
                !is.na(SettingID))$E_score)

################################################
################################################
```

```
##################################################


# Section B: Primary outcome model

### 1. Run the multi-level model.

pr_model <- lmer(E_score ~
                   treatment +
                   B_score +
                   SettingRegion +
                   Settingtype +
                   (1 | SettingID),
                 data=data, REML = FALSE)

summary(pr_model) #Produce the results

performance::icc(pr_model) #ICC

#########

### 2. Testing OLS assumptions

#Residual diagnostics: Testing normality of residuals OLS assumption
resid_pr <- resid(pr_model) #Create object which stores the residuals of the model
plot_resid_pr <- plot(density(resid_pr)) #Kernel density plot to explore normality
shapiro.test(resid_pr) #Shapiro-wilk test for normality. !FLAG potentially K-S test better with larger sample sizes!
(p<0.05 rejects H0 that data is normally distributed)
ks.test(resid_pr, "pnorm") #K-S test (p<0.05 rejects H0 that data is normally distributed

##Both of these tests reject H0

# QQ line
qq_line <- ggplot(data = data.frame(resid = resid_pr), aes(sample = resid)) +
  stat_qq() +
  stat_qq_line() +
  labs(title = "Q-Q Plot of Residuals", x = "Theoretical Quantiles", y = "Sample Quantiles") +
  theme_minimal()


#Plotting residual density against a normal distribution for comparison. NB this is using the residual objects created
from the primary model above
pr_resid_df <- as.data.frame(resid_pr)
sd_resid <- sd(resid_pr)
max_resid <- max(resid_pr)
min_resid <- min(resid_pr)

resid_pr_kd<- ggplot(pr_resid_df, aes(x = resid_pr)) +
  geom_density(aes(color = "Kernel density"), linewidth = 1) +
  stat_function(aes(color = "Normal density"),
                fun = function(x) dnorm(x, mean = 0, sd = sd_resid),
                linetype = "dotted", linewidth = 1) +
  theme_minimal() +
```

```
  scale_x_continuous(limits = c(min_resid, max_resid)) +
  scale_color_manual(name = "Legend", values = c("Kernel density" = "#000080", "Normal density" = "red")) +
  labs(
    title = "Early Numeracy - Residuals Density Plot",
    x = "Residuals",
    y = "Density"
  )

#Testing linearity and assumption that residual errors have a mean of 0
plot(pr_model, col = "red") #We want the line here to be horizontal and at 0


### Figures to export:
resid_pr_kd
qq_line

plots <- list(resid_pr_kd, qq_line)

### Export graphs to word

# Create a Word document
doc <- read_docx()

# Save each plot as an image and add to the Word document
for (i in seq_along(plots)) {
  # Export the plot to a temporary file
  file_name <- tempfile(fileext = ".png")
  ggplot2::ggsave(filename = file_name, plot = plots[[i]])

  # Add the plot to the Word document
  doc <- doc %>%
    body_add_par(value = paste("Plot", i), style = "heading 1") %>%
    body_add_img(src = file_name, width = 6, height = 4)
}

# Save the document
print(doc, target = "//Nevis/RE/Projects/In_Progress/022807.014_P2899_EEF_The_ONE/WIP/09. Impact/Graphs &
Summary Tables/primary_outcome_exmathschampions.docx")

##########


### 3.  Bootstrapping CIs and p-value
#+ The QQ residual plot, Shapiro-Wilk, and K-S test
#+ results suggest that the residuals are not normally distributed. We will
#+ therefore need to re-estimate CIs and p-values using bootstrapping.

set.seed(999)
boot_pr_model <- boot.pval::boot_summary(pr_model,
                                         type = "norm",
                                         method = NULL,
                                         conf.level = 0.95)
boot_pr_model
```

```
####################


### 4.  Effect-size estimation

#Create Hedges g function.
hedges.g <- function(c, n, m, v, w){
  c/sqrt((((n - 1)*v)+((m - 1)* w))/(n + m-2))
}



#We use the output from the model above to calculate the effect size: (NB: This needs to be done for confidence intervals too!)
#We need to define the objects needed to be inputted into the Hedges G function created above (c, n, m, v, & w)

coefs <- data.frame(summary(pr_model)$coefficients) #create data frame of all coefficients from primary outcome model

c <- coefs["treatment", "Estimate"] #Extract the treatment coefficient

n <- nrow(dplyr::filter(data,
                        treatment == 1 &
                          !is.na(E_score) &
                          !is.na(B_score) &
                          !is.na(SettingRegion) &
                          !is.na(Settingtype) &
                          !is.na(SettingID))) #Number of individuals in treatment group in the model

m <- nrow(dplyr::filter(data,
                        treatment == 0 &
                          !is.na(E_score) &
                          !is.na(B_score) &
                          !is.na(SettingRegion) &
                          !is.na(Settingtype) &
                          !is.na(SettingID))) #Number of individuals in control group in the model

v <- var(dplyr::filter(data,
                       treatment == 1 &
                         !is.na(B_score) &
                         !is.na(SettingRegion) &
                         !is.na(Settingtype) &
                         !is.na(E_score) &
                         !is.na(SettingID))$E_score) #Variance in outcome among treatment group

w <- var(dplyr::filter(data,
                       treatment == 0 &
                         !is.na(B_score) &
                         !is.na(SettingRegion) &
                         !is.na(Settingtype) &
                         !is.na(E_score) &
                         !is.na(SettingID))$E_score)  #Variance in outcome among control group
```

```
primary_effect_size <- hedges.g(c, n, m, v, w)
primary_effect_size

#LOWER CI EFFECT SIZE using lower 95% CI from the bootstrapped model:

#We redefine c to equal the treatment coefficient minus the SE created above
c <- boot_pr_model$Lower.bound[2]

pr_effect_size_low <- hedges.g(c, n, m, v, w) #Re-run hedges G on lower CI treatment coefficient
pr_effect_size_low #This is the lower confidence interval for the treatment effect on the following directions subtest.



#UPPER CI EFFECT SIZE using bootrapped 95% upper CI from the above model:

#We redefine c again to equal the treatment coefficient plus the SE created above. Remember c has already been
defined to be the lower confidence interval, we therefore need to add 2 SEs to it.
c <- boot_pr_model$Upper.bound[2]

pr_effect_size_high <- hedges.g(c, n, m, v, w) #Re-run hedges G on upper CI treatment coefficient
pr_effect_size_high #This is the upper confidence interval for the treatment effect on the following directions subtest.

## P-value
p_values <- boot_pr_model$p.value[2]
p_value <- round(p_values, digits = 2)


### 5. Creating primary analysis table (to paste output into report)

##+ This part of the code is to create a formatted output table which can just be copied
##+ into the table empty table in word

## Find missing numbers for model (number of obs with outcome but missing covariates)

# Treatment
n_mis <- nrow(dplyr::filter(data,
                            treatment == 1 &
                              !is.na(E_score) &
                              (is.na(B_score) |
                                 is.na(SettingRegion) |
                                 is.na(Settingtype) |
                                 is.na(SettingID))))

# Control
m_mis <- nrow(dplyr::filter(data,
                            treatment == 0 &
                              !is.na(E_score) &
                              (is.na(B_score) |
                                 is.na(SettingRegion) |
                                 is.na(Settingtype) |
                                 is.na(SettingID))))

# Paste in numbers (non-missing and missing) for table
```

```
num_t <- paste0(as.character(n), " (", as.character(n_mis), ")")
num_c <- paste0(as.character(m), " (", as.character(m_mis), ")")

# Total numbers
total_n <- n + m
total_n_t_c <- paste0(as.character(total_n), " (", n, ";", m, ")")

## Generate means and CIs

# Treatment
me_t <- t.test(dplyr::filter(data,
                        treatment == 1 &
                          !is.na(E_score) &
                          !is.na(B_score) &
                          !is.na(SettingRegion) &
                          !is.na(Settingtype) &
                          !is.na(SettingID))$E_score) #Run t-test on outcome in analytical sample

mean_t <- as.character(round(me_t$estimate, digits = 2)) #extract mean from t-test object (me), to 2dp and as character to facilitate pasting in CIs.

me_ci_t <- as.character(round(me_t$conf.int[c(1, 2)], digits = 2)) #extract CIs (95%) from t-test object (me), to 2dp and as character to facilitate pasting

mean_ci_t <- paste0(mean_t, " (", me_ci_t[1], " - ", me_ci_t[2], ")") #Paste CIs in brackets next to mean figure

# Control
me_c <- t.test(dplyr::filter(data,
                        treatment == 0 &
                          !is.na(E_score) &
                          !is.na(B_score) &
                          !is.na(SettingRegion) &
                          !is.na(Settingtype) &
                          !is.na(SettingID))$E_score)

mean_c <- as.character(round(me_c$estimate, digits = 2))
me_ci_c <- as.character(round(me_c$conf.int[c(1, 2)], digits = 2))
mean_ci_c <- paste0(mean_c, " (", me_ci_c[1], " - ", me_ci_c[2], ")")


## Effect size with CIs
ES_CI <- paste0(as.character(round(primary_effect_size, digits = 3)),
            " (", as.character(round(pr_effect_size_low, digits = 2)),
            " -- ",
            as.character(round(pr_effect_size_high, digits = 2)),
            ")")

## Create table

primary_analysis_table <- t(as.data.frame(c(num_t, mean_ci_t, num_c, mean_ci_c, total_n_t_c, ES_CI, p_value)))
colnames(primary_analysis_table) <- c("N (intervention)", "Mean (intervention", "N (control)", "Mean (control)", "Total (T;C)", "Hedges g (95% CIs)", "p-value")

rownames(primary_analysis_table) <- "EYTN2 Numeracy"
```

```
## Display results - TABLE 1
View(primary_analysis_table)

###########

### 6. Creating effect size estimation table. Most of the required fields have already been created

u_m <- as.character(round((mean(dplyr::filter(data, #Unadjusted difference in means
                                      treatment == 1 &
                                        !is.na(B_score) &
                                        !is.na(SettingRegion) &
                                        !is.na(Settingtype) &
                                        !is.na(SettingID))$E_score, na.rm = TRUE)) - (mean(dplyr::filter(data,
                                                                                        treatment == 0 &
                                                                                          !is.na(B_score) &
                                                                                          !is.na(SettingRegion) &
                                                                                          !is.na(Settingtype) &
                                                                                          !is.na(SettingID))$E_score, na.rm = TRUE)),
digits = 2))

c <- round(coefs["treatment", "Estimate"], digits = 2) #Adjusted difference in means

var_t <- round(v, digits = 2) #Variance of outcome in treatment

var_c <- round(w, digits = 2) #Variance of outcome in control

pooled_var <- round(((((n - 1)*v)+((m - 1)* w))/(n + m - 2)), digits = 2) #pooled variance



# Create table 2
pr_es_est <- t(as.data.frame(c(u_m, c, num_t, var_t, num_c, var_c, pooled_var)))
colnames(pr_es_est) <- c("Unadj. diff in means", "Adj diff in means", "N (treatment)", "Variance (treatment)", "N
(control)", "Variance (control)", "Pooled Variance")
rownames(pr_es_est) <- "EYTN Numeracy Test"
View(pr_es_est)


######################  The ONE Analysis: Additional compliance analysis ###################


##############################


# Clear workspace

rm(list=ls())

set.seed(999) #NB this is the same seed set in the randomisation code at the start of the project so we shall
retain it here for consistency.


# Load packages
```

```
library("haven")
library("dplyr")
library("ggplot2")
library("lme4")
library("sjstats")
library("Hmisc")
library("performance")
library("lmtest")
library("eeptools")
library("lmerTest")
library("officer")
library("boot")
library("boot.pval")
library("PowerUpR")

# Load cleaned data
data <- read_dta("//Nevis/RE/Projects/In_Progress/022807.014_P2899_EEF_The_ONE/WIP/06. Data collection/02. Endline/02. Data_NotMirrored/BaselineEndlineComplianceClean2.dta")

data <- filter(data, !is.na(SettingID))
data <- filter(data, endlineonly != 1)

#Dropping non-compliant setting
table(dplyr::filter(data,
        treatment == 1)$`_Practitioner_Consistently_Pres`)
table(filter(data, `_Practitioner_Consistently_Pres`==0)$SettingID)

data<-filter(data, SettingID != "S123")
table(dplyr::filter(data,
        treatment  ==  1)$`_Practitioner_Consistently_Pres`)

##################################################
```

```
# Section A: Descriptive statistics


#1. Means, SDs, Min and Max (for whole sample and analytical sample)


#Overall sample
Hmisc::describe(dplyr::filter(data,
                  !is.na(E_score))$E_score)
sd(dplyr::filter(data,
           !is.na(E_score))$E_score)



##Analytical sample
Hmisc::describe(dplyr::filter(data,
                  !is.na(B_score)  &
                   !is.na(Settingtype) &
                   !is.na(SettingRegion) &
                   !is.na(E_score)  &
                   !is.na(SettingID))$E_score)
sd(dplyr::filter(data,
           !is.na(B_score) &
            !is.na(Settingtype)  &
            !is.na(SettingRegion) &
            !is.na(E_score) &
            !is.na(SettingID))$E_score)



#And by treatment and control groups:
Hmisc::describe(dplyr::filter(data, #treatment
                  treatment == 1 &
                   !is.na(B_score)  &
                   !is.na(Settingtype)  &
```

```
                !is.na(SettingRegion) &
                !is.na(E_score)  &
                !is.na(SettingID))$E_score)

sd(dplyr::filter(data, #treatment
          treatment == 1 &
           !is.na(B_score) &
           !is.na(Settingtype)  &
           !is.na(SettingRegion) &
           !is.na(E_score) &
           !is.na(SettingID))$E_score)

Hmisc::describe(dplyr::filter(data, #control
                 treatment == 0 &
                  !is.na(B_score)  &
                  !is.na(Settingtype)  &
                  !is.na(SettingRegion) &
                  !is.na(E_score)  &
                  !is.na(SettingID))$E_score)
sd(dplyr::filter(data, #control
          treatment == 0 &
           !is.na(B_score) &
           !is.na(Settingtype)  &
           !is.na(SettingRegion) &
           !is.na(E_score) &
           !is.na(SettingID))$E_score)

#####################################################
#####################################################
#####################################################
```

```
# Section B: Primary outcome model

### 1. Run the multi-level model.

pr_model <- lmer(E_score ~
                 treatment +
                 B_score +
                 SettingRegion +
                 Settingtype +
                 (1 | SettingID),
                data=data, REML = FALSE)

summary(pr_model) #Produce the results

performance::icc(pr_model, ci = 0.95) #ICC

#########

### 2. Testing OLS assumptions

#Residual diagnostics: Testing normality of residuals OLS assumption

resid_pr <- resid(pr_model) #Create object which stores the residuals of the model

plot_resid_pr <- plot(density(resid_pr)) #Kernel density plot to explore normality

shapiro.test(resid_pr) #Shapiro-wilk test for normality. !FLAG potentially K-S test better with larger sample sizes! (p<0.05 rejects H0 that data is normally distributed)

ks.test(resid_pr, "pnorm") #K-S test (p<0.05 rejects H0 that data is normally distributed

##Both of these tests reject H0

# QQ line

qq_line <- ggplot(data = data.frame(resid = resid_pr), aes(sample = resid)) +
```

```
 stat_qq() +
 stat_qq_line() +
 labs(title = "Q-Q Plot of Residuals", x = "Theoretical Quantiles", y = "Sample Quantiles") +
 theme_minimal()
```

```
#Plotting residual density against a normal distribution for comparison. NB this is using the residual objects created from the primary model above
pr_resid_df <- as.data.frame(resid_pr)
sd_resid <- sd(resid_pr)
max_resid <- max(resid_pr)
min_resid <- min(resid_pr)

resid_pr_kd<- ggplot(pr_resid_df, aes(x = resid_pr)) +
 geom_density(aes(color = "Kernel density"), linewidth = 1) +
 stat_function(aes(color = "Normal density"),
          fun = function(x) dnorm(x, mean = 0, sd = sd_resid),
          linetype = "dotted", linewidth = 1) +
 theme_minimal() +
 scale_x_continuous(limits = c(min_resid, max_resid)) +
 scale_color_manual(name = "Legend", values = c("Kernel density" = "#000080", "Normal density" = "red")) +
 labs(
  title = "Early Numeracy - Residuals Density Plot",
  x = "Residuals",
  y = "Density"
 )

#Testing linearity and assumption that residual errors have a mean of 0
plot(pr_model, col = "red") #We want the line here to be horizontal and at 0

### Figures to export:
```

```
resid_pr_kd

qq_line


plots <- list(resid_pr_kd, qq_line)


### Export graphs to word


# Create a Word document

doc <- read_docx()


# Save each plot as an image and add to the Word document

for (i in seq_along(plots)) {

 # Export the plot to a temporary file

 file_name <- tempfile(fileext = ".png")

 ggplot2::ggsave(filename = file_name, plot = plots[[i]])


 # Add the plot to the Word document

 doc <- doc %>%

  body_add_par(value = paste("Plot", i), style = "heading 1") %>%

  body_add_img(src = file_name, width = 6, height = 4)

}


# Save the document

print(doc, target = "//Nevis/RE/Projects/In_Progress/022807.014_P2899_EEF_The_ONE/WIP/09.
Impact/Graphs & Summary Tables/primary_outcome.docx")


##########



### 3.  Bootstrapping CIs and p-value

#+ The QQ residual plot, Shapiro-Wilk, and K-S test

#+ results suggest that the residuals are not normally distributed. We will
```

```
#+ therefore need to re-estimate CIs and p-values using bootstrapping.


set.seed(999)
boot_pr_model <- boot.pval::boot_summary(pr_model,
                          type = "norm",
                          method = NULL,
                          conf.level = 0.95)
boot_pr_model



######################


### 4.  Effect-size estimation


#Create Hedges g function.
hedges.g <- function(c, n, m, v, w){
  c/sqrt((((n - 1)*v)+((m - 1)* w))/(n + m - 2))
}




#We use the output from the model above to calculate the effect size: (NB: This needs to be done for confidence intervals too!)
#We need to define the objects needed to be inputted into the Hedges G function created above (c, n, m, v, & w)


coefs <- data.frame(summary(pr_model)$coefficients) #create data frame of all coefficients from primary outcome model


c <- coefs["treatment", "Estimate"] #Extract the treatment coefficient


n <- nrow(dplyr::filter(data,
```

```
                treatment == 1 &
                 !is.na(E_score) &
                 !is.na(B_score) &
                 !is.na(SettingRegion) &
                 !is.na(Settingtype)  &
                 !is.na(SettingID))) #Number of individuals in treatment group in the model

m <- nrow(dplyr::filter(data,
                treatment == 0 &
                 !is.na(E_score) &
                 !is.na(B_score) &
                 !is.na(SettingRegion) &
                 !is.na(Settingtype)  &
                 !is.na(SettingID))) #Number of individuals in control group in the model

v <- var(dplyr::filter(data,
                treatment == 1 &
                 !is.na(B_score) &
                 !is.na(SettingRegion) &
                 !is.na(Settingtype)  &
                 !is.na(E_score) &
                 !is.na(SettingID))$E_score) #Variance in outcome among treatment group

w <- var(dplyr::filter(data,
                treatment == 0 &
                 !is.na(B_score) &
                 !is.na(SettingRegion) &
                 !is.na(Settingtype)  &
                 !is.na(E_score) &
                 !is.na(SettingID))$E_score)  #Variance in outcome among control group

primary_effect_size <- hedges.g(c, n, m, v, w)
```

```
primary_effect_size

#LOWER CI EFFECT SIZE using lower 95% CI from the bootstrapped model:

#We redefine c to equal the treatment coefficient minus the SE created above

c <- boot_pr_model$Lower.bound[2]

pr_effect_size_low <- hedges.g(c, n, m, v, w) #Re-run hedges G on lower CI treatment coefficient

pr_effect_size_low #This is the lower confidence interval for the treatment effect on the following directions subtest.

#UPPER CI EFFECT SIZE using bootrapped 95% upper CI from the above model:

#We redefine c again to equal the treatment coefficient plus the SE created above. Remember c has already been defined to be the lower confidence interval, we therefore need to add 2 SEs to it.

c <- boot_pr_model$Upper.bound[2]

pr_effect_size_high <- hedges.g(c, n, m, v, w) #Re-run hedges G on upper CI treatment coefficient

pr_effect_size_high #This is the upper confidence interval for the treatment effect on the following directions subtest.

## P-value

p_values <- boot_pr_model$p.value[2]

p_value <- round(p_values, digits = 2)

### 5. Creating primary analysis table (to paste output into report)

##+ This part of the code is to create a formatted output table which can just be copied

##+ into the table empty table in word
```

```
## Find missing numbers for model (number of obs with outcome but missing covariates)


# Treatment
n_mis <- nrow(dplyr::filter(data,
                    treatment == 1 &
                     !is.na(E_score) &
                     (is.na(B_score) |
                     is.na(SettingRegion) |
                     is.na(Settingtype) |
                     is.na(SettingID))))


# Control
m_mis <- nrow(dplyr::filter(data,
                    treatment == 0 &
                     !is.na(E_score) &
                     (is.na(B_score) |
                     is.na(SettingRegion) |
                     is.na(Settingtype) |
                     is.na(SettingID))))


# Paste in numbers (non-missing and missing) for table
num_t <- paste0(as.character(n), " (", as.character(n_mis), ")")
num_c <- paste0(as.character(m), " (", as.character(m_mis), ")")


# Total numbers
total_n <- n + m
total_n_t_c <- paste0(as.character(total_n), " (", n, ";", m, ")")


## Generate means and CIs


# Treatment
me_t <- t.test(dplyr::filter(data,
```

```
                treatment == 1 &
                 !is.na(E_score) &
                 !is.na(B_score)  &
                 !is.na(SettingRegion) &
                 !is.na(Settingtype)  &
                 !is.na(SettingID))$E_score) #Run t-test on outcome in analytical sample


mean_t <- as.character(round(me_t$estimate, digits = 2)) #extract mean from t-test object (me), to 2dp and as character to facilitate pasting in CIs.


me_ci_t <- as.character(round(me_t$conf.int[c(1, 2)], digits = 2)) #extract CIs (95%) from t-test object (me), to 2dp and as character to facilitate pasting


mean_ci_t <- paste0(mean_t, " (", me_ci_t[1], " - ", me_ci_t[2], ")") #Paste CIs in brackets next to mean figure


# Control
me_c <- t.test(dplyr::filter(data,
                treatment == 0 &
                 !is.na(E_score) &
                 !is.na(B_score)  &
                 !is.na(SettingRegion) &
                 !is.na(Settingtype)  &
                 !is.na(SettingID))$E_score)


mean_c <- as.character(round(me_c$estimate, digits = 2))
me_ci_c <- as.character(round(me_c$conf.int[c(1, 2)], digits = 2))
mean_ci_c <- paste0(mean_c, " (", me_ci_c[1], " - ", me_ci_c[2], ")")



## Effect size with CIs
ES_CI <- paste0(as.character(round(primary_effect_size, digits = 3)),
          " (", as.character(round(pr_effect_size_low, digits = 2)),
          " -- ",
```

```
        as.character(round(pr_effect_size_high, digits = 2)),
        ")")


## Create table


primary_analysis_table <- t(as.data.frame(c(num_t, mean_ci_t, num_c, mean_ci_c, total_n_t_c, ES_CI,
p_value)))
colnames(primary_analysis_table) <- c("N (intervention)", "Mean (intervention", "N (control)", "Mean (control)",
"Total (T;C)", "Hedges g (95% CIs)", "p-value")


rownames(primary_analysis_table) <- "EYTN2 Numeracy"


## Display results - TABLE 1
View(primary_analysis_table)


############


### 6. Creating effect size estimation table. Most of the required fields have already been created


u_m <- as.character(round((mean(dplyr::filter(data, #Unadjusted difference in means
                               treatment == 1 &
                                !is.na(B_score) &
                                !is.na(SettingRegion) &
                                !is.na(Settingtype) &
                                !is.na(SettingID))$E_score, na.rm = TRUE)) - (mean(dplyr::filter(data,
                                                                                  treatment == 0 &
                                                                                   !is.na(B_score) &
                                                                                   !is.na(SettingRegion) &
                                                                                   !is.na(Settingtype) &
                                                                                   !is.na(SettingID))$E_score, na.rm = TRUE)), digits
= 2))


c <- round(coefs["treatment", "Estimate"], digits = 2) #Adjusted difference in means
```

```
var_t <- round(v, digits = 2) #Variance of outcome in treatment


var_c <- round(w, digits = 2) #Variance of outcome in control


pooled_var <- round((((n - 1)*v)+((m - 1)* w))/(n + m - 2), digits = 2) #pooled variance



# Create table 2

pr_es_est <- t(as.data.frame(c(u_m, c, num_t, var_t, num_c, var_c, pooled_var)))

colnames(pr_es_est) <- c("Unadj. diff in means", "Adj diff in means", "N (treatment)", "Variance (treatment)", "N (control)", "Variance (control)", "Pooled Variance")

rownames(pr_es_est) <- "EYTN Numeracy Test"

View(pr_es_est)
```

# Appendix O: Privacy Notice

## Privacy Notice

For parents of children in settings participating in The ONE evaluation
Updated 12/01/2024

### Project Title

The ONE

RAND Europe Ref: 23002

University of Oxford Central University Research Ethics Approval Ref: R85139/RE003

### Who we are and what does this document include?

**RAND Europe** is an independent not-for-profit research institute based in Cambridge, Brussels and the Netherlands, whose mission is to help improve policy and decision making through evidence-based research. The **University of Oxford and Sheffield University** provide world-class research and education to benefit society on a local, regional, national and global scale.

This document provides information on the evaluation of The ONE programme. It sets out how your child's personal information will be collected, used and looked after in accordance with the UK General Data Protection Regulations (UK GDPR) and the Data Protection Act 2018.

### About the evaluation

The ONE Project is a programme delivered by the University of Oxford and Sheffield University and is funded by both the **Education Endowment Foundation (EEF)** and the **Department for Education (DfE)** as part of their Stronger Practice Hubs initiative. The aim of the programme is to support Early Years practitioners to run play-based maths activities that support maths development by embedding executive function skills into maths learning. The programme will be rolled out across approximately 150 settings.

To see if The ONE influences learning, the EEF is also funding an evaluation and has commissioned RAND Europe to conduct the evaluation. You are receiving this privacy notice because your child attends one of the settings taking part in The ONE evaluation.

### Who is holding your data?

RAND Europe will act as Data Controller during the evaluation (from January 2023 until publication of the evaluation report in early 2025). Before the programme is evaluated, the University of Oxford (with the University of Sheffield) will act as Data Processors and collect data on behalf of RAND Europe (see details below). QA Research will collect data on your child's maths and executive functioning skills using short, child-friendly assessments on behalf of RAND Europe.

After the evaluation report is published (early 2025), RAND Europe will transfer the data by secure cloud to the EEF where it will be stored in the EEF data archive, and EEF will become Data Controller. Data collected as part of all EEF evaluations are archived for research purposes. The EEF are the data controller for the data archive which is

managed by FFT Education as a data processor working on EEF's behalf. At the end of the evaluation, RAND Europe will submit the data from the evaluation (child assessment scores and some personal data of children, like full names, dates of birth, Unique Pupil Numbers) directly to FFT through a secure portal, such as Egress.

In the future, researchers might link the data deposited in the EEF's data archive to information held in the Department for Education's (DfE) National Pupil Database and other datasets. To do so, FFT Education would use the direct identifiers of children (like names, dates of birth) to request Pupil Matching References (PMR) from the DfE, who would transfer PMRs directly to the Office for National Statistics' Secure Research Space (SRS). FFT Education would use the SRS to match the data from the evaluation (i.e., child assessment scores) with the PMRs. This process ensures that future researchers will be able to link data from the evaluation to the National Pupil Database and other datasets without accessing direct identifiers of the children.

After the evaluation report is published, RAND Europe will also share all evaluation data from the project with the University of Oxford securely via Egress in order for them to collect follow-up data. During this period the EEF and the University of Oxford will be independent controllers.

Finally, once the data has been collected, analysed and published in scientific articles and doctoral theses by the University of Oxford, the EEF will become the final Data Controller.

### What data are we collecting and how we are collecting them?

By 'personal data', we mean any information about an individual from which that individual can be identified.

QA Research will email settings to send the following personal data on children at settings to RAND Europe via Egress: name, date of birth, setting postcode, number of hours attendance and attendance patterns (i.e., which days and times your child attends nursery), eligibility for Early Years Pupil Premium (EYPP), and whether your child speaks English as an additional language, because this will help RAND Europe evaluate whether The ONE benefits young children's numeracy more or less in relation to their attendance at school, economic disadvantage, and speaking English as an additional language. Name, setting postcode and date of birth will help EEF and the University of Oxford evaluate whether The ONE has benefits that are sustained in reception and later in primary school.

Before the intervention, and at the end of the project, RAND Europe will ask QA Research to administer the 'Early Years Toolbox Early Numeracy' and two executive functions assessments to children. The tests are child friendly and have been used in a number of other evaluations. They will help us understand whether The ONE has an effect on children's maths and executive functioning skills.

### How we share and store the data?

Data collected by RAND Europe and QA research will be transferred using Egress – a secure file sharing platform – and will be stored securely on the RAND Europe server in a password protected folder, accessible only to members of the research project. RAND Europe will use this data at an aggregate level (i.e., not individually) to analyse the effects of The ONE and to write a report that will be publicly available. We will not identify any individuals in our report and will ensure that presented data does not provide information on small cohorts of children so the risk that any child can be identified is minimised.

QA Research will administer the assessments which will then be stored on the Early Years Toolbox (EYT) secure servers in the EU or on QA Research's internal GDPR-compliant clouds and servers. Where possible, when in transit to settings, QA test administrators will store personal data in a GDPR-compliant cloud or locally on secure, protected tablets. Occasionally, child data may be stored in paper copy during assessors transit to the setting, where digital or

cloud access may be limited. Paper copies will then be retained in the setting during the course of evaluation, and left with the setting for destruction upon completion of testing. QA staff are trained in data protection and processing.

RAND Europe will also share the evaluation data with the University of Oxford via Egress or OneDrive The University of Oxford will then contact parents to ask permission to follow their children's numeracy outcomes in school. This additional data will allow long-term follow-up which will help us understand if The ONE is effective in the longer term.

The EEF will then become the final data controllers.

### What is the legal basis for processing your data?

The legal basis for RAND Europe to process your child's personal data is legitimate interests detailed in Article 6(1)(f) of the UK GDPR. The legal bases for processing your child's special category data is for reasons of substantial public interest and because it is necessary for archiving purposes in the public interest, scientific or historical research purposes as detailed in Article 9(2)(g)& (j) respectively of the UK GDPR. To ensure that all processing is fair and lawful, RAND Europe have also completed a Legitimate Interest Assessment and a Data Protection Impact Assessment and completed an application to the RAND internal review board for ethical approval. RAND Europe will process only what is required to meet these legal bases and will ensure security and safeguards are in place to protect the information.

After the evaluation report is published, the University of Oxford will become the data controller with respect to your personal data, and as such will determine how your child's personal data is used in the research. The University will process your child's personal data for the purpose of the research outlined above. Research is a task that we perform in the public interest. Further information about your rights with respect to your personal data is available from https://compliance.web.ox.ac.uk/individual-rights . The EEF will also become a data controller once the evaluation report is published for the purposes of archiving the trial data with FFT.

### What are we using the data for?

The evaluation team is collecting your child's data to aid the evaluation. The evaluation aims to find out more about the impact that The ONE intervention has on a number of pupil outcomes including maths attainment and executive functioning. We will look at how the outcomes of children taking part in The ONE in the 2023/2024 academic year compare to those of children who do not take part in The ONE in the same year. We will analyse the data to see if there is a difference. This will help us to understand if and how the ONE intervention makes a difference to children.

### How do we keep your data secure?

The evaluation team have put various security measures in place to keep personal data secure and to prevent any unauthorised access to or use of it in accordance with the Data Protection Act (2018) and UK GDPR requirements. All data collected by RAND Europe (or by QA Research on behalf of RAND Europe) will be stored on secure servers, accessed only by relevant project team. All data will be shared using specialist software (Egress). No data will be saved on servers or shared with processors outside the UK and the EU. The Department for Education will keep the data it processes securely and strictly in accordance with Government standards.

### How long do we keep your data?

RAND Europe and QA Research will securely delete all data held on its secure server six months after the end of their involvement in the project.

Oxford will hold the data with the highest level of security in encrypted and password protected storage. Data will be deleted after September 2028, to allow time for doctoral thesis completion. This time duration is dependent on the infrequent occasions in which reviewers or readers of publications ask researchers to re-analyse data in a different way.

Data about all pupils will be deposited into the EEF's data archive at the end of the project for the purpose of research. You can find out more about the archive in the Project Information Sheet.

## What are your rights?

RAND Europe, the University of Oxford, and the Education Endowment Foundation operate in accordance with the Data Protection Act 2018 and UK GDPR 2016 requirements. You are provided with certain rights that you may have the right to exercise through us. In summary those rights are:

- To access your data ("data subject access request") (Article 15 of the GDPR)
- To have inaccurate personal data rectified (Article 16 of the GDPR)
- To have your data erased (Article 17 of the GDPR)
- To restrict the processing of your data (Article 18 of the GDPR)
- Request the transfer of your personal data to you or to a third party (Article 20 of the GDPR)
- Object to processing of your personal data (Article 21 of the GDPR).

## How do you contact us?

- Before early 2025 you should contact RAND Europe, by:
    - Sending an email to the evaluation team at TheONE@randeurope.org
    - Contacting RAND Europe's Data Protection officer, Rani Viknaraja rviknara@randeurope.org and quoting Ref: 23002 "The ONE" or in writing to Data Protection Officer, RAND Europe, Eastbrook House, Shaftesbury Road, Cambridge, CB2 8BF
- From early 2025 until 2028, you should contact the University of Oxford, by:
    - Sending an e-mail to the delivery team at TheONE@psy.ox.ac.uk

If you would like to withdraw your child from the intervention, you must complete the attached opt-out form. If you would like to withdraw your child from the evaluation, you must complete the attached withdrawal form and return it to your setting. This will mean that your child's data will not be collected or will be deleted if already collected. If already collected, the evaluation team will erase their data. Withdrawing the data collection will have no negative repercussions. We will maintain a log to store any such requests that we receive.

Where you request information from us, we will need to confirm your identity to ensure the security of your data. We will endeavour to respond within 30 days but our response time may vary depending on the complexity of your request. A fee will not normally be charged unless a request is considered to be without basis, repetitive or excessive. Where we request a fee it shall always be reasonable.

You also have the right to contact the Information Commissioner's Office in the UK if you have any concerns about the processing of your data by RAND Europe or the University of Oxford. You can visit ico.org.uk or email casework@ico.org.uk. Alternatively, you can write or telephone to Information Commissioner's Office, Wycliffe House. Water Lane, Wilmslow. Cheshire SK9 5AF.

# Appendix P: Participant Information Sheet

**The ONE Project**

**Information for Parents / Guardians of Children**

This document is for parents of children who attend Early Years settings (many of which will be preschools) that are taking part in a study of the ONE (Orchestrating Numeracy and the Executive) programme. The ONE programme involves training for preschool staff, and guided play activities delivered by preschool staff. The aim of the ONE programme is to support children's numeracy and thinking skills. In this study, we want to find out if the ONE programme improves these skills.

**What will happen if my child takes part?**

Staff in your child's preschool will be trained to run 25 short and child-friendly games that aim to support children's numeracy and thinking skills, over a 12-week period. At the start of the study, and again at the end of the study, up to 15 children in each preschool will be asked to play some games to assess their maths and thinking skills. Some of the games are played on a tablet, some of the games involve objects. The games are designed to be fun, and child friendly. In total, the games will take around 30 minutes. The assessment will be led by trained members of Qa research. All Qa Research assessors will hold a current Disclosure and Barring Service (DBS) certificate and will never be alone with your child. The assessment will take place in a quiet space within the preschool classroom, either during class time or break time. Children will not be pressured to take part if they do not want to.

Children's scores on the assessment games will be averaged together to look at whether the ONE programme leads to improvements in numeracy and thinking skills across a large group of children. Scores will be compared between children in preschools that have implemented the ONE programme with those that have not yet received the ONE. In order to see if there has been an improvement, assessment scores will be analysed alongside some personal data about your child. Qa research will collect your child's name and date of birth; your child's eligibility for Early Years Pupil Premium (EYPP) and whether they speak English as an additional language; the preschool postcode, and the days and times your child attends preschool.

This information will help us work out who the intervention works best for. All data for each child will be kept strictly confidential. We will never publish personal information about your child, and we will not be making assessments about individual children based on their scores. Your child's performance will not be used to make judgements about your or their education or access to services. While lower executive functions can be linked to special educational needs, the assessment games are not diagnostic. They have been chosen to assess change over time in preschoolers with a range of abilities, and cannot be used as an individual educational assessment or to identify a diagnosis. The data analysis will be carried out by RAND Europe.

In addition, Oxford University will receive the personal data above to find out if there are longer-term benefits for children when they are in school. Finally, and only if you allow us to do so, Oxford University will contact you to understand if their numeracy and executive functions have continued to improve in school.

You can find out more about how we will collect, use and safely store your child's data from the accompanying privacy notice here.

**What happens to the results of the study?**

The research may be published in academic publications and websites. In addition, the findings will be shared in a report required by the EEF, who are funding this study. The ONE team will also share summary findings with preschools who have taken part in the study. Parents can sign up to find out about the results of the study by emailing theONE@psy.ox.ac.uk. Part of this work will be written up as a student's doctoral thesis. It will be deposited both in print and online in the Oxford University Research Archive, where it will be openly accessible.

Results from your child will be combined with the results from other children and will not be used in any way that would identify your child. We will not use your child's name or the name of the school in any report. We will not provide information that identifies your child to anyone outside the study team.

**Does my child need to take part in the assessment games?**
No. All children will be included in the new activities their preschool practitioners will play with them, unless they do not wish to play them. However, it is entirely up to you whether you want your child to take part in the numeracy and executive functions assessment games. Qa Research and RAND Europe will use the games to assess whether the ONE professional development and activities help children improve their numeracy and executive functions. If you or your child prefers that they do not take part in the assessment games described above, please complete the opt-out form your setting has given to you, and return by the date specified on that form. We use a parent opt-out form, rather than opt-in consent, to ensure that we understand whether our programme is helpful for children from a range of backgrounds, representative of the UK. It is important that we understand whether the programme helps children from all walks of life. There is evidence that requiring opt-in consent from parents can unduly lead to some parents and children being excluded, which would skew the research data such that it would not be usable.

Before the assessment games are started, a trained assessor from Qa Research will describe the project to your child and explain that they only have to play the games if they want to. If your child does not want to take part in those games, they do not have to.

If your child has already completed the assessment games and you decide afterwards that you no longer wish for your child's data to be part of the evaluation of The ONE programme, you can still withdraw your child's data using the data withdrawal form until September 2028. This will mean that the RAND Europe team and the ONE team will delete their data and will not use them in the evaluation of the ONE, or its follow-up.

There are no disadvantages or risks to your child from taking part in the assessment games. The evaluation of The ONE practitioner development and activities may help future children and preschool practitioners in the future. You are free to refuse or withdraw your child's involvement in the study until September 2028. Your decision will not affect your child's access to services or programmes at your child's preschool outside the research.

THE ONE
ORCHESTRATING NUMERACY AND THE EXECUTIVE

**More about The ONE Project**
The ONE is an Early Years programme involving guided play preschool activities to support the development of children's numeracy and thinking skills. The thinking skills we are aiming to support in the ONE are sometimes called executive functions. Executive functions, such as focusing attention, ignoring distractions and thinking flexibly, are important for Early Years numeracy skills even before going to school.

The ONE was co-designed by a team of researchers from the University of Oxford, University of Sheffield, Ulster University, collaborators in Australia and Canada, and Early Years colleagues in UK preschools. The ONE involves training educators to run 25 short and child-friendly games that support children's numeracy and thinking skills for 12 weeks. The activities are embedded into preschool routines such as small group activities, outdoor play, and free play. The activities are delivered in preschools by the preschool staff. Children will take part in three activities a week as part of a group, and these games will include key early maths skills including numbers, counting, ordering, patterns, and spatial awareness. The ONE was recently tested in a small-scale study that showed that educators and children liked the activities and that they also improved children's maths, but we do not yet know if this will work on a large scale.

**Who is involved in the current project and what are we trying to find out?**
The current study involves a partnership between:

- The ONE team (staff at the Universities of Oxford and Sheffield) who are delivering the professional development to preschool staff and supporting preschool staff to play the activities in their preschool;
- Researchers at RAND Europe (a not-for-profit research institute) who are responsible for the independent evaluation of the ONE professional development and activities, along with their partners Qa Research who will assess children's maths and thinking skills;
- The Education Endowment Foundation (EEF) as well as the Stronger Practice Hubs (SPH), which are providing the resources for the project.

Staff based at the University of Oxford and Sheffield University will provide the training to educators in each preschool. Educators and children will take part in activities embedded in their weekly routines. The intervention activities will be carried out by setting staff, so children will not be interacting with researchers at this point of the study (although researchers might observe a lesson).

**What are the EEF archive and the Stronger Practice Hubs? Who will access personal data?**
The Education Endowment Foundation (EEF) is a Foundation established by the Department of Education to evaluate what interventions work across preschools in England. The EEF is funding the evaluation of the ONE. As part of their work, the EEF hold a data archive. At the end of the evaluation, they will become 'controllers' of the data obtained throughout the project. Data will be securely transferred into their archive. You can find more information about the EEF archive on the EEF's website. Data collected as part of all EEF evaluations are archived for research purposes. The EEF are the data controller for the data archive which is managed by FFT Education as a data processor working on EEF's behalf. At the end of the evaluation, RAND Europe will submit the data from the evaluation (child assessment scores and some personal data of children, like full names, dates of birth, and Unique Pupil Numbers) directly to FFT Education through a secure portal.

In the future, researchers might link the data deposited in the EEF's data archive to information held in the Department for Education's (DfE) National Pupil Database and other datasets. To do so, FFT Education would use the direct identifiers of children (names and date of birth) to request Pupil Matching References (PMR) from the DfE, who would transfer PMRs directly to the Office for National Statistics' Secure Research Space (SRS). FFT Education would use the SRS to match the data from the evaluation (i.e., child assessment scores) with the PMRs. This process ensures that future researchers will be able to link data from the evaluation to the National Pupil Database and other datasets without accessing direct identifiers of the children. This means that no one who looks at the information in the EEF archive will eventually know who it related to. In the future, people can ask to use the EEF archive to carry out more studies and find out if this project has helped children. Only researchers who are approved by the EEF will be able to look at the archive.

The Stronger Practice Hubs support the project by contacting preschools about the ONE, and by covering costs for the intervention so that your preschool receives the intervention at no cost. The Stronger Practice Hub will not receive personal data.

After RAND Europe publish the report evaluating the immediate benefits of the ONE, the University of Oxford, on behalf of the ONE team, will also become 'controllers' of the data, to understand if there are long-term benefits of the intervention in primary schools. Data will be stored securely on a University approved server. Access to personal data will be restricted to the ONE Team Oxford University staff. Access may also be granted to authorised members of the University of Oxford for the purposes of monitoring and/or audit. Data will be held until September 2028 to allow for doctoral thesis completion and publication of research reports.

**What should I do next?**

It is your preschool's decision to deliver the ONE intervention, but it is your decision whether we can include your child in this research. If you **do not** wish your child to take part in this study, please fill in the opt-out form and return it to your child's preschool as soon as possible (and by the date indicated on the form). If you are happy for your child to take part, you do not need to do anything. You are free to withdraw your child at any time until September 2028, without penalty and without giving a reason, by notifying your child's preschool, RAND Europe or the University of Oxford researchers via our withdrawal form.

**Any questions?**

If you would like to discuss the research with someone beforehand (or if you have questions afterwards), please contact: **Principal Investigator:** Professor Gaia Scerif, University of Oxford; Tel: 01865 271403; Email: gaia.scerif@psy.ox.ac.uk or theONE@psy.ox.ac.uk. For further information, please see the project website here.

If you wish to make a formal complaint, please contact the Chair of the Medical Sciences Interdivisional Research Ethics Committee at the University of Oxford who will seek to resolve the matter as soon as possible: Email: ethics@medsci.ox.ac.uk ; Address: Research Services, University of Oxford, Boundary Brook House, Churchill Drive, OX3 7GB.

Please keep this information sheet for your record.

# Appendix Q: Data Withdrawal Form

## Orchestrating Numeracy and the Executive – The ONE Project

## DATA WITHDRAWAL FORM

Ethics Approval Reference: R85139/RE001

If you are happy for your child's data to contribute to this trial, to be archived by the Education Endowment Foundation, and for follow up data on numeracy to be analysed by the University of Oxford, then you do not have to do anything.

If you **do NOT** want your child's data to be used for this study or their data to be archived, please complete the form below and return to your child's school.

☐ I, the undersigned, hereby do NOT give permission for my child's data to be used for the purposes of the trial of The One.

☐ I, the undersigned, hereby do NOT give permission for my child's data to be archived by the Education Endowment Foundation.

☐ I, the undersigned, hereby do NOT give permission for my child's data to be analysed by the University of Oxford.

---

Child's full name: _____

School name: _____

Parent/guardian's name: _____

Parent/guardian's signature: _____

Date: _____

This document is available for download at https://educationendowmentfoundation.org.uk

Education Endowment Foundation

The Education Endowment Foundation
5th Floor, Millbank Tower
21–24 Millbank
London
SW1P 4QP

https://educationendowmentfoundation.org.uk
@EducEndowFoundn
Facebook.com/EducEndowFoundn