Concept Cat: A two-armed cluster randomised controlled trial

Further appendices

July 2025

Miguel Subosa, Fin Oades, Erin Dysart, Louise Tracey, Anna Brian, and Elena Rosa Speciani

The Education Endowment Foundation (EEF) 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

☎ 0207 802 1653

info@eefoundation.org.uk

www.educationendowmentfoundation.org.uk

Contents

Appendix G: Randomisation Syntax

```
**# Concept Cat randomisation code


/*


The aim of this script is to produce treatment and control assignments to settings for the Concept Cat project.

First the script sets up the variables which will be used as Strata during the randomisation.

Next the script generates two random numbers for each observation, and adds them together, to generate a randomly and uniquely assigned identifier for each setting - this is done by another researcher to ensure that randomisation is blind.

Then, after checking for duplicates in the unique identifier, the script saves the dataset, drops the information about the setting and saves the dataset again - this creates two datasets (1) containing the school information and the unique identifier and (2) containing no information about the setting but retaining variables required for randomisation and the unique identifier.

Randomisation is conducted on dataset (2) and the script checks for balance between conditions.

Finally, the script merges the school information back into dataset (2) from dataset (1) to complete the blind randomisation process and end with a list of settings and their assignment.


*/


**# Load the data


*we load in a list of settings with information required for randomisation

use "[FILEPATH REDACTED]"


**# Preparing strata variables


*check how many settings have PVI status and how many are in each region

*we should be looking for balance in these since they are the strata in the randomisation

tab settingtypepviormaintained

tab region


*settingtypepviormaintained and region are string variables and need to be numeric for the randomisation code to work properly
```

```
encode settingtypepviormaintained, gen(pviormaintained)

encode region, gen(regionnum)


**# Generating random unique identifier


*now we generate a random number for each observation - since Stata generates random numbers with replacement, we generate two random numbers and add them together to significantly reduce the change that the same number is assigned to two observations

*set the seed

set seed 991

gen double rand1 = runiform()

gen double rand2 = runiform()

gen uniqueobs = rand1+rand2

*we check that there are no duplicates (the statistical probability of getting duplicates is below 0.00001 since we generated two random numbers independently for each observation)

duplicates list uniqueobs


**# Save the dataset


*now we save the data with the uniqueobs variable including

save "[FILEPATH REDACTED]", replace


**# Prepare the dataset for randomisation


*now we drop the setting name and address, so that we can randomise based on the unique identifier

drop settingname settingaddress mainsettingcontactname mainsettingcontactemail hassentpupildemographics hasstartedtestingwithelklan hascompletedtestingwithelklan testingdatebookedwithelklan numberofpupilsegress numberofpupilstestedegress numberofpupilstestedelklan dsaforwardedtokerin dsacompletedsignedbysettingandke notes hassentdsa


*and we save the data again - this is the data we will do the randomisation on

save "[FILEPATH REDACTED]", replace


**# Run the randomisation
```

```
*we use the randtreat package to run the randomisation

ssc install randtreat, replace

*we use randtreat to generate a new variable "treatment" which is stratified by eia status (equal (non-)eia in treatment and control), creating two different groups (multiple(2)) and using a seed of 999 to ensure that this code generates the same assignment each time

*since we have an unequal number of settings in some regions, we need to assign a misfit strategy to deal with the settings left out of the equal group assignment - here we assign the strata method which effectively randomly assigns it to one group in this context

randtreat, generate(treatment) setseed(990) strata(pviormaintained regionnum) multiple(2) misfits(strata)


**# Check the balance of the randomisation


*now we check that we have balance in the two strata by tabulating their counts in both conditions

tab treatment pviormaintained

tab treatment regionnum


**# Merge setting information back into treatment allocation


*we remerge to bring the setting information back into the dataset

merge 1:1 uniqueobs using [FILEPATH REDACTED]"


**# Save the final dataset


*finally we save the resulting dataset which contains all of the required information

save "[FILEPATH REDACTED]"
```

Appendix H: Analysis Syntax

```
####################  Concept Cat Analysis: Primary Analysis #################


## [PURPOSE OF SCRIPT]:
```

```
####+  This script conducts the main primary analysis for concept cat.
####+  We are exploring the impact of the intervention on 1 primary outcome:
####+    1. CELF P2 'Basic Concepts' subtest

####+    For this outcome, 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


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


# Clear workspace
rm(list=ls())
set.seed(991) #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("dplyr")
library("lmerTest")
library("officer")
library("boot")
library("boot.pval")
```

```
# Load cleaned data
data <- read.csv("[FILEPATH REDACTED]Concept_Cat_Analysis_Data.csv")



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

# 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
basic_concept_ovr_hist <- hist(data$basicconceptendline,
     main="Histogram of CELF-P2 Basic Concepts subtest at endline (overall sample)",
     xlab = "Basic Concepts scores",
     xlim = c(0, 18),
     breaks = seq(0, 18, 1),
     xaxp = c(0, 18, 9))

#+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)
basic_concept_model_hist <- hist(dplyr::filter(data,
                   !is.na(basicconcepttotal) &
                     !is.na(PVI) &
                     !is.na(region) &
                     !is.na(Setting_id))$basicconceptendline,
     main="Histogram of CELF-P2 Basic Concepts subtest at endline (analytical sample)",
     xlab = "Basic Concepts scores",
     xlim = c(0, 18),
```

```
   breaks = seq(0, 18, 1),
   xaxp = c(0, 18, 9))


##+ We can see we have a pretty severe ceiling effect in the primary outcome.
##+ It's so pronounced there is no need to run any further sensitivity checks
##+ for this- we will need to run a Tobit model (see sensitivity analysis code
##+ script)


#2. Means, SDs, Min and Max (for whole sample and analytical sample)

### Overall sample
Hmisc::describe(dplyr::filter(data,
                              !is.na(basicconceptendline))$basicconceptendline)
sd(dplyr::filter(data,
                 !is.na(basicconceptendline))$basicconceptendline)


# Overall sample by treatment and control groups

Hmisc::describe(dplyr::filter(data, #treatment
                              !is.na(basicconceptendline) &
                                treatment == 1)$basicconceptendline)

sd(dplyr::filter(data, #treatment
                 !is.na(basicconceptendline) &
                   treatment == 1)$basicconceptendline)

Hmisc::describe(dplyr::filter(data, #control
                              !is.na(basicconceptendline) &
                                treatment == 0)$basicconceptendline)

sd(dplyr::filter(data, #control
                 !is.na(basicconceptendline) &
                   treatment == 0)$basicconceptendline)
```

```
### Analytical sample

Hmisc::describe(dplyr::filter(data,

                              !is.na(basicconcepttotal) &

                                !is.na(PVI) &

                                !is.na(region) &

                                !is.na(basicconceptendline) &

                                !is.na(Setting_id))$basicconceptendline)

sd(dplyr::filter(data,

                 !is.na(basicconcepttotal) &

                   !is.na(PVI) &

                   !is.na(region) &

                   !is.na(basicconceptendline) &

                   !is.na(Setting_id))$basicconceptendline)



# Analytical sample by treatment and control groups:

Hmisc::describe(dplyr::filter(data, #treatment

                              treatment == 1 &

                              !is.na(basicconcepttotal) &

                                !is.na(PVI) &

                                !is.na(region) &

                                !is.na(basicconceptendline) &

                                !is.na(Setting_id))$basicconceptendline)


sd(dplyr::filter(data, #treatment

                 treatment == 1 &

                 !is.na(basicconcepttotal) &

                   !is.na(PVI) &

                   !is.na(region) &

                   !is.na(basicconceptendline) &

                   !is.na(Setting_id))$basicconceptendline)


Hmisc::describe(dplyr::filter(data, #control

                              treatment == 0 &
```

```
                    !is.na(basicconcepttotal) &
                      !is.na(PVI) &
                      !is.na(region) &
                      !is.na(Setting_id))$basicconceptendline)
sd(dplyr::filter(data, #control
                treatment == 0 &
                !is.na(basicconcepttotal) &
                  !is.na(PVI) &
                  !is.na(region) &
                  !is.na(basicconceptendline) &
                  !is.na(Setting_id))$basicconceptendline)

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



# Section B: Primary outcome model


###+ 1. Run the multi-level model. This is a random intercept model, with random effects at setting level.
#+ Basic Concept scores at endline as the outcome
#+ Covariates: treatment assignment, Basic Concept scores at baseline, region, PVI.
#+ Random effects: Setting


pr_model <- lmer(basicconceptendline ~
                  treatment +
                  basicconcepttotal +
                  region +
                  PVI +
                  (1 | Setting_id),
                data=data, REML = FALSE)


summary(pr_model) #Produce the model coefficients (fixed effects) with SEs and p-values


performance::icc(pr_model) #ICC
```

```r
#########


### 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<0.5 rejects H0 that data is normally distributed)
ks.test(resid_pr, "pnorm") #K-S test (p<0.5 rejects H0 that data is normally distributed



# 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 = "[FILEPATH REDACTED]primary_outcome.docx")


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

```
### 3.  Bootstrapping CIs and p-value

#+ The QQ residual plot, kernal density plot of residuals, 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(991) # Set seed- this is set to 991, the same seed set for randomisation

# Bootstrap residuals
boot_pr_model <- boot.pval::boot_summary(pr_model,
                                         type = "norm",
                                         method = NULL,
                                         conf.level = 0.95)
boot_pr_model # View bootsrapped SEs and p-values


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


### 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(basicconceptendline) &
                          !is.na(basicconcepttotal) &
                          !is.na(region) &
                          !is.na(PVI) &
                          !is.na(Setting_id))) #Number of individuals in treatment group in the model

m <- nrow(dplyr::filter(data,
                        treatment == 0 &
                          !is.na(basicconceptendline) &
                          !is.na(basicconcepttotal) &
                          !is.na(region) &
                          !is.na(PVI) &
                          !is.na(Setting_id))) #Number of individuals in control group in the model

v <- var(dplyr::filter(data,
                       treatment == 1 &
                         !is.na(basicconcepttotal) &
                         !is.na(region) &
                         !is.na(PVI) &
                         !is.na(basicconceptendline) &
                         !is.na(Setting_id))$basicconceptendline) #Variance in outcome among treatment group

w <- var(dplyr::filter(data,
                       treatment == 0 &
                         !is.na(basicconcepttotal) &
                         !is.na(region) &
                         !is.na(PVI) &
                         !is.na(basicconceptendline) &
```

```
                    !is.na(Setting_id))$basicconceptendline)  #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 bootstrapped 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 empty table in the report


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


# Treatment
n_mis <- nrow(dplyr::filter(data,
                    treatment == 1 &
                      !is.na(basicconceptendline) &
                      is.na(basicconcepttotal) &
                      is.na(region) &
                      is.na(PVI) &
                      is.na(Setting_id)))


# Control
m_mis <- nrow(dplyr::filter(data,
                        treatment == 0 &
                          !is.na(basicconceptendline) &
                          is.na(basicconcepttotal) &
                          is.na(region) &
                          is.na(PVI) &
                          is.na(Setting_id)))


# 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(basicconceptendline) &
                            !is.na(basicconcepttotal) &
                            !is.na(region) &
                            !is.na(PVI) &
                            !is.na(Setting_id))$basicconceptendline) #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(basicconceptendline) &
                            !is.na(basicconcepttotal) &
                            !is.na(region) &
                            !is.na(PVI) &
                            !is.na(Setting_id))$basicconceptendline)


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 = 2)),
                " (", 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) <- "CELF-P2 Basic Concepts"

## Display results for copying

View(primary_analysis_table)

###########

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

# Unadjusted difference in means

u_m <- as.character(round((mean(dplyr::filter(data,
                              treatment == 1 &
                                !is.na(basicconcepttotal) &
                                !is.na(region) &
                                !is.na(PVI) &
                                !is.na(Setting_id))$basicconceptendline, na.rm = TRUE)) - (mean(dplyr::filter(data,
                                                                              treatment == 0 &
                                                                                !is.na(basicconcepttotal) &
                                                                                !is.na(region) &
                                                                                !is.na(PVI) &
                                                                                !is.na(Setting_id))$basicconceptendline, 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

pr_es_est <- t(as.data.frame(c(u_m, c, num_t, var_t, num_c, var_c, pooled_sd)))

colnames(pr_es_est) <- c("Unadj. diff in means", "Adj diff in means", "N (treatment)", "Variance (treatment)", "N (control)",
"Variance (control)", "Pooled SD")

rownames(pr_es_est) <- "CELF P2 Basic Concepts"

View(pr_es_est) # View table


##################### Concept Cat: Secondary Analysis Code ##################


####+ This script conducts the main secondary analysis for concept cat.

####+ We are exploring the impact of the intervention on 2 secondary outcomes:

####+   1. CELF P2 'Concepts and following directions' subtest

####+   2. EYT Early Numeracy Assessment

#+

#+

#+  For each of these two outcomes, 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



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


#Section A: Loading the data

rm(list = ls())

set.seed(991) #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("dplyr")
library("lmerTest")
library("officer")
library("boot")
library("boot.pval")
library("plyr")
library("crctStepdown")



data <- read.csv("[FILEPATH REDACTED]Concept_Cat_Analysis_Data.csv")
###########################


######### Secondary Outcome 1: Following directions:

#Section A: Descriptive statistics

#1. Histograms of outcome

#Use this to check distribution of endline scores for our CELF P2 basic concepts and following directions subtest.
#We are mainly checking for floor effects, as flagged at baseline stage.

# In overall sample
fd_hist_ovr_sample <- hist(data$followdirectionendline,
   main="Histogram of CELF-P2 Following Directions subtest at endline",
```

```
    xlab = "Following Directions scores",

    xlim = c(0, 21),

    breaks = seq(0, 22, 1), # Set bins to match number of possible scores for outcome

    xaxp = c(0, 20, 10))


#In analytical sample

fd_hist_ana_sample <- hist(dplyr::filter(data,

                !is.na(followdirectiontotal) &

                  !is.na(region) &

                  !is.na(PVI) &

                  !is.na(Setting_id))$followdirectionendline,

    main="Histogram of CELF-P2 Following Directions subtest at endline",

    xlab = "Following Directions scores",

    xlim = c(0, 21),

    breaks = seq(0, 22, 1), # Set bins to match number of possible scores for outcome

    xaxp = c(0, 20, 10))


#2. Means, SDs, Min and Max (for whole sample and analytical sample). This will just go into the text in the report.



#Overall sample

Hmisc::describe(dplyr::filter(data,

                        !is.na(followdirectionendline))$followdirectionendline)

sd(dplyr::filter(data,

              !is.na(followdirectionendline))$followdirectionendline)


##Analytical sample

Hmisc::describe(dplyr::filter(data,

                        !is.na(followdirectiontotal) &

                          !is.na(region) &

                          !is.na(PVI) &

                          !is.na(followdirectionendline) &

                          !is.na(Setting_id))$followdirectionendline)

sd(dplyr::filter(data,

              !is.na(followdirectiontotal) &

                !is.na(region) &
```

```
          !is.na(PVI) &
          !is.na(followdirectionendline) &
          !is.na(Setting_id))$followdirectionendline)


#And by treatment group
Hmisc::describe(dplyr::filter(data,
                treatment == 1 &
                !is.na(followdirectiontotal) &
                 !is.na(region) &
                 !is.na(PVI) &
                 !is.na(followdirectionendline) &
                 !is.na(Setting_id))$followdirectionendline) #Treatment
sd(dplyr::filter(data,
         treatment == 1 &
         !is.na(followdirectiontotal) &
          !is.na(region) &
          !is.na(PVI) &
          !is.na(followdirectionendline) &
          !is.na(Setting_id))$followdirectionendline)


#And by control group
Hmisc::describe(dplyr::filter(data,
                treatment == 0 &
                !is.na(followdirectiontotal) &
                 !is.na(region) &
                 !is.na(PVI) &
                 !is.na(followdirectionendline) &
                 !is.na(Setting_id))$followdirectionendline) #Control
sd(dplyr::filter(data,
         treatment == 0 &
         !is.na(followdirectiontotal) &
          !is.na(region) &
          !is.na(PVI) &
          !is.na(followdirectionendline) &
          !is.na(Setting_id))$followdirectionendline)
```

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



### Section B: Secondary outcome model


#+ Run the multi-level model. Concepts and Following directions endline scores as outcome

#+ Controlling for baseline scores, region, and PVI status, and random intercept using setting_id

fd_model <- lmer(followdirectionendline ~

                 treatment +

                 followdirectiontotal +

                 region +

                 PVI +

                 (1 | Setting_id),

               data=data, REML = FALSE)


summary(fd_model) # Produce model coefficients, SEs and p-values


performance::icc(fd_model) #ICC




#########


###Testing OLS assumptions


#Residual diagnostics: Testing normality of residuals OLS assumption

resid_fd <- resid(fd_model) #Create object which stores the residuals of the model

plot_resid_fd <- plot(density(resid_fd)) #Kernel density plot to explore normality

shapiro.test(resid_fd) #Shapiro-wilk test for normality. !FLAG potentially K-S test better with larger sample sizes! (p<0.5 rejects H0 that data is normally distributed)

ks.test(resid_fd, "pnorm") #K-S test (p<0.5 rejects H0 that data is normally distributed)



# QQ line

qq_line_fd <- ggplot(data = data.frame(resid = resid_fd), 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
fd_resid_df <- as.data.frame(resid_fd)
sd_resid <- sd(resid_fd)
max_resid <- max(resid_fd)
min_resid <- min(resid_fd)


resid_fd_kd <- ggplot(fd_resid_df, aes(x = resid_fd)) +
  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 = "Following Directions - Residuals Density Plot",
    x = "Residuals",
    y = "Density"
  )




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


#Testing independent predictors
#dwtest(fd_model)


### Figures to export
resid_fd_kd
qq_line_fd
```

```
plots <- list(resid_fd_kd, qq_line_fd)

### 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 = "[FILEPATH REDACTED]follow_directions.docx")



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

### Bootstrapping CIs and p-value

#+ The QQ residual plot, kernal density plot of residuals, 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(991) # Setting seed to match seed set at randomisation

# Bootstrapping residuals
```

```
boot_fd_model <- boot.pval::boot_summary(fd_model,
                        type = "norm",
                        method = NULL,
                        conf.level = 0.95)



boot_fd_model # View model with bootsrapped CIs and p-values




###+ 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, in the same way we did for the primary analysis:


coefs <- data.frame(summary(fd_model)$coefficients) #Save all coefficients from model
c <- coefs["treatment", "Estimate"] #Extract treatment coefficient
n <- nrow(dplyr::filter(data,
                    treatment == 1 &
                      !is.na(followdirectionendline) &
                      !is.na(followdirectiontotal) &
                      !is.na(region) &
                      !is.na(PVI) &
                      !is.na(Setting_id))) #Number of individuals in treatment in analytical sample
m <- nrow(dplyr::filter(data,
                    treatment == 0 &
                      !is.na(followdirectionendline) &
                      !is.na(followdirectiontotal) &
                      !is.na(region) &
                      !is.na(PVI) &
```

```
                  !is.na(Setting_id))) #Number of individuals in control in analytical sample


v <- var(dplyr::filter(data,
                  treatment == 1 &
                    !is.na(followdirectiontotal) &
                    !is.na(region) &
                    !is.na(PVI) &
                    !is.na(followdirectionendline) &
                    !is.na(Setting_id))$followdirectionendline) #Variance in outcome among treatment group in the analytical sample
w <- var(dplyr::filter(data,
                  treatment == 0 &
                    !is.na(followdirectiontotal) &
                    !is.na(region) &
                    !is.na(PVI) &
                    !is.na(followdirectionendline) &
                    !is.na(Setting_id))$followdirectionendline) #Variance in outcome among control group in the analytical sample



fd_effect_size <- hedges.g(c, n, m, v, w)
fd_effect_size #This is the main effect size for the following directions subtest.



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


# We redefine c to equal the lower bound bootstrapped 95% CI
c <- boot_fd_model$Lower.bound[2]


# Calculate lower CI for effect size:
fd_effect_size_low <- hedges.g(c, n, m, v, w)
fd_effect_size_low #This is the lower confidence interval for the treatment effect.



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


# We redefine c to equal the upper bound bootstrapped 95% CI
c <- boot_fd_model$Upper.bound[2]
```

```
# Calculate upper CI for effect size:
fd_effect_size_high <- hedges.g(c, n, m, v, w)
fd_effect_size_high #This is the upper confidence interval for the treatment effect.

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




#### Creating FD 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(followdirectionendline) &
                     is.na(followdirectiontotal) &
                     is.na(region) &
                     is.na(PVI) &
                     is.na(Setting_id)))

# Control
m_mis <- nrow(dplyr::filter(data,
                    treatment == 0 &
                     !is.na(followdirectionendline) &
                     is.na(followdirectiontotal) &
                     is.na(region) &
                     is.na(PVI) &
```

```
                              is.na(Setting_id)))


# 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(followdirectionendline) &

                               !is.na(followdirectiontotal) &

                               !is.na(region) &

                               !is.na(PVI) &

                               !is.na(Setting_id))$followdirectionendline) #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(followdirectionendline) &

                               !is.na(followdirectiontotal) &

                               !is.na(region) &

                               !is.na(PVI) &

                               !is.na(Setting_id))$followdirectionendline)
```

```
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(fd_effect_size, digits = 2)),

            " (", as.character(round(fd_effect_size_low, digits = 2)),

            " - ",

            as.character(round(fd_effect_size_high, digits = 2)),

            ")")


## Create table. We will combine this with the ENA table at the end of this script, as is to be presented in the report.


fd_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(fd_analysis_table) <- c("N (intervention)", "Mean (intervention)", "N (control)", "Mean (control)", "Total (T;C)", "Hedges g (95% CIs)", "p-value")


rownames(fd_analysis_table) <- "CELF-P2 Concepts and Following Directions"


View(fd_analysis_table)

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

###########


### 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(followdirectiontotal) &

                                !is.na(region) &

                                !is.na(PVI) &

                                !is.na(Setting_id))$followdirectionendline, na.rm = TRUE)) - (mean(dplyr::filter(data,

                                                                  treatment == 0 &

                                                                    !is.na(followdirectiontotal) &

                                                                    !is.na(region) &
```

```
                                                                     !is.na(PVI) &

                                                                     !is.na(Setting_id))$followdirectionendline, 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

fd_es_est <- t(as.data.frame(c(u_m, c, num_t, var_t, num_c, var_c, pooled_sd)))

colnames(fd_es_est) <- c("Unadj. diff in means", "Adj diff in means", "N (treatment)", "Variance (treatment)", "N (control)", "Variance (control", "Pooled SD")

rownames(fd_es_est) <- "CELF P2 Concepts and Following Directions"



######### Secondary Analysis: EYT ENA


#Section A: Descriptive statistics




#1. Histograms of outcomes


#Use this to check distribution of endline scores for EYT ENA scores.


# In overall data

ENA_ovr_hist <- hist(data$ena_endline,

    main="Histogram of ENA Early Numeracy Task scores at endline",

    xlab = "early Numeracy Task scores",

    xlim = c(0, 88),

    breaks = seq(0, 88, 1), # Adjusting bins to match number of possible values for endline scores

    xaxp = c(0, 88, 8))


#In analytical data
```

```
ENA_ovr_hist <- hist(dplyr::filter(data,
                 !is.na(basicconcepttotal) &
                   !is.na(region) &
                   !is.na(PVI) &
                   !is.na(Setting_id))$ena_endline,
     main="Histogram of ENA Early Numeracy Task scores at endline",
     xlab = "Early Numeracy Task scores",
     xlim = c(0, 88),
     breaks = seq(0, 88, 1), # Adjusting bins to match number of possible values for endline scores
     xaxp = c(0, 88, 8))


#2. Means, SDs, Min and Max (for whole sample and analytical sample)


#Overall sample
Hmisc::describe(dplyr::filter(data,
                            !is.na(ena_endline))$ena_endline)
sd(dplyr::filter(data,
               !is.na(ena_endline))$ena_endline)


##Analytical sample
Hmisc::describe(dplyr::filter(data,
                              !is.na(basicconcepttotal) &
                              !is.na(region) &
                              !is.na(PVI) &
                              !is.na(ena_endline) &
                              !is.na(Setting_id))$ena_endline)
sd(dplyr::filter(data,
                 !is.na(basicconcepttotal) &
                 !is.na(region) &
                 !is.na(PVI) &
                 !is.na(ena_endline) &
                 !is.na(Setting_id))$ena_endline)


#And by treatment and control groups:
Hmisc::describe(dplyr::filter(data, #Treatment
                            treatment == 1 &
```

```
                     !is.na(basicconcepttotal) &

                     !is.na(region) &

                     !is.na(PVI) &

                     !is.na(ena_endline) &

                     !is.na(Setting_id))$ena_endline)

sd(dplyr::filter(data,

            treatment == 1 &

              !is.na(basicconcepttotal) &

              !is.na(region) &

              !is.na(PVI) &

              !is.na(ena_endline) &

              !is.na(Setting_id))$ena_endline)




Hmisc::describe(dplyr::filter(data, #Control

                    treatment == 0 &

                     !is.na(basicconcepttotal) &

                     !is.na(region) &

                     !is.na(PVI) &

                     !is.na(ena_endline) &

                     !is.na(Setting_id))$ena_endline)

sd(dplyr::filter(data,

            treatment == 0 &

              !is.na(basicconcepttotal) &

              !is.na(region) &

              !is.na(PVI) &

              !is.na(ena_endline) &

              !is.na(Setting_id))$ena_endline)



### Section B: EYT ENA Secondary outcome model



#+ Run the multi-level model. ENA EYT endline scores as outcome

#+ Controlling for baseline scores, region, and PVI status, and random intercept using setting_id
```

```
ena_model <- lmer(ena_endline ~
           treatment +
           basicconcepttotal +
          region +
           PVI +
           (1 | Setting_id),
          data=data, REML = FALSE)



summary(ena_model) # Produce model coefficients, SEs, and p-values


performance::icc(ena_model) #ICC


#########


###Testing OLS assumptions


#Residual diagnostics: Testing normality of residuals OLS assumption

resid_ena <- resid(ena_model) #Create object which stores the residuals of the model

plot_resid_ena <- plot(density(resid_ena)) #Kernel density plot to explore normality


# QQ line

ena_qq_line <- ggplot(data = data.frame(resid = resid_ena), aes(sample = resid)) +

  stat_qq() +

  stat_qq_line() +

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

  theme_minimal()



shapiro.test(resid_ena) #Shapiro-wilk test for normality. !FLAG potentially K-S test better with larger sample sizes! (p<0.5 rejects H0 that data is normally distributed)

ks.test(resid_ena, "pnorm") #K-S test (p<0.5 rejects H0 that data is normally distributed)


#Plotting residual density against a normal distribution for comparison. NB this is using the residual objects created from the primary model above

ena_resid_df <- as.data.frame(resid_ena)
```

```
sd_resid <- sd(resid_ena)

max_resid <- max(resid_ena)

min_resid <- min(resid_ena)


ena_resid_kd <- ggplot(ena_resid_df, aes(x = resid_ena)) +

  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 = "EYT ENA - Residuals Density Plot",

    x = "Residuals",

    y = "Density"

  )


#Testing linearity and assumption that residual errors have a mean of 0

plot(ena_model, col = "red") #We want the line here to be horizontal and at 0


#Figures to be exported

ena_qq_line

ena_resid_kd


plots <- list(ena_resid_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 = "[FILEPATH REDACTED]ENA.docx")

### Bootstrapping CIs and p-value

#+ The QQ residual plot, kernal density plot of residuals, 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(991)
boot_ena_model <- boot.pval::boot_summary(ena_model,
                                          type = "norm",
                                          method = NULL,
                                          conf.level = 0.95)

boot_ena_model


### Effect-size estimation

# Using the hedges g function created earlier.

#EFFECT SIZE:
#We use the output from the model above to calculate the effect size:
coefs <- data.frame(summary(ena_model)$coefficients)
c <- coefs["treatment", "Estimate"]
n <- nrow(dplyr::filter(data,
                        treatment == 1 &
```

```
                        !is.na(ena_endline) &
                        !is.na(basicconcepttotal) &
                        !is.na(region) &
                        !is.na(PVI) &
                        !is.na(Setting_id))) #number of individuals in treatment in model
m <- nrow(dplyr::filter(data,
                      treatment == 0 &
                        !is.na(ena_endline) &
                        !is.na(basicconcepttotal) &
                        !is.na(region) &
                        !is.na(PVI) &
                        !is.na(Setting_id)))#number of individuals in control in model


v <- var(dplyr::filter(data,
                     treatment == 1 &
                       !is.na(ena_endline) &
                       !is.na(basicconcepttotal) &
                       !is.na(region) &
                       !is.na(PVI) &
                       !is.na(Setting_id))$ena_endline) #Variance in outcome among treatment group
w <- var(dplyr::filter(data,
                     treatment == 0 &
                       !is.na(ena_endline) &
                       !is.na(basicconcepttotal) &
                       !is.na(region) &
                       !is.na(PVI) &
                       !is.na(Setting_id))$ena_endline) #Variance in outcome among control group




ena_effect_size <- hedges.g(c, n, m, v, w)
ena_effect_size #This is the main effect size for the following directions subtest.


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


# We redefine c to equal the lower bound bootstrapped 95% CI
```

```
c <- boot_ena_model$Lower.bound[2]

ena_effect_size_low <- hedges.g(c, n, m, v, w)
ena_effect_size_low #This is the lower confidence interval for the treatment effect.

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

# We redefine c to equal the upper bound bootstrapped 95% CI
c <- boot_ena_model$Upper.bound[2]

ena_effect_size_high <- hedges.g(c, n, m, v, w)
ena_effect_size_high #This is the upper confidence interval for the treatment effect.

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

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

#### Creating EYT ENA 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(ena_endline) &
                     is.na(basicconcepttotal) &
                     is.na(region) &
                     is.na(PVI) &
                     is.na(Setting_id)))

# Control
```

```
m_mis <- nrow(dplyr::filter(data,
                         treatment == 0 &
                           !is.na(ena_endline) &
                           is.na(basicconcepttotal) &
                           is.na(region) &
                           is.na(PVI) &
                           is.na(Setting_id)))


# 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(ena_endline) &
                            !is.na(basicconcepttotal) &
                            !is.na(region) &
                            !is.na(PVI) &
                            !is.na(Setting_id))$ena_endline) #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(ena_endline) &
                          !is.na(basicconcepttotal) &
                          !is.na(region) &
                          !is.na(PVI) &
                          !is.na(Setting_id))$ena_endline)


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(ena_effect_size, digits = 2)),
         " (", as.character(round(ena_effect_size_low, digits = 2)),
         " - ",
         as.character(round(ena_effect_size_high, digits = 2)),
         ")")


## Create table
ena_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(ena_analysis_table) <- c("N (intervention)", "Mean (intervention", "N (control)", "Mean (control)", "Total (T;C)", "Hedges g (95% CIs)", "p-value")


rownames(ena_analysis_table) <- "EYT ENA"


##### Combine secondary analysis results into 1 table:
secondary_analysis_table <- rbind(fd_analysis_table, ena_analysis_table)
View(secondary_analysis_table) # View secondary analysis table




### Creating effect size estimation table for appendix. 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(basicconcepttotal) &

                          !is.na(region) &

                          !is.na(PVI) &

                          !is.na(Setting_id))$ena_endline, na.rm = TRUE)) - (mean(dplyr::filter(data,

                                                                     treatment == 0 &

                                                                      !is.na(basicconcepttotal) &

                                                                      !is.na(region) &

                                                                      !is.na(PVI) &

                                                                      !is.na(Setting_id))$ena_endline, 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

ena_es_est <- t(as.data.frame(c(u_m, c, num_t, var_t, num_c, var_c, pooled_sd)))

colnames(ena_es_est) <- c("Unadj. diff in means", "Adj diff in means", "N (treatment)", "Variance (treatment)", "N (control)",
"Variance (control)", "Pooled SD")

rownames(ena_es_est) <- "EYT ENA"


#Combine with following directions ES estimation table

secondary_es_estimation <- rbind(fd_es_est, ena_es_est)

View(secondary_es_estimation) # View effect size estimation table




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


###+ Romano-Wolff Correction for Multiple comparisons problem.

###+    We states in the SAP that we would address the potential for an increased

###+    FWER arising from using multiple secondary outcomes through a RW

###+    correction on the p-values and CIs of the secondary models
```

```
###+    This is problematic in this instance as it assumes that the models are
###+    run on the exact same dataset (this issue revolves around having multiple
###+    hypotheses from the same data). Given that we have different extents of
###+    missingness across these 2 outcomes, we can't truly run this correction.
###+    However, for completeness, we have run the analysis below, but only on
###+    a subset of the data where we have complete observations for all included
###+    variables in both models.

# Subset data into only complete cases for both secondary analysis models
data_sec <- dplyr::filter(data,
                !is.na(ena_endline),
                !is.na(basicconcepttotal),
                !is.na(followdirectionendline),
                !is.na(followdirectiontotal))



# Convert PVI variable to a binary
data_sec$PVI_num <- ifelse(data_sec$PVI == "PVI", 1, 0)



#+ Create dummies for regions. This is because LMER creates dummies when it runs
#+ a model using a categorical independent variable. The stepdown function breaks
#+ when it encounters this as the dummy variables don't exist in the original data.
#+ We don't create one for bright futures, as this is the reference category.
data_sec$region_HS <- ifelse(data_sec$region == "Heart South", 1, 0)
data_sec$region_HN <- ifelse(data_sec$region == "Heart North", 1, 0)
data_sec$region_EV <- ifelse(data_sec$region == "Everton", 1, 0)



# Run FD model on non-missing subset
fd_model_2 <- lme4::lmer(followdirectionendline ~
                treatment +
                followdirectiontotal +
                region_HS  +
                region_HN +
                region_EV +
```

```
                     PVI_num +
                     (1 | Setting_id),
                   data=data_sec, REML = FALSE)


# Run ENA model on non-missing subset
ena_model_2 <- lmer(ena_endline ~
                 treatment +
                 basicconcepttotal +
                 region_HS  +
                 region_HS +
                 region_HS +
                 PVI_num +
                 (1 | Setting_id),
               data=data_sec, REML = FALSE)



# Run RW correction command
results <- stepdown(
  fitlist = list(fd_model_2, ena_model_2),
  tr_var = "treatment",
  cl_var = "Setting_id",
  data = data_sec,
  alpha = 0.05,
  plots = FALSE,
  n_permute = 1000,
  nsteps = 1000,
  type = "rw",
  confint = TRUE,
  verbose = TRUE
)

###Print results
print(results)


####################  Concept Cat: Subgroup Analysis: EYPP ####################
```

Please note the following code is identical to code run for EAL and SEND sub-groups. Code for these groups has been omitted for brevity.

```
######+ The purpose of this script is to conduct the necessary EYPP subgroup
######+ analysis for the concept cat trial.
######+
######+ For each of the EYPP sub-group we need to do the following analysis
######+     1. Descriptive analysis (mean/sd/min/max)
######+     2. Histograms of outcome distributions for both overall
######+        and analytical samples
######+     3. Run a multi-level regression model on the sub-group sample
######+     4. Run a multi-level regression model with an interaction term
######+     5. (Test model assumptions)
######+     6. Calculate effect sizes for each model type.


# Boilerplate
rm(list=ls())
set.seed(991)

# 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("officer")
library("boot")
library("boot.pval")
```

```
# Load cleaned data

data <- read.csv("[FILEPATH REDACTED]Concept_Cat_Analysis_Data.csv") #<-Insert main cleaned dataset filepath here.


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



##### SUBGROUP 1: EYPP Subgroup Analysis


## A: Descriptive analysis of CELF-P2 among EYPP subgroup:


#Leaving descriptive stats for the moment.


# Create subsaet of data of only EYPP children

data_EYPP <- dplyr::filter(data, EYPP == "Yes")


# Generate histogram of variable distribution in overall EYPP sample

EYPP_ovr_hist <- hist(data_EYPP$basicconceptendline,

                      main="Histogram of CELF-P2 Basic Concepts subtest at endline (EYPP overall subgroup)",

                      xlab = "Basic Concepts scores",

                      xlim = c(0, 18),

                      breaks = seq(0, 18, 1),

                      xaxp = c(0, 18, 9))



# Generate histogram of variable distribution in analytical EYPP sample

EYPP_model_hist <- hist(dplyr::filter(data_EYPP,

                                       !is.na(basicconcepttotal) &

                                        !is.na(PVI) &

                                        !is.na(region) &

                                        !is.na(Setting_id))$basicconceptendline,

                        main="Histogram of CELF-P2 Basic Concepts subtest at endline (EYPP analytical subgroup)",

                        xlab = "Basic Concepts scores",

                        xlim = c(0, 18),

                        breaks = seq(0, 18, 1),
```

```
                         xaxp = c(0, 18, 9))


###########


## B: Run primary analysis model within EYPP subgroup


# Run multi-level model
EYPP_sg_model <- lmer(basicconceptendline ~
          treatment +
          basicconcepttotal +
         region +
          PVI +
          (1 | Setting_id),
         data=data_EYPP, REML = FALSE)


summary(EYPP_sg_model) #Produce the results


performance::icc(EYPP_sg_model) #ICC


## Testing OLS assumptions Check distribution of residuals
#Residual diagnostics: Testing normality of residuals OLS assumption
resid_EYPP_sg <- resid(EYPP_sg_model) #Create object which stores the residuals of the model
plot_resid_EYPP_sg <- plot(density(resid_EYPP_sg)) #Kernel density plot to explore normality
shapiro.test(resid_EYPP_sg) #Shapiro-wilk test for normality. !FLAG potentially K-S test better with larger sample sizes! (p<0.5 rejects H0 that data is normally distributed)
ks.test(resid_EYPP_sg, "pnorm") #K-S test (p<0.5 rejects H0 that data is normally distributed)


# QQ line
qq_line_sg <- ggplot(data = data.frame(resid = resid_EYPP_sg), 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
resid_EYPP_sg_df <- as.data.frame(resid_EYPP_sg)
```

```
sd_resid <- sd(resid_EYPP_sg)

max_resid <- max(resid_EYPP_sg)

min_resid <- min(resid_EYPP_sg)


resid_EYPP_sg_kd  <- ggplot(resid_EYPP_sg_df, aes(x = resid_EYPP_sg)) +

  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 = "EYPP Subgroup Model - Residuals Density Plot",

    x = "Residuals",

    y = "Density"

  )



### Bootstrapping CIs and p-value


#+ The QQ residual plot, kernal density plot of residuals, 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(991)

boot_EYPP_sg_model <- boot.pval::boot_summary(EYPP_sg_model,

                                    type = "norm",

                                    method = NULL,

                                    conf.level = 0.95)

boot_EYPP_sg_model


# EFFECT SIZE:


#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(EYPP_sg_model)$coefficients) #Save all coefficients from model


c <- coefs["treatment", "Estimate"] #extract treatment coefficient
n <- nrow(dplyr::filter(data,
                       treatment == 1 &
                         EYPP == "Yes" &
                         !is.na(basicconceptendline) &
                         !is.na(basicconcepttotal) &
                         !is.na(region) &
                         !is.na(PVI) &
                         !is.na(Setting_id))) #number of individuals in treatment in model


m <- nrow(dplyr::filter(data,
                       treatment == 1 &
                         EYPP == "Yes" &
                         !is.na(basicconceptendline) &
                         !is.na(basicconcepttotal) &
                         !is.na(region) &
                         !is.na(PVI) &
                         !is.na(Setting_id))) #number of individuals in control in model
v <- var(dplyr::filter(data,
                      treatment == 1 &
                        EYPP == "Yes" &
                        !is.na(basicconcepttotal) &
                        !is.na(region) &
                        !is.na(PVI) &
                        !is.na(basicconceptendline) &
                        !is.na(Setting_id))$basicconceptendline) #Variance in outcome among treatment group
w <- var(dplyr::filter(data,
                      treatment == 0 &
                        EYPP == "Yes" &
                        !is.na(basicconcepttotal) &
                        !is.na(region) &
```

```
        !is.na(PVI) &
        !is.na(basicconceptendline) &
        !is.na(Setting_id))$basicconceptendline) #Variance in outcome among control group


EYPP_sg_effect_size <- hedges.g(c, n, m, v, w)
EYPP_sg_effect_size #This is the main effect size on primary outcome for EYPP subgroup.

#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_EYPP_sg_model$Lower.bound[2]

EYPP_sg_effect_size_low <- hedges.g(c, n, m, v, w)
EYPP_sg_effect_size_low #This is the lower confidence interval for the treatment effect.



#UPPER CI EFFECT SIZE using SEs from the above model:

c <- boot_EYPP_sg_model$Upper.bound[2]

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

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


#### Creating EYPP sub-group analysis table (to paste output into report)

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

# Treatment
n_mis <- nrow(dplyr::filter(data,
```

```
                         treatment == 1 &
                           EYPP == "Yes" &
                           !is.na(basicconceptendline) &
                           is.na(basicconcepttotal) &
                           is.na(region) &
                           is.na(PVI) &
                           is.na(Setting_id)))

# Control
m_mis <- nrow(dplyr::filter(data,
                         treatment == 0 &
                           EYPP == "Yes" &
                           !is.na(basicconceptendline) &
                           is.na(basicconcepttotal) &
                           is.na(region) &
                           is.na(PVI) &
                           is.na(Setting_id)))

# 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 &
                            EYPP == "Yes" &
                            !is.na(basicconceptendline) &
                            !is.na(basicconcepttotal) &
                            !is.na(region) &
```

```
                    !is.na(PVI) &

                    !is.na(Setting_id))$basicconceptendline) #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 &

                    EYPP == "Yes" &

                    !is.na(basicconceptendline) &

                    !is.na(basicconcepttotal) &

                    !is.na(region) &

                    !is.na(PVI) &

                    !is.na(Setting_id))$basicconceptendline)


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(EYPP_sg_effect_size, digits = 2)),

          " (", as.character(round(EYPP_sg_effect_size_low, digits = 2)),

          " - ",

          as.character(round(EYPP_sg_effect_size_high, digits = 2)),

          ")")


## Create table

EYPP_sg_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(EYPP_sg_analysis_table) <- c("N (intervention)", "Mean (intervention", "N (control)", "Mean (control)", "Total (T;C)",
"Hedges g (95% CIs)", "p-value")


rownames(EYPP_sg_analysis_table) <- "CELF-P2 Basic Concepts"


## Display results

View(EYPP_sg_analysis_table)


#########


### 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 &

                              EYPP == "Yes" &

                              !is.na(basicconcepttotal) &

                              !is.na(region) &

                              !is.na(PVI) &

                              !is.na(Setting_id))$basicconceptendline, na.rm = TRUE)) - (mean(dplyr::filter(data,

                                                                              treatment == 0 &

                                                                                EYPP == "Yes" &

                                                                                !is.na(basicconcepttotal) &

                                                                                !is.na(region) &

                                                                                !is.na(PVI) &

                                                                                !is.na(Setting_id))$basicconceptendline, 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
EYPP_sg_es_est <- t(as.data.frame(c(u_m, c, num_t, var_t, num_c, var_c, pooled_sd)))
colnames(EYPP_sg_es_est) <- c("Unadj. diff in means", "Adj diff in means", "N (treatment)", "Variance (treatment)", "N (control)", "Variance (control", "Pooled SD")
rownames(EYPP_sg_es_est) <- "CELF P2 Basic Concepts"


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


#B: Primary analysis model with FSM interaction (also need to include individual terms for treatment and EYPP)
EYPP_int_model <- lmer(basicconceptendline ~
           treatment +
            EYPP +
            treatment*EYPP +
           basicconcepttotal +
          region +
           PVI +
           (1 | Setting_id),
          data=data, REML = FALSE)


summary(EYPP_int_model)


## Testing OLS assumptions Check distribution of residuals
#Residual diagnostics: Testing normality of residuals OLS assumption
resid_EYPP_int <- resid(EYPP_int_model) #Create object which stores the residuals of the model
plot_resid_EYPP_int <- plot(density(resid_EYPP_int)) #Kernel density plot to explore normality


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


shapiro.test(resid_EYPP_int) #Shapiro-wilk test for normality. !FLAG potentially K-S test better with larger sample sizes! (p<0.5 rejects H0 that data is normally distributed)
ks.test(resid_EYPP_int, "pnorm") #K-S test (p<0.5 rejects H0 that data is normally distributed)
```

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

resid_EYPP_int_kd <- ggplot(resid_EYPP_int_df, aes(x = resid_EYPP_int)) +
  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 = "EYPP Interaction Model - Residuals Density Plot",
    x = "Residuals",
    y = "Density"
  )


# Figures to output
qq_line_sg
resid_EYPP_sg_kd
qq_line_int
resid_EYPP_int_kd


plots <- list(resid_EYPP_sg_kd, qq_line_sg, resid_EYPP_int_kd, qq_line_int)

### 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 = "[FILEPATH REDACTED]EYPP.docx")

### Bootstrapping CIs and p-value (int model)

#+ The QQ residual plot, kernal density plot of residuals, 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(991)
boot_EYPP_int_model <- boot.pval::boot_summary(EYPP_int_model,
                                                type = "norm",
                                                method = NULL,
                                                conf.level = 0.95)
boot_EYPP_int_model

#We use the output from the model above to calculate the effect size: (NB: This needs to be done for confidence intervals too!)
c <- boot_EYPP_int_model["treatment:EYPPYes", "Estimate"]
n <- nrow(dplyr::filter(data,
                       treatment == 1 &
                         !is.na(basicconceptendline) &
                         !is.na(EYPP) &
                         !is.na(basicconcepttotal) &
```

```
                  !is.na(region) &
                  !is.na(PVI) &
                  !is.na(Setting_id))) #number of individuals in treatment in model


m <- nrow(dplyr::filter(data,
                 treatment == 0 &
                  !is.na(basicconceptendline) &
                  !is.na(EYPP) &
                  !is.na(basicconcepttotal) &
                  !is.na(region) &
                  !is.na(PVI) &
                  !is.na(Setting_id))) #number of individuals in control in model


v <- var(dplyr::filter(data,
                treatment == 1 &
                 !is.na(EYPP) &
                 !is.na(basicconcepttotal) &
                 !is.na(region) &
                 !is.na(PVI) &
                 !is.na(basicconceptendline) &
                 !is.na(Setting_id))$basicconceptendline) #Variance in outcome among treatment group


w <- var(dplyr::filter(data,
                treatment == 0 &
                 !is.na(EYPP) &
                 !is.na(basicconcepttotal) &
                 !is.na(region) &
                 !is.na(PVI) &
                 !is.na(basicconceptendline) &
                 !is.na(Setting_id))$basicconceptendline) #Variance in outcome among control group



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

EYPP_int_effect_size #This is the main effect size for the interaction effect between EYPP status and treatment assignment on the primary outcome.
```

```
#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_EYPP_int_model["treatment:EYPPYes", "Lower.bound"]


EYPP_int_effect_size_low <- hedges.g(c, n, m, v, w) #Running hedges G function.

EYPP_int_effect_size_low #This is the lower confidence interval for the treatment effect.



#UPPER CI EFFECT SIZE using SEs from the above model:


c <- boot_EYPP_int_model["treatment:EYPPYes", "Upper.bound"]


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

EYPP_int_effect_size_high


## P-value

p_values <- boot_EYPP_int_model["treatment:EYPPYes", "p.value"]

p_value <- round(p_values, digits = 2)


#### Creating EYPP interaction analysis table (to paste output into report)


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


# Treatment

n_mis <- nrow(dplyr::filter(data,

                           treatment == 1 &

                             !is.na(basicconceptendline) &

                             is.na(basicconcepttotal) &

                             is.na(region) &

                             is.na(PVI) &

                             is.na(EYPP) &

                             is.na(Setting_id)))


# Control

m_mis <- nrow(dplyr::filter(data,
```

```
                    treatment == 0 &
                     !is.na(basicconceptendline) &
                     is.na(basicconcepttotal) &
                     is.na(region) &
                     is.na(PVI) &
                     is.na(EYPP) &
                     is.na(Setting_id)))


# 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(EYPP) &
                      !is.na(basicconceptendline) &
                      !is.na(basicconcepttotal) &
                      !is.na(region) &
                      !is.na(PVI) &
                      !is.na(Setting_id))$basicconceptendline) #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(EYPP) &
                 !is.na(basicconceptendline) &
                 !is.na(basicconcepttotal) &
                 !is.na(region) &
                 !is.na(PVI) &
                 !is.na(Setting_id))$basicconceptendline)

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(EYPP_int_effect_size, digits = 2)),
        " (", as.character(round(EYPP_int_effect_size_low, digits = 2)),
        " - ",
        as.character(round(EYPP_int_effect_size_high, digits = 2)),
        ")")

#Create table
EYPP_int_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(EYPP_int_analysis_table) <- c("N (intervention)", "Mean (intervention", "N (control)", "Mean (control)", "Total (T;C)",
"Hedges g (95% CIs)", "p-value")


rownames(EYPP_int_analysis_table) <- "CELF-P2 Basic Concepts"


View(EYPP_int_analysis_table)


###########


### 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(EYPP) &
                         !is.na(basicconcepttotal) &
                         !is.na(region) &
                         !is.na(PVI) &
                         !is.na(Setting_id))$basicconceptendline, na.rm = TRUE)) - (mean(dplyr::filter(data,
                                                                  treatment == 0 &
                                                                   !is.na(EYPP) &
                                                                   !is.na(basicconcepttotal) &
                                                                   !is.na(region) &
                                                                   !is.na(PVI) &
                                                                   !is.na(Setting_id))$basicconceptendline, 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
EYPP_int_es_est <- t(as.data.frame(c(u_m, c, num_t, var_t, num_c, var_c, pooled_sd)))
colnames(EYPP_int_es_est) <- c("Unadj. diff in means", "Adj diff in means", "N (treatment)", "Variance (treatment)", "N (control)",
"Variance (control", "Pooled SD")
rownames(EYPP_int_es_est) <- "CELF P2 Basic Concepts"
View(EYPP_int_es_est)


############## Concept Cat Primary Analysis Tobit Model ##############


#+ Purpose of this script:
###+ 1. Exploration of the magnitude of the ceiling effect in 'Basic Concepts
###+    scores (as laid out in the SAP)
###+ 2. To account for censoring at the upper level of the primary outcome
###+    using Tobit regression, re-estimating the effect-size
###+
```

```
# Clear workspace
rm(list=ls())

# Set seed
set.seed(991)

# Load packages
library("dplyr")
library("plm")
library("censReg")


# load data
data <- read.csv("[FILEPATH REDACTED]Concept_Cat_Analysis_Data.csv")

######


#### 1. Inspection of the magnitude of ceiling effect in primary outcome (in analytical sample)


## A. Visual inspection (histogram)
basic_concept_model_hist <- hist(dplyr::filter(data,
                                !is.na(basicconcepttotal) &
                                 !is.na(PVI) &
                                 !is.na(region) &
                                 !is.na(Setting_id))$basicconceptendline,
                     main="Histogram of CELF-P2 Basic Concepts subtest at endline (analytical sample)",
                     xlab = "Basic Concepts scores",
                     xlim = c(0, 18),
                     breaks = seq(0, 18, 1),
                     xaxp = c(0, 18, 9)) # As we've already done this- we know the ceiling effect is pretty huge!

####
```

```
## B. Calculation of Pearson's coefficient of skewness


#+ This is calculated by finding the difference between the mean and the median,
#+ multiplying by 3, then dividing by the SD. See function written below:


pearson.skew <- function(d, v){
  m <- mean(d[, v], na.rm = T)
  med <- median(d[, v], na.rm = T)
  sdv <- sd(d[, v], na.rm = T)
  m_med <- 3*(m - med)
  p <- m_med/sdv
  return(p)
}


skew_coeff <- pearson.skew(data, "basicconceptendline")
skew_coeff


# Skew coefficient of -0.779


####


## C. Exploration of proportion of sample within 1 SD of the maximum value


# Calculate SD
sd <- sd(data$basicconceptendline, na.rm = T)


# Create df which is a copy of analytical data
data_sd <- dplyr::filter(data,
                          !is.na(basicconceptendline) &
                            !is.na(PVI) &
                            !is.na(region) &
                            !is.na(Setting_id))


#+ Create a variable which takes on value of 1 if value of outcome is less than 1
#+ SD from 18 (maximum value)
data_sd <- data_sd %>%
```

```
  mutate(sd_dummy = 0)


data_sd$sd_dummy <- ifelse((18 - data_sd$basicconceptendline) < sd, 1, 0)


mean(data_sd$sd_dummy) #Find mean of dummy to find proportion of obs with


# 58.2% of the sample is within 1 SD of the maximum value- this is a serious ceiling effect...




###########




#### 2. Create multi-level censored regression model (tobit model)


# Establish hierarchical structure of data (Setting_id as setting identifier)

h.data <- pdata.frame(data,

                    index = "Setting_id")


h.data$Setting_id <- as.numeric(h.data$Setting_id) #Convert to numeric


# Specify tobit model

#+ Note: This command in R can be replicated in STATA using the metobit command as follows:

#+ metobit basicconcepttotal treatment followdirectiontotal || Setting_id:, ll(0) ul(18) intpoints(8) intmethod(ghermite)

#+ We're using 8 integration points, with Newton-Raphson MLE, and gaussian-hermite integration method.


pr_tobit <- censReg(formula = basicconceptendline ~ treatment + basicconcepttotal + region + PVI, #No need to specify the nested structure of the dataset, as this has been done already as part of the p.data command.

                data = h.data,

                method = "Newton-Raphson",

                left = 0, #Lower limit of var

                right = 18) #Upper limit of var



summary(pr_tobit) #View results
```

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


## Calculating effect size


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



#Extract data frame of results
coefs <- data.frame(summary(pr_tobit)$estimate)


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


# P-value
p_values <- coefs[, "Pr...t."]
p_value <- round(p_values[2], digits = 2) #Round to 2dp


# Obtain other relevant information to generate effect size
n <- nrow(dplyr::filter(data,
                      treatment == 1 &
                        !is.na(basicconceptendline) &
                        !is.na(basicconcepttotal) &
                        !is.na(region) &
                        !is.na(PVI) &
                        !is.na(Setting_id))) #Number of individuals in treatment group in the model
m <- nrow(dplyr::filter(data,
                      treatment == 0 &
                        !is.na(basicconceptendline) &
                        !is.na(basicconcepttotal) &
                        !is.na(region) &
                        !is.na(PVI) &
                        !is.na(Setting_id))) #Number of individuals in control group in the model
```

```
v <- var(dplyr::filter(data,

                      treatment == 1 &

                        !is.na(basicconcepttotal) &

                        !is.na(region) &

                        !is.na(PVI) &

                        !is.na(basicconceptendline) &

                        !is.na(Setting_id))$basicconceptendline) #Variance in outcome among treatment group

w <- var(dplyr::filter(data,

                      treatment == 0 &

                        !is.na(basicconcepttotal) &

                        !is.na(region) &

                        !is.na(PVI) &

                        !is.na(basicconceptendline) &

                        !is.na(Setting_id))$basicconceptendline)  #Variance in outcome among control group



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

pr_tb_effect_size



#LOWER CI EFFECT SIZE using SEs from the above model:

se <- sqrt(diag(vcov(pr_tobit))) #save the standard errors for all

se <- data.frame(se)

se <- se["treatment", ] #Extract the treatment SE


#We redefine c to equal the treatment coefficient minus the SE created above

c <- c - se


pr_tb_effect_size_low <- hedges.g(c, n, m, v, w) #Re-run hedges G on lower CI treatment coefficient

pr_tb_effect_size_low #This is the lower confidence interval for the treatment effect on the following directions subtest.



#UPPER CI EFFECT SIZE using SEs 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 <- c + (se*2)

pr_tb_effect_size_high <- hedges.g(c, n, m, v, w) #Re-run hedges G on upper CI treatment coefficient
pr_tb_effect_size_high #This is the upper confidence interval



#### Creating primary analysis tobit model 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(basicconceptendline) &
                          is.na(basicconcepttotal) &
                          is.na(region) &
                          is.na(PVI) &
                          is.na(Setting_id)))

# Control
m_mis <- nrow(dplyr::filter(data,
                        treatment == 0 &
                          !is.na(basicconceptendline) &
                          is.na(basicconcepttotal) &
                          is.na(region) &
                          is.na(PVI) &
                          is.na(Setting_id)))

# 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(basicconceptendline) &

                     !is.na(basicconcepttotal) &

                     !is.na(region) &

                     !is.na(PVI) &

                     !is.na(Setting_id))$basicconceptendline) #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(basicconceptendline) &

                     !is.na(basicconcepttotal) &

                     !is.na(region) &

                     !is.na(PVI) &

                     !is.na(Setting_id))$basicconceptendline)


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(pr_tb_effect_size, digits = 2)),

        " (", as.character(round(pr_tb_effect_size_low, digits = 2)),

        " - ",

        as.character(round(pr_tb_effect_size_high, digits = 2)),

        ")")


## Create table


tobit_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(tobit_table) <- c("N (intervention)", "Mean (intervention", "N (control)", "Mean (control)", "Total (T;C)", "Hedges g (95% CIs)", "p-value")


rownames(tobit_table) <- "CELF-P2 Basic Concepts (tobit regression)"


## Display results

View(tobit_table)


#########




### Tobit model effect size estimation


### 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(basicconcepttotal) &

                       !is.na(region) &

                       !is.na(PVI) &

                       !is.na(Setting_id))$basicconceptendline, na.rm = TRUE)) - (mean(dplyr::filter(data,

                                                  treatment == 0 &

                                                   !is.na(basicconcepttotal) &

                                                   !is.na(region) &

                                                   !is.na(PVI) &

                                                   !is.na(Setting_id))$basicconceptendline, 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

tobit_es_est <- t(as.data.frame(c(u_m, c, num_t, var_t, num_c, var_c, pooled_sd)))

colnames(tobit_es_est) <- c("Unadj. diff in means", "Adj diff in means", "N (treatment)", "Variance (treatment)", "N (control)", "Variance (control", "Pooled SD")

rownames(tobit_es_est) <- "CELF P2 Basic Concepts"


View(tobit_es_est)




################  Concept Cat: CACE Analysis ###################


###+ Purpose of script is to explore the degree to which fidelity to the intervention

###+ affects its impact. This will involve:

######+ 1. Generating the binary compliance indicator

######+ 2. Descriptive overview of compliance

######+ 3. 2SLS/IV approach


# Boilerplate

rm(list = ls())

set.seed(991)


# Load necessary libraries

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")


# Load main data
data <- read.csv("[FILEPATH REDACTED]Concept_Cat_Analysis_Data.csv") #<-

### Overview of words learned compliance (setting-level compliance)
############

### Create school-level compliance

# Create variable
data <- data %>%
  dplyr::mutate(school_compliance = "Yes")

# Recode all non-compliant schools
data$school_compliance <- ifelse(data$Setting_name == "Daisychain Day Nursery", NA, data$school_compliance)
data$school_compliance <- ifelse(data$Setting_name == "Nursery In The Park", "No", data$school_compliance)
data$school_compliance <- ifelse(data$Setting_name == "Osborne Nursery School", "No", data$school_compliance)
data$school_compliance <- ifelse(data$Setting_name == "Percy Shurmer Academy", "No", data$school_compliance)
data$school_compliance <- ifelse(data$Setting_name == "Alexandra Park Primary School", "No", data$school_compliance)
data$school_compliance <- ifelse(data$Setting_name == "Barlow Hall Primary", "No", data$school_compliance)
data$school_compliance <- ifelse(data$Setting_name == "Bloomsbury Nursery School", "No", data$school_compliance)
data$school_compliance <- ifelse(data$Setting_name == "East Didsbury", "No", data$school_compliance)
data$school_compliance <- ifelse(data$Setting_name == "Kirkway Nursery", "No", data$school_compliance)
data$school_compliance <- ifelse(data$Setting_name == "Pleasant Street", "No", data$school_compliance)
data$school_compliance <- ifelse(data$Setting_name == "Rose Hill Primary School Nursery", "No", data$school_compliance)
data$school_compliance <- ifelse(data$Setting_name == "Rosegrove Nursery School", "No", data$school_compliance)
```

```
data$school_compliance <- ifelse(data$Setting_name == "Stoneyholme Nursery School", "No", data$school_compliance)
data$school_compliance <- ifelse(data$Setting_name == "Villiers Primary", "No", data$school_compliance)

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

###+ Create combined compliance metric which accounts for both school and pupil
###+ level compliance

# Create variable
data <- data %>%
  dplyr::mutate(compliance_num = 0)

# Recode for both minimum school and pupil compliance levels
data$compliance_num <- ifelse(data$school_compliance == "Yes" & data$min_hrs_attended == "Yes", 1, data$compliance_num)

# Recode to missing for missing obs for either of the 2 measures
data$compliance_num <- ifelse(is.na(data$school_compliance), NA, data$compliance_num)
data$compliance_num <- ifelse(is.na(data$min_hrs_attended), NA, data$compliance_num)

# Recode to 0 if control setting
data$compliance_num <- ifelse(data$treatment == 0, 0, data$compliance_num)

###########

### Descriptive tables of compliance within treatment group

# Overall
table(dplyr::filter(data,
                    treatment == 1)$compliance_num)

# Pupil
table(dplyr::filter(data,
                    treatment == 1)$min_hrs_attended)

# School
table(dplyr::filter(data,
```

```
          treatment == 1)$school_compliance)

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

##### 2SLS Analysis (CACE):

# Subset data into only data with no missing compliance data
data_compl <- dplyr::filter(data,
               !is.na(compliance_num))

###+1.  First stage: Run a multi-level model with binary compliance as the
###+    outcome, but with the same covariates as the main primary model
first_stage <- lmer(compliance_num ~
          treatment +
          basicconcepttotal +
          region +
          PVI +
          (1 | Setting_id),
         data=data_compl, REML = FALSE)

summary(first_stage)



# Predicted compliance- ignoring random effects
data_compl$predicted_compliance <- predict(first_stage, data_compl, re.form = ~(1|Setting_id),
                    allow.new.levels = T)

###+2. Second stage: Model the primary outcome with predicted compliance in
###+   place of treatment assignment (treatment instrumented by compliance)
second_stage <- lmer(basicconceptendline ~
            predicted_compliance +
            basicconcepttotal +
            region +
            PVI +
            (1 | Setting_id),
           data=data_compl, REML = FALSE)
```

```
summary(second_stage)


#########


##### Inspecting distribution of scores among compliant vs non-compliant group (it's a very large effect size!)


# Overall treatment group mean/variance
overall_t_mean <- mean(dplyr::filter(data,
                                     treatment == 1)$basicconceptendline, na.rm = T)


overall_t_var <- var(dplyr::filter(data,
                                   treatment == 1)$basicconceptendline, na.rm = T)


# Overall compliant group mean
compliant_mean <- mean(dplyr::filter(data_compl,
                  compliance_num==1)$basicconceptendline, na.rm = T)


compliant_var <- var(dplyr::filter(data_compl,
                                   compliance_num==1)$basicconceptendline, na.rm = T)


# Overall non-compliant treatment group mean
non_compliant_t_mean <- mean(dplyr::filter(data_compl,
                                   compliance_num==0 &
                                     treatment == 1)$basicconceptendline, na.rm = T)



# Mean/variance in treatment group with missing compliance data
missing_compliance_t_mean <- mean(dplyr::filter(data,
                                            is.na(compliance_num) &
                                              treatment == 1)$basicconceptendline, na.rm = T)


missing_compliance_t_var <- var(dplyr::filter(data,
                                             is.na(compliance_num) &
                                               treatment == 1)$basicconceptendline, na.rm = T)
```

```
##+ We can see from the above that there is a fairly sizeable difference both
##+ in terms of the means and variances of the outcomes between the compliant
##+ group, and the treatment group missing compliance data. This is likely
##+ the reason why we're observing a significant increase in effect size for the
##+ CACE analysis, and therefore these results should be interpreted with caution.


###Testing OLS assumptions


#Residual diagnostics: Testing normality of residuals OLS assumption
resid_pr_comp <- resid(second_stage) #Create object which stores the residuals of the model
plot_resid_pr_comp <- plot(density(resid_pr_comp)) #Kernel density plot to explore normality
shapiro.test(resid_pr_comp) #Shapiro-wilk test for normality. !FLAG potentially K-S test better with larger sample sizes! (p<0.5 rejects H0 that data is normally distributed)
ks.test(resid_pr_comp, "pnorm") #K-S test (p<0.5 rejects H0 that data is normally distributed


# QQ line
qq_line <- ggplot(data = data.frame(resid = resid_pr_comp), 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_comp_resid_df <- as.data.frame(resid_pr_comp)
sd_resid <- sd(resid_pr_comp)
max_resid <- max(resid_pr_comp)
min_resid <- min(resid_pr_comp)


resid_pr_comp_kd <- ggplot(pr_comp_resid_df, aes(x = resid_pr_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 = "Basic Concepts - Residuals Density Plot",
    x = "Residuals",
    y = "Density"
  )



### Figures to export
resid_pr_comp_kd
qq_line

plots <- list(resid_pr_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 = "[FILEPATH REDACTED]CACE_second_stage.docx")


### Bootstrapping CIs and p-value


#+ The QQ residual plot, kernal density plot of residuals, 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(991)
boot_CACE_model <- boot.pval::boot_summary(second_stage,
                                           type = "norm",
                                           method = NULL,
                                           conf.level = 0.95)
boot_CACE_model


# Calculate Hedges' g for the CACE estimate
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(second_stage)$coefficients) # Create data frame of all coefficients from IV model
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(data_compl,
                        treatment== 1 &
                          !is.na(basicconceptendline) &
                          !is.na(basicconcepttotal) &
                          !is.na(region) &
                          !is.na(PVI) &
                          !is.na(Setting_id))) # Number of individuals in treatment group in the model
m <- nrow(dplyr::filter(data_compl,
                        treatment == 0 &
                          !is.na(basicconceptendline) &
```

```
                 !is.na(basicconcepttotal) &

                 !is.na(region) &

                 !is.na(PVI) &

                 !is.na(Setting_id))) # Number of individuals in control group in the model

v <- var(dplyr::filter(data_compl,

                treatment == 1 &

                 !is.na(basicconcepttotal) &

                 !is.na(region) &

                 !is.na(PVI) &

                 !is.na(basicconceptendline) &

                 !is.na(Setting_id))$basicconceptendline) # Variance in outcome among treatment group

w <- var(dplyr::filter(data_compl,

                treatment == 0 &

                 !is.na(basicconcepttotal) &

                 !is.na(region) &

                 !is.na(PVI) &

                 !is.na(basicconceptendline) &

                 !is.na(Setting_id))$basicconceptendline)  # 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:

c <- boot_CACE_model$Lower.bound[2]

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


# Upper 95% CI

c <- boot_CACE_model$Upper.bound[2]# Extract the compliance coefficient

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


# P-values

p_values <- boot_CACE_model$p.value[2]

p_value <- round(p_values, digits = 2)


#### Creating CACE analysis table (to paste output into report)
```

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


# Treatment
n_mis <- nrow(dplyr::filter(data_compl,
                            treatment == 1 &
                              !is.na(basicconceptendline) &
                              is.na(basicconcepttotal) &
                              is.na(region) &
                              is.na(PVI) &
                              is.na(Setting_id)))


# Control
m_mis <- nrow(dplyr::filter(data_compl,
                            treatment == 0 &
                              !is.na(basicconceptendline) &
                              is.na(basicconcepttotal) &
                              is.na(region) &
                              is.na(PVI) &
                              is.na(Setting_id)))


# 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_compl,
                             treatment == 1 &
                               !is.na(basicconceptendline) &
                               !is.na(basicconcepttotal) &
                               !is.na(region) &
```

```
                        !is.na(PVI) &
                        !is.na(Setting_id))$basicconceptendline) #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_compl,
                       treatment == 0 &
                        !is.na(basicconceptendline) &
                        !is.na(basicconcepttotal) &
                        !is.na(region) &
                        !is.na(PVI) &
                        !is.na(Setting_id))$basicconceptendline)


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 = 2)),
             " (", as.character(round(pr_effect_size_low, digits = 2)),
             " - ",
             as.character(round(pr_effect_size_high, digits = 2)),
             ")")


## Create table


primary_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(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) <- "CELF-P2 Basic Concepts (Compliance)"


## Display results

View(primary_analysis_table)


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


### Effect size estimation table

u_m <- as.character(round((mean(dplyr::filter(data_compl, #Unadjusted difference in means- within the sample with non-missing compliance data

                         treatment == 1 &

                           !is.na(basicconcepttotal) &

                           !is.na(region) &

                           !is.na(PVI) &

                           !is.na(Setting_id))$basicconceptendline, na.rm = TRUE)) - (mean(dplyr::filter(data_compl,

                                                                     treatment == 0 &

                                                                       !is.na(basicconcepttotal) &

                                                                       !is.na(region) &

                                                                       !is.na(PVI) &

                                                                       !is.na(Setting_id))$basicconceptendline, 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_sd <- round(sqrt((((n - 1)*v)+((m - 1)* w))/n + (m-2)), digits = 2) #pooled standard deviation




# Create table

pr_es_est <- t(as.data.frame(c(u_m, c, num_t, var_t, num_c, var_c, pooled_sd)))

colnames(pr_es_est) <- c("Unadj. diff in means", "Adj diff in means", "N (treatment)", "Variance (treatment)", "N (control)", "Variance (control)", "Pooled SD")
```

```
rownames(pr_es_est) <- "CELF P2 Basic Concepts (compliance)"
View(pr_es_est)



#### Checks
mean(dplyr::filter(data_compl,
                  compliance_num == 1)$basicconceptendline,
    na.rm = T)

var(dplyr::filter(data_compl,
                 compliance_num == 1)$basicconceptendline,
    na.rm = T)



mean(dplyr::filter(data,
                  treatment == 1)$basicconceptendline,
    na.rm = T)

var(dplyr::filter(data,
                 treatment == 1)$basicconceptendline,
    na.rm = T)

############## Concept Cat: Attrition and Missing Data ##################

### Purpose of this script:
###+ 1. Explore the degree of missingness in the primary outcome
###+ 2. Generate figures for the attrition flow diagram
###+ 3. Create a balance at baseline/analysis table
###+ 4. Run logit models to explore possible determinants of missingness
###+ 5. Re-run primary analysis model (check similarity of results to main model)
###+ 6. (If suggested that data is MNAR in section 5): Additional sensitivity analysis
###+ 7. (If covariates in second primary analysis model are MAR): Multiple imputation
###+ 8. MDES table (at analysis section)
```

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

# Boilerplate

rm(list=ls())

set.seed(991) #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("dplyr")

library("ggplot2")

library("lme4")

library("sjstats")

library("Hmisc")

library("performance")

library("lmtest")

library("eeptools")

library("dplyr")

library("lmerTest")

library("jtools")

# Load cleaned data

data <- read.csv("[FILEPATH REDACTED]Concept_Cat_Analysis_Data.csv")

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

### 1. Pupil attrition in primary outcome- table 7 in report

### Number of pupils as randomised

table(data$treatment)

### Number of pupils as analysed

table(dplyr::filter(data,

                    !is.na(basicconceptendline))$treatment)
```

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


### 2. Attrition flow diagram figures
#(Ignore 'recruitment' section of diagram for now)

#Create datasets for each level of flow diagram
data_allocation <- data #Data at allocation
data_follow <- dplyr::filter(data,
                             !is.na(basicconceptendline)) #Data at follow-up
data_analysis <- dplyr::filter(data,
                               !is.na(basicconcepttotal) &
                                 !is.na(PVI) &
                                 !is.na(region) &
                                 !is.na(basicconceptendline) &
                                 !is.na(Setting_id)) #Data at analysis

## A. Settings

# Allocation (we've already got this info)
allocation_settings_ovr <- 89 #Overall settings allocated
allocation_settings_t <- 45 #Treatment settings allocated
allocation_settings_c <- 44 #Control settings allocated

# Follow-up
follow_settings_ovr <- length(unique(data_follow$Setting_id)) #Number of schools at follow-up (overall)
follow_settings_t <- length(unique(dplyr::filter(data_follow,
                                                 treatment == 1)$Setting_id)) #Number of schools at follow-up (treatment)
follow_settings_c <- length(unique(dplyr::filter(data_follow,
                                                 treatment == 0)$Setting_id)) #Number of schools at follow-up (control)


# Analysis
```

```
analysis_settings_ovr <- length(unique(data_analysis$Setting_id))

analysis_settings_t <- length(unique(dplyr::filter(data_analysis,

                                        treatment == 1)$Setting_id)) #Number of schools at analysis (treatment)

analysis_settings_c <- length(unique(dplyr::filter(data_analysis,

                                        treatment == 0)$Setting_id)) #Number of schools at analysis (control)


## B. Children


# Allocation (we've already got this info)

allocation_children_ovr <- 1040

allocation_children_t <- 527

allocation_children_c <- 513


# Follow-up

follow_children_ovr <- nrow(data_follow) #Number of children at follow-up (overall)

follow_children_t <- nrow(dplyr::filter(data_follow,

                                 treatment == 1)) #Number of children at follow-up (treatment)

follow_children_c <- nrow(dplyr::filter(data_follow,

                                 treatment == 0)) #Number of children at follow-up (control)


# Analysis

analysis_children_ovr <- nrow(data_analysis) #Number of children at analysis (overall)

analysis_children_t <- nrow(dplyr::filter(data_analysis,

                                   treatment == 1))  #Number of children at analysis (control)

analysis_children_c <- nrow(dplyr::filter(data_analysis,

                                   treatment == 0)) #Number of children at analysis (treatment)




# Make into table

setting_numbers <- as.data.frame(t(c(allocation_settings_ovr,

                  allocation_settings_t,

                  allocation_settings_c,

                  follow_settings_ovr,

                  follow_settings_t,

                  follow_settings_c,

                  analysis_settings_ovr,
```

```
                    analysis_settings_t,
                    analysis_settings_c)))


child_numbers <- as.data.frame(t(c(allocation_children_ovr,
                   allocation_children_t,
                   allocation_children_c,
                   follow_children_ovr,
                   follow_children_t,
                   follow_children_c,
                   analysis_children_ovr,
                   analysis_children_t,
                   analysis_children_c)))


colnames <- c("at allocation (overall)",
              "at allocation (treatment)",
              "at allocation (control)",
              "at follow-up (overall)",
              "at follow-up (treatment)",
              "at follow-up (control)",
              "at analysis (overall)",
              "at analysis (treatment)",
              "at analysis (control)")


colnames(setting_numbers) <- colnames
colnames(child_numbers) <- colnames


trial_numbers_table <- dplyr::bind_rows(setting_numbers, child_numbers)
rownames(trial_numbers_table) <- c("Settings (N)", "Children (N)")


# View table
View(trial_numbers_table)




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

```
##### 3. Balance at baseline/analysis table


### Figures at baseline:

# For this, we need to obtain the relevant figures to populate the balance of baseline characteristics

# Convert EYPP, EAL, SEND, and gender to dummy variables
data <- data %>%
  mutate(EYPP_num = as.numeric(factor(EYPP, levels = c("No", "Yes")))-1)

data <- data %>%
  mutate(EAL_num = as.numeric(factor(EAL, levels = c("No", "Yes")))-1)

data <- data %>%
  mutate(SEND_num = as.numeric(factor(SEND, levels = c("No", "Yes")))-1)

data <- data %>%
  mutate(gender_num = as.numeric(factor(Gender, levels = c("Male", "Female")))-1) # Male = 0, Female = 1


####+ School-level characteristics

########+ 1. PVI

## Obtained from SAP

########  2. Region

## Obtained from SAP

########+ 3. EYPP (mean proportion)
```

```
# Create empty objects for baseline EYPP proportions

ovr_EYPP_baseline <- c()

treatment_EYPP_baseline <- c()

control_EYPP_baseline <- c()


## Create loops to calculate mean proportion of EYPP


# Overall

for(i in unique(data$Setting_id)){

  d <- dplyr::filter(data,

                     Setting_id == i)

  p <- mean(d$EYPP_num, na.rm = T)

  ovr_EYPP_baseline <- c(ovr_EYPP_baseline, p)

}


EYPP_ovr_baseline_prop <- mean(ovr_EYPP_baseline, na.rm = T)



# Treatment

for(i in unique(dplyr::filter(data,

                              treatment == 1)$Setting_id)){

  d <- dplyr::filter(data,

                     Setting_id == i)

  p <- mean(d$EYPP_num, na.rm = T)

  treatment_EYPP_baseline <- c(treatment_EYPP_baseline, p)

}


EYPP_tr_baseline_prop <- mean(treatment_EYPP_baseline, na.rm = T)



# Control

for(i in unique(dplyr::filter(data,

                              treatment == 0)$Setting_id)){

  d <- dplyr::filter(data,

                     Setting_id == i)

  p <- mean(d$EYPP_num, na.rm = T)
```

```
  control_EYPP_baseline <- c(control_EYPP_baseline, p)

}


EYPP_ct_baseline_prop <- mean(control_EYPP_baseline, na.rm = T)




# Display proportions

EYPP_ovr_baseline_prop # overall

EYPP_tr_baseline_prop # treatment

EYPP_ct_baseline_prop # control


########+ 4. EAL (mean proportion)


# Create empty objects for baseline EAL proportions

ovr_EAL_baseline <- c()

treatment_EAL_baseline <- c()

control_EAL_baseline <- c()


## Create loops to calculate mean proportion of EAL


# Overall

for(i in unique(data$Setting_id)){

  d <- dplyr::filter(data,

                     Setting_id == i)

  p <- mean(d$EAL_num, na.rm = T)

  ovr_EAL_baseline <- c(ovr_EAL_baseline, p)

}


EAL_ovr_baseline_prop <- mean(ovr_EAL_baseline, na.rm = T)




# Treatment

for(i in unique(dplyr::filter(data,

                              treatment == 1)$Setting_id)){

  d <- dplyr::filter(data,

                     Setting_id == i)
```

```
  p <- mean(d$EAL_num, na.rm = T)

  treatment_EAL_baseline <- c(treatment_EAL_baseline, p)

}


EAL_tr_baseline_prop <- mean(treatment_EAL_baseline, na.rm = T)



# Control

for(i in unique(dplyr::filter(data,

                              treatment == 0)$Setting_id)){

  d <- dplyr::filter(data,

                     Setting_id == i)

  p <- mean(d$EAL_num, na.rm = T)

  control_EAL_baseline <- c(control_EAL_baseline, p)

}


EAL_ct_baseline_prop <- mean(control_EAL_baseline, na.rm = T)



# Display proportions

EAL_ovr_baseline_prop # overall

EAL_tr_baseline_prop # treatment

EAL_ct_baseline_prop # control




########+ 5. SEND (mean proportion)



# Create empty objects for baseline SEND proportions

ovr_SEND_baseline <- c()

treatment_SEND_baseline <- c()

control_SEND_baseline <- c()


## Create loops to calculate mean proportion of SEND
```

```
# Overall
for(i in unique(data$Setting_id)){
  d <- dplyr::filter(data,
                     Setting_id == i)
  p <- mean(d$SEND_num, na.rm = T)
  ovr_SEND_baseline <- c(ovr_SEND_baseline, p)
}

SEND_ovr_baseline_prop <- mean(ovr_SEND_baseline, na.rm = T)


# Treatment
for(i in unique(dplyr::filter(data,
                              treatment == 1)$Setting_id)){
  d <- dplyr::filter(data,
                     Setting_id == i)
  p <- mean(d$SEND_num, na.rm = T)
  treatment_SEND_baseline <- c(treatment_SEND_baseline, p)
}

SEND_tr_baseline_prop <- mean(treatment_SEND_baseline, na.rm = T)


# Control
for(i in unique(dplyr::filter(data,
                              treatment == 0)$Setting_id)){
  d <- dplyr::filter(data,
                     Setting_id == i)
  p <- mean(d$SEND_num, na.rm = T)
  control_SEND_baseline <- c(control_SEND_baseline, p)
}

SEND_ct_baseline_prop <- mean(control_SEND_baseline, na.rm = T)


# Display proportions
```

```
SEND_ovr_baseline_prop # overall
SEND_tr_baseline_prop # treatment
SEND_ct_baseline_prop # control

### Gender


# Create empty objects for baseline gender proportions
ovr_gender_baseline <- c()
treatment_gender_baseline <- c()
control_gender_baseline <- c()

## Create loops to calculate mean proportion of gender

# Overall
for(i in unique(data$Setting_id)){
  d <- dplyr::filter(data,
                     Setting_id == i)
  p <- mean(d$gender_num, na.rm = T)
  ovr_gender_baseline <- c(ovr_gender_baseline, p)
}

gender_ovr_baseline_prop <- mean(ovr_gender_baseline, na.rm = T)


# Treatment
for(i in unique(dplyr::filter(data,
                              treatment == 1)$Setting_id)){
  d <- dplyr::filter(data,
                     Setting_id == i)
  p <- mean(d$gender_num, na.rm = T)
  treatment_gender_baseline <- c(treatment_gender_baseline, p)
}

gender_tr_baseline_prop <- mean(treatment_gender_baseline, na.rm = T)
```

```
# Control
for(i in unique(dplyr::filter(data,
                              treatment == 0)$Setting_id)){
  d <- dplyr::filter(data,
                     Setting_id == i)
  p <- mean(d$gender_num, na.rm = T)
  control_gender_baseline <- c(control_gender_baseline, p)
}

gender_ct_baseline_prop <- mean(control_gender_baseline, na.rm = T)


# Display proportions
gender_ovr_baseline_prop # overall
gender_tr_baseline_prop # treatment
gender_ct_baseline_prop # control



####+ Child-level:
########+ 1. Gender (percentage)

# Treatment
table(dplyr::filter(data,
                    treatment == 1)$Gender)

prop.table(table(dplyr::filter(data,
                               treatment == 1)$Gender))

# Control
table(dplyr::filter(data,
                    treatment == 0)$Gender)

prop.table(table(dplyr::filter(data,
                               treatment == 0)$Gender))
```

```
########+ 2. EYPP (percentage)

# treatment

table(dplyr::filter(data,

                    treatment == 1)$EYPP)


prop.table(table(dplyr::filter(data,

                               treatment == 1)$EYPP))


# Control

table(dplyr::filter(data,

                    treatment == 0)$EYPP)


prop.table(table(dplyr::filter(data,

                               treatment == 0)$EYPP))




########+ 3. EAL

# treatment

table(dplyr::filter(data,

                    treatment == 1)$EAL)


prop.table(table(dplyr::filter(data,

                               treatment == 1)$EAL))


# Control

table(dplyr::filter(data,

                    treatment == 0)$EAL)


prop.table(table(dplyr::filter(data,

                               treatment == 0)$EAL))

########+ 4. SEND (percentage)


# treatment

table(dplyr::filter(data,

                    treatment == 1)$SEND)
```

```
prop.table(table(dplyr::filter(data,

                               treatment == 1)$SEND))


# control

table(dplyr::filter(data,

                    treatment == 0)$SEND)


prop.table(table(dplyr::filter(data,

                               treatment == 0)$SEND))


########+ 5. Baseline scores (average)


# Available from SAP



########  6. Age in months


##Generating age in months variable. This is age at the end of summer term ()

data <- data%>%

  dplyr::mutate(DoB_num = ifelse(is.na(data$pupildob), 0, data$pupildob))


data <- data%>%

  dplyr::mutate(DoB_num = as.Date(data$pupildob))


data <- data %>%

  dplyr::mutate(age_months =  eeptools::age_calc(data$DoB_num, enddate = as.Date("2024-07-24"), units = "months", precise =
TRUE))


# Treatment

mean(dplyr::filter(data,

                   treatment == 1)$age_months, na.rm = T)

sum(is.na(dplyr::filter(data,

                        treatment == 1)$age_months))


# Control
```

```
mean(dplyr::filter(data,

                 treatment == 0)$age_months, na.rm = T)

sum(is.na(dplyr::filter(data,

                     treatment == 0)$age_months))




#########


### At analysis


# Create analysis data

data_analysis <- dplyr::filter(data,

                             !is.na(basicconcepttotal) &

                              !is.na(PVI) &

                              !is.na(region) &

                              !is.na(basicconceptendline) &

                              !is.na(Setting_id)) #Data at analysis


#+ No whole-school drop out so we are not exploring balance of school-level

#+ variables


####+ Child-level:

########+ 1. Gender (percentage)


# Treatment

table(dplyr::filter(data_analysis,

                  treatment == 1)$Gender)


prop.table(table(dplyr::filter(data_analysis,

                             treatment == 1)$Gender))


# Control

table(dplyr::filter(data_analysis,

                  treatment == 0)$Gender)
```

```
prop.table(table(dplyr::filter(data_analysis,
                          treatment == 0)$Gender))


########+ 2. EYPP (percentage)
# treatment
table(dplyr::filter(data_analysis,
                 treatment == 1)$EYPP)


prop.table(table(dplyr::filter(data_analysis,
                          treatment == 1)$EYPP))


# Control
table(dplyr::filter(data_analysis,
                 treatment == 0)$EYPP)


prop.table(table(dplyr::filter(data_analysis,
                          treatment == 0)$EYPP))



########+ 3. EAL
# treatment
table(dplyr::filter(data_analysis,
                 treatment == 1)$EAL)


prop.table(table(dplyr::filter(data_analysis,
                          treatment == 1)$EAL))


# Control
table(dplyr::filter(data_analysis,
                 treatment == 0)$EAL)


prop.table(table(dplyr::filter(data_analysis,
                          treatment == 0)$EAL))
########+ 4. SEND (percentage)


# treatment
```

```
table(dplyr::filter(data_analysis,

                    treatment == 1)$SEND)


prop.table(table(dplyr::filter(data_analysis,

                               treatment == 1)$SEND))


# control

table(dplyr::filter(data_analysis,

                    treatment == 0)$SEND)


prop.table(table(dplyr::filter(data_analysis,

                               treatment == 0)$SEND))



########  5. Age in months


# Treatment

mean(dplyr::filter(data_analysis,

                   treatment == 1)$age_months, na.rm = T)

sum(is.na(dplyr::filter(data_analysis,

                        treatment == 1)$age_months))


# Control

mean(dplyr::filter(data_analysis,

                   treatment == 0)$age_months, na.rm = T)

sum(is.na(dplyr::filter(data_analysis,

                        treatment == 0)$age_months))




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



############ 4. Exploring patterns in missingness through logit models


# We will also explore the degree to which baseline characteristics affect the probability that data is missing at endline
```

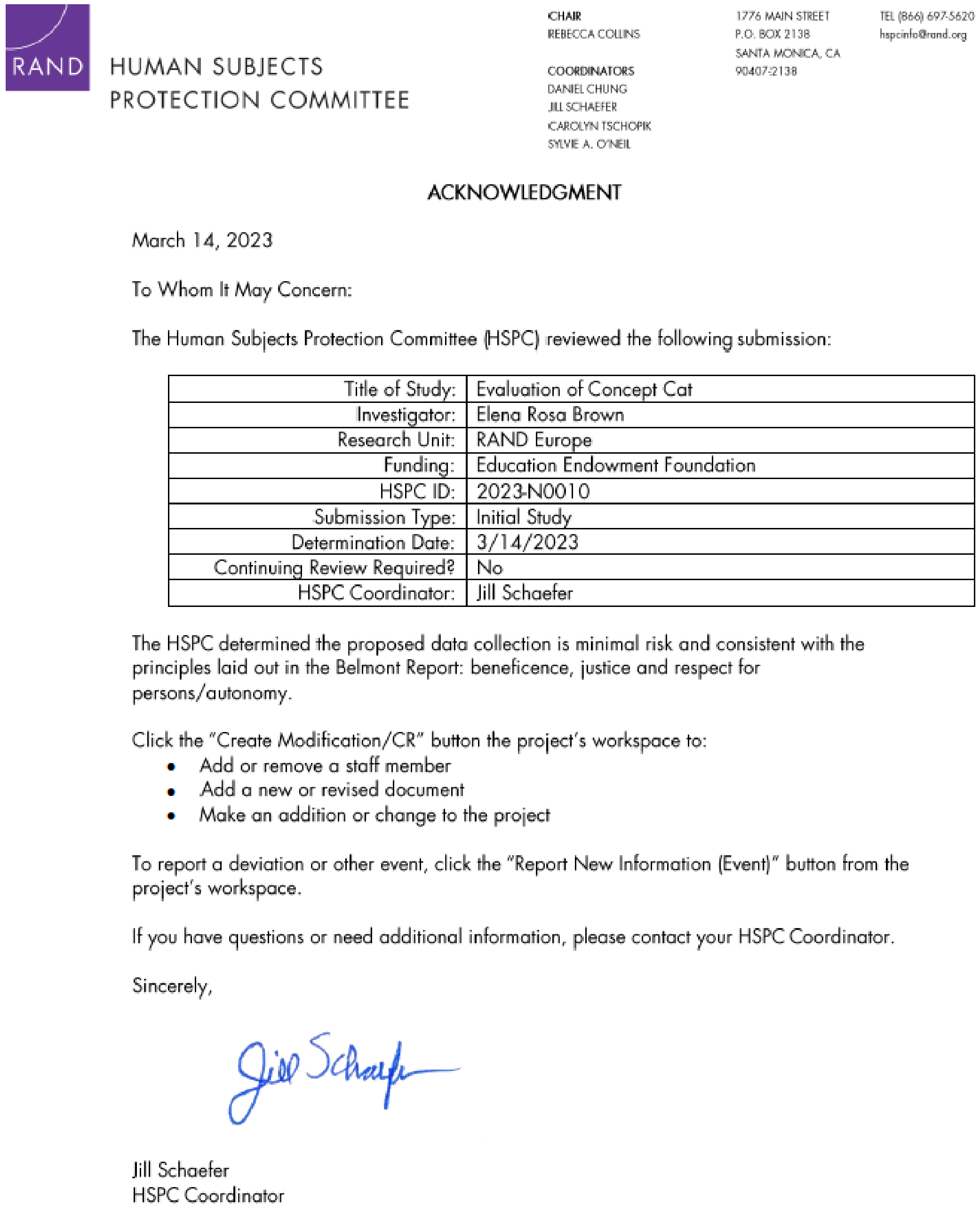
```
# Create a binary flag which is equal to 1 if endline data for the primary outcome is missing.
data <- data %>%
  dplyr::mutate(missing_flag = ifelse(is.na(basicconceptendline), 1, 0))


# Run a hierarchical logit model with the primary analysis model covariates included
logit1 <- glmer(missing_flag ~
          treatment +
          basicconcepttotal +
          region +
          PVI +
          (1 | Setting_id),
         data = data,
         family = binomial)


summary(logit1)
summ(logit1)


# Run a hierarchical logit model with gender, FSM, and SEND also included as covariates
logit2 <- glmer(missing_flag ~
          treatment +
          basicconcepttotal +
          region +
          PVI +
          EAL +
          EYPP +
          SEND +
          (1 | Setting_id),
         data = data,
         family = binomial)


summary(logit2)#
summ(logit2)



# Run a hierarchical logit model with gender, FSM, SEND, and Gender also included as covariates
```

```
logit3 <- glmer(missing_flag ~

         treatment +

         basicconcepttotal +

         region +

         PVI +

         EAL +

         EYPP +

         SEND +

         Gender +

         (1 | Setting_id),

        data = data,

        family = binomial)


summary(logit3)

summ(logit3)


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



### 5. Introduce covariates found to be associated with missingness into analysis model



#### The only covariate that has a statistically significant association with missingness at endline is PVI- this is already in the main analysis model. Therefore, no further analysis is required.




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



### 6. (If suggested that data is MNAR in section 5): Additional sensitivity analysis


##### Very unlikely data MNAR as the mechanism for missingness is largely known (pupils dropping out of settings)


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

7. (If covariates in second primary analysis model are MAR): Multiple imputation

Not required

Appendix I: RAND Europe ethical approval



RAND
HUMAN SUBJECTS
PROTECTION COMMITTEE

CHAIR
REBECCA COLLINS

COORDINATORS
DANIEL CHUNG
JILL SCHAEFER
CAROLYN TSCHOPIK
SYLVIE A. O'NEIL

1776 MAIN STREET
P.O. BOX 2138
SANTA MONICA, CA
90407-2138

TEL (866) 697-5620
hspcinfo@rand.org

ACKNOWLEDGMENT

March 14, 2023

To Whom It May Concern:

The Human Subjects Protection Committee (HSPC) reviewed the following submission:

Title of Study:	Evaluation of Concept Cat
Investigator:	Elena Rosa Brown
Research Unit:	RAND Europe
Funding:	Education Endowment Foundation
HSPC ID:	2023-N0010
Submission Type:	Initial Study
Determination Date:	3/14/2023
Continuing Review Required?	No
HSPC Coordinator:	Jill Schaefer

The HSPC determined the proposed data collection is minimal risk and consistent with the principles laid out in the Belmont Report: beneficence, justice and respect for persons/autonomy.

Click the "Create Modification/CR" button the project's workspace to:

- Add or remove a staff member
- Add a new or revised document
- Make an addition or change to the project

To report a deviation or other event, click the "Report New Information (Event)" button from the project's workspace.

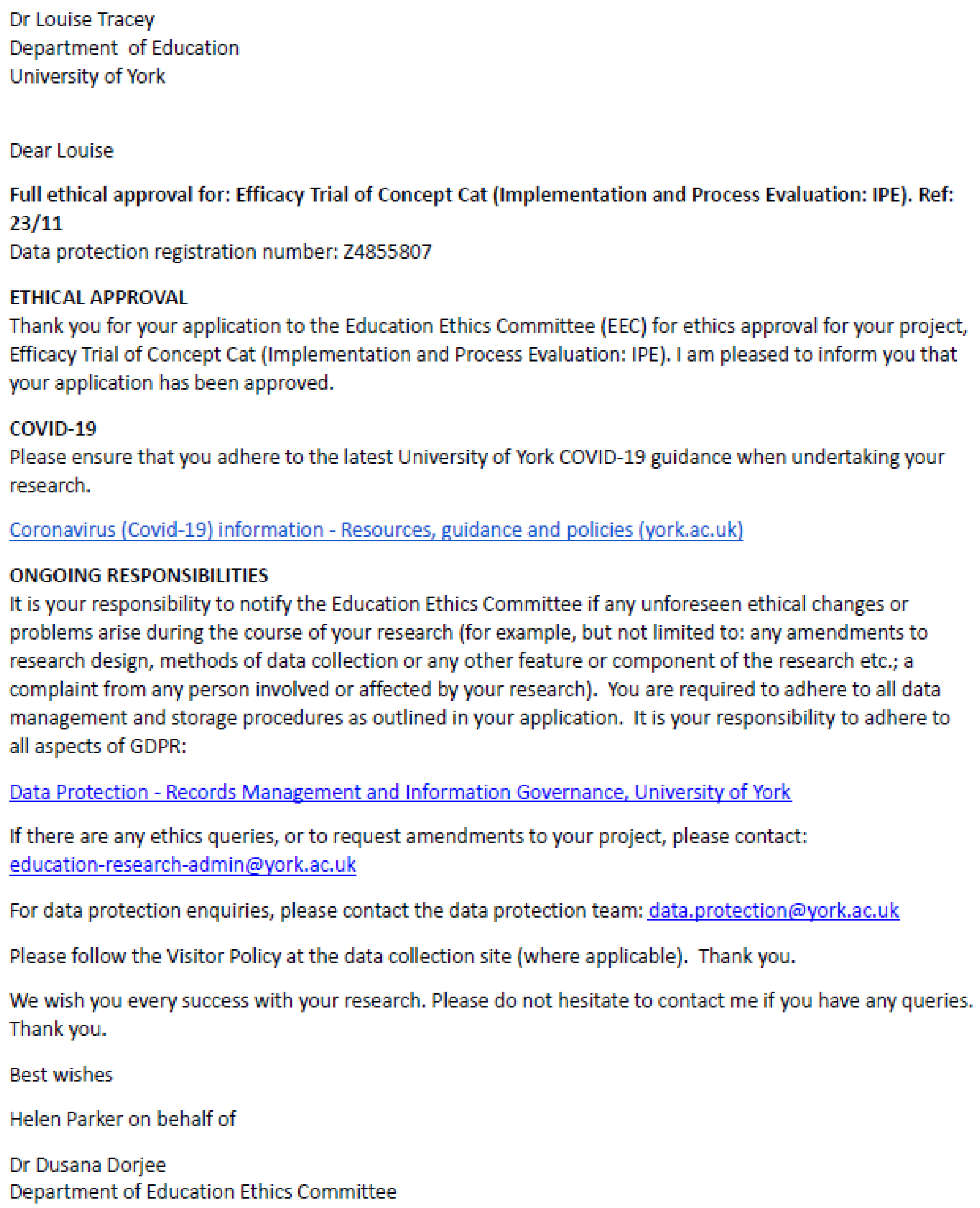
If you have questions or need additional information, please contact your HSPC Coordinator.

Sincerely,

Jill Schaefer

Jill Schaefer
HSPC Coordinator

Appendix J: University of York ethical approval



Dr Louise Tracey
Department of Education
University of York

Dear Louise

Full ethical approval for: Efficacy Trial of Concept Cat (Implementation and Process Evaluation: IPE). Ref: 23/11
Data protection registration number: Z4855807

ETHICAL APPROVAL
Thank you for your application to the Education Ethics Committee (EEC) for ethics approval for your project, Efficacy Trial of Concept Cat (Implementation and Process Evaluation: IPE). I am pleased to inform you that your application has been approved.

COVID-19
Please ensure that you adhere to the latest University of York COVID-19 guidance when undertaking your research.

Coronavirus (Covid-19) information - Resources, guidance and policies (york.ac.uk)

ONGOING RESPONSIBILITIES
It is your responsibility to notify the Education Ethics Committee if any unforeseen ethical changes or problems arise during the course of your research (for example, but not limited to: any amendments to research design, methods of data collection or any other feature or component of the research etc.; a complaint from any person involved or affected by your research). You are required to adhere to all data management and storage procedures as outlined in your application. It is your responsibility to adhere to all aspects of GDPR:

Data Protection - Records Management and Information Governance, University of York

If there are any ethics queries, or to request amendments to your project, please contact:
education-research-admin@york.ac.uk

For data protection enquiries, please contact the data protection team: data.protection@york.ac.uk

Please follow the Visitor Policy at the data collection site (where applicable). Thank you.

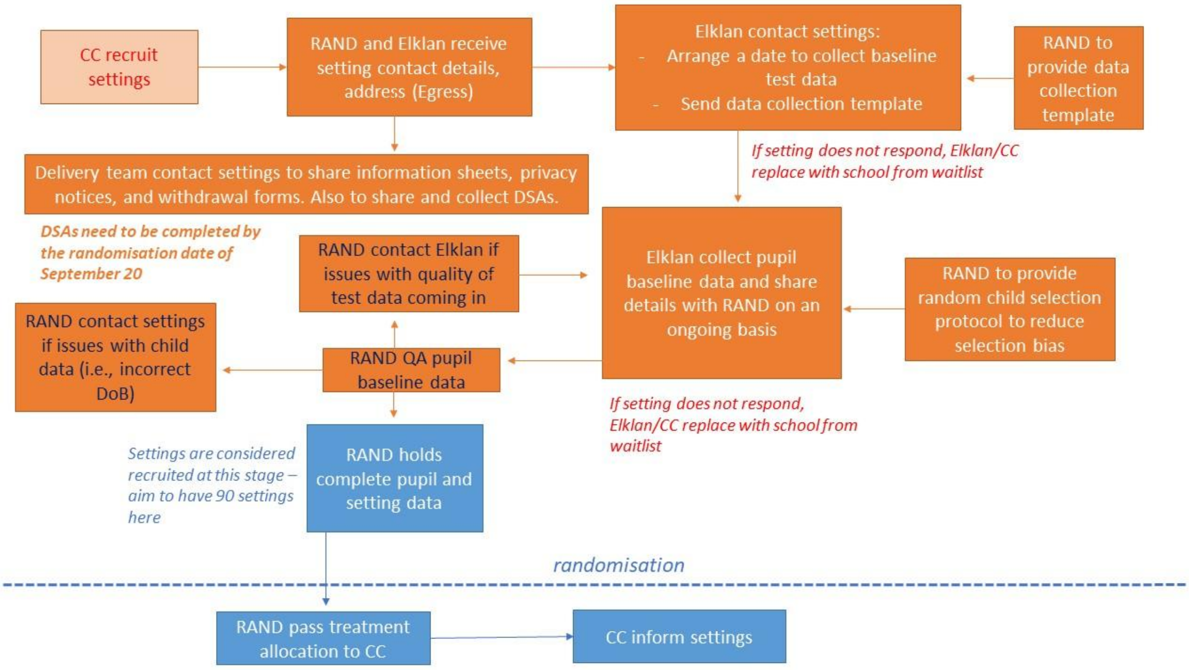
We wish you every success with your research. Please do not hesitate to contact me if you have any queries. Thank you.

Best wishes

Helen Parker on behalf of

Dr Dusana Dorjee
Department of Education Ethics Committee

Appendix K: Data flow diagram



Appendix L: Memorandum of Understanding

Concept Cat Randomised Controlled Trial

MEMORANDUM OF UNDERSTANDING

This study is designed to evaluate the impact of the Concept Cat programme on the language development and numeracy skills of children aged three to four years old. In addition, it seeks to identify teaching practices and strategies that promote early language development and numeracy.

Concept Cat is a whole-class teaching methodology for early verbal concepts (part of the widely used book Word Aware 2). This methodology is delivered by Stephen Parsons and Anna Branagan, authors of Word Aware 2, in collaboration with Better Communication CIC. The evaluation is funded by the Education Endowment Foundation (EEF) and supported by the Department for Education (DfE) as part of the Stronger Practice Hub initiative, which forms part of the latter's Early Years Educational Recovery Programme.[1] The results of this study will help us understand the impact of the Concept Cat programme on language development and numeracy outcomes in the early years.

This memorandum of understanding (MoU) explains what your setting's participation in the study will entail. If you agree to take part and accept the terms and conditions outlined, please sign a copy of this form and send to **ConceptCat@bettercommunication.org.uk** .

Who will be my main contacts during the study?

Throughout your participation in this study, you will be contacted by both the **delivery team** (Better Communication CIC), who will implement the Concept Cat programme and all related training, and by the independent **evaluation team** (RAND Europe and the University of York), who are carrying out an evaluation of its effectiveness. The evaluation team have partnered with Elklan, who have been sub-contracted to administer standardised tests to measure children's outcomes before and after delivery of Concept Cat.

If, after randomisation, your school is assigned to the intervention group, you should plan to release your staff to attend training administered by the Delivery Team, which is scheduled for the fourth week of September 2023.

What is Concept Cat?

The Concept Cat approach teaches children concepts such as 'first', 'wide' and 'empty'. This is designed to be taught in a structured and engaging manner with explicit teaching of vocabulary followed by implicit teaching in play-based learning. The teaching process includes staff acting out a scripted story with a toy cat (that we will send to each participating setting). Small changes are made to the environment so that children have chances to experience the new word an increased number of times. For instance, if the target word for the week was 'empty', then sand and water trays would be out that week. Families are engaged with simple home activities which will be available in a number of community languages.

This sequence is designed to be accessible to a wide range of children, including those with limited language. One word is taught per week, allowing children an opportunity to develop deep understanding. Rather than general vocabulary, Concept Cat specifically teaches early verbal concepts (such as 'before', 'early' and 'through') that are central to maths and science curricula. This is underscored by the notion that learning these early verbal concepts will

---

[1] The new network of Stronger Practice Hubs, launched in Autumn 2022, is part of the Department for Education's (DfE) **Early Years Education COVID-19 Recovery Package**. They aim to address the impact of the pandemic on young children by supporting early years settings to build local networks and share evidence-informed practices to ultimately improve the quality of education and care. As Evidence Partner, the EEF will support the launch of the network.

consequently improve maths and science attainment at Key Stage 1. The delivery of the full programme is expected to span 30 weeks.

What is involved in delivering Concept Cat?

Settings that participate in the programme will receive the following:

- resources needed to run Concept Cat;
- a free three-hour training session for a lead practitioner and a one-hour training session for all staff, both delivered remotely by Stephen Parsons and Anna Branagan;
- 50% of funding needed to cover staff attendance in the said training sessions;
- visits from specially trained ‘Concept Cat Coaches’ every half-term to provide modelling and in-school support sessions, accompanied by an additional session in the first half-term;
- an hour-long online peer support session for lead practitioners, administered every half-term and facilitated by Concept Cat Coaches; and
- a £400 incentive provided in two tranches: £200 at the beginning of the academic year 2023/2024 once pupil baseline testing is finished, with the remaining £200 provided once all evaluation activities have been completed

More information about this programme is available on the Concept Cat programme page at **https://tinyurl.com/ConceptCatEEF**.

What will the evaluation entail?

The evaluation will involve 90 early years settings in England. Half of the settings will be randomly assigned to receive the Concept Cat training and subsequently implement the programme (hereafter called the intervention group). The remaining half will continue with business as usual (hereafter called the control group). Random allocation is essential to the evaluation as it is the best way of investigating the effect of Concept Cat training and implementation on children’s language development. Note that, if a setting is allocated to the control group, they will still receive the Concept Cat training, albeit in the following academic year (i.e., 2024-2025). It is important that settings understand and consent to this process.

At the start of the study, the evaluation team will ask settings to distribute information letters to parents or carers of participating children aged three to four (i.e., Foundation 1, the year before children enter Reception classes). Although the programme takes a whole-setting approach, this year group has been chosen for the purposes of this evaluation, in order to minimise burden on settings.

Where parents or carers have given their consent to include their children in the study, the evaluation team, via its sub-contractor Elklan, will liaise with settings in the first half-term of academic year 2023-2024 to arrange for baseline test administration. Elklan will use the Clinical Evaluation of Language Fundamentals® Preschool-2 (CELF Preschool-2) for baseline assessment. This assessment will take about 20-25 minutes per child. Elklan will repeat this assessment, as well as the Early Numeracy Assessment from the Early Years Toolbox, with the same children at the end of the academic year. These assessments will be scheduled based on your availability. All administrators will be checked against the Disclosure and Barring Service (DBS).

To measure the impact of Concept Cat on children’s early language and numeracy development, the evaluation team will use setting- and child-level data, provided directly by settings, in conjunction with baseline and endline results of the CELF Preschool-2 and the Early Numeracy Assessment from the Early Years Toolbox. For a more in-depth understanding of how the programme is implemented in participating settings, the evaluation team will select a sample of intervention group settings, where they will conduct interviews with staff and observe two Concept Cat sessions during the academic year. In addition, setting staff in both the intervention and control group will be asked to complete online surveys at the start and end of the academic year. Information sheets, privacy notices, and withdrawal forms will be shared with staff before any data is collected.

For more information on the data that will be collected and shared over the course of this evaluation, please see *How will data be used?*).

Is my setting eligible to participate?

Both school-based and private, independent, and voluntary (PVI) settings can participate, as long as they:

- have a minimum of 15 children aged three to four (i.e., in Foundation 1) enrolled to attend for at least 15 hours a week in the academic year 2023/2024;

- coordinate schedules so that Elklan may administer all baseline assessments prior to randomisation (which will take place in September 2023);
- agree to participate fully in the evaluation, including completing the programme (as outlined above) if selected to be in the intervention group, and completing all evaluation requirements (both intervention and control group) in the academic year 2023/2024;
- do not have any staff who have attended the Word Aware Early Years training in the last three years; and
- have not implemented Concept Cat in the last two years.[2]
- have not accessed Concept Cat resources through Lift Lessons (liftlessons.co)

Since this study is part of the DfE's Stronger Practice Hub initiative, each participating setting will be allocated to a Stronger Practice Hub located either in the West Midlands or the North West. As such, initial geographical restrictions may apply when selecting settings.

Children will be eligible for inclusion in the study[3] (i.e., have their data collected) if they:

- are in Foundation 1 and aged three to four years old when they begin Reception in academic year 2024/2025; and
- are registered to attend the setting for a minimum of 15 hours per week.

On the other hand, a child cannot be recruited into the study if:

- their parent has indicated that they wish to withdraw their child's data from the study; or
- at baseline assessment, the child is judged by the setting staff to be unable to sit and follow a short adult-selected task, since this would prevent them from completing the assessments.

If you have any queries about these requirements, please contact Elena Rosa Brown (erbrown@randeurope.org) or Miguel Subosa (msubosa@randeurope.org) of the Evaluation Team.

How will data be used in this study?

This study will be conducted in compliance with current data protection legislation. As such, all child data and any other personal data used for the project will be treated with the strictest confidence and will be used and stored in accordance with the General Data Protection Regulation (GDPR) (2018) and the Data Protection Act (2018). The types of data that will be collected and how that data will be managed are described in detail below.

RAND Europe will act as data controllers and will be the main point of contact for any matters relating to the protection of all personal data once the recruitment is completed and will make decisions about how and what personal data is used in the evaluation.

Setting- and child-level data

Better Communication CIC, acting as data processors, will collect administrative setting- and child-level data directly from settings. Before collecting any data, Better Communication CIC will co-sign a Data Sharing Agreement (DSA) with individual settings. In addition, Better Communication CIC will provide information sheets, privacy notices, and withdrawal forms to settings. Settings will then distribute these documents to parents or carers of participating children. Once the necessary documents have been signed and returned by parents or carers to settings, data collection can begin.

Settings will be invited to share the following setting- and child-level information with Better Communication CIC:

- Full name
- Date of birth
- Gender
- Home postcode

---

[2] If a setting has the Word Aware 2 Early Years book but are not actively using it, they are still eligible. Similarly, settings are eligible if individual children have been participating in speech and language programmes containing Concept Cat. In this case, however, the Delivery team should be alerted.

[3] The programme should, however, be used as a whole setting approach, not just with those children participating in the evaluation.

- Early years setting postcode (or Local Authority Establishment number and the school's Unique Reference Number, if school-based)
- Attendance data (number of hours and frequency)

After collecting the data listed above, Thinking Talking will share the information with RAND Europe, through a secure electronic file transfer software (i.e., Egress). Data sharing between Thinking Talking and RAND Europe will follow the stipulations of a DSA signed by both parties.

Elklan, RAND Europe's testing partner, will act as data processors throughout the evaluation and will process data in accordance with the ways and purposes set by RAND Europe. These ways and purposes are specified in Elklan's sub-contract.

Using the CELF Preschool-2 (including its *Basic Concepts* and *Concepts and Following Directions* sub-tests), Elklan will collect standardised assessment data before randomisation. After delivery of Concept Cat, Elklan will administer the same test (i.e., CELF Preschool-2) together with the Early Numeracy Assessment from the Early Years Toolbox. The baseline and endline test scores of participating children will be shared by Elklan with RAND Europe via Egress.

Administrative setting- and child-level data, as well as attainment data collected by Elklan, will be matched against the following data extracted from the National Pupil Database (NPD):

- Early Years Pupil Premium (EYPP) status
- English as an additional language (EAL) status

Observations, surveys, and interviews

The Evaluation Team, namely the University of York, may ask to visit your setting to conduct observations of the delivery of Concept Cat. This will allow the Evaluation Team to understand different elements that influence implementation as well as adaptations made by participating settings to the intervention.

In addition, the evaluation team will be asking parents or carers of participating children, as well as designated setting staff, to complete a short survey at the start and at the end of the academic year. The University of York will share links to electronic versions of these surveys with settings for dissemination to children's parents or carers, as well as setting staff. Select staff may also be invited to participate in interviews.

RAND Europe and the University of York will act as joint controllers of data collected through observations, surveys, and interviews. The latter will share said data with the former through Egress. This data sharing will follow the stipulations of a DSA co-signed by the two parties.

Data removal

Any setting that wishes to have their data removed at any point before the study has concluded should contact Elena Rosa Brown at **erbrown@randeurope.org** or Miguel Subosa at **msubosa@randeurope.org**, who will notify all relevant parties to remove the data.

What will happen to the data at the end of the study?

At the end of the study, RAND Europe will submit data to the Office for National Statistics Secure Research Service (ONS SRS) for archiving in the EEF data archive (managed by EEF's archive manager)[4] and will include data only individually identifiable to the Department for Education (DfE). At this point, the EEF and DfE will act as data controllers for the archive, along with contractors appointed to manage the archive. In future, this data may be matched to the National Pupil Database (NPD) and other administrative data for subsequent research. Anonymous data will be kept indefinitely by the University of York.

Access to information will be restricted on a need-to-know basis and security arrangements are regularly reviewed to ensure their continued suitability. The evaluation team will collect and store all personal and special category data in

---

[4] You can find more information about the EEF archive on the EEF's website: **https://educationendowmentfoundation.org.uk/privacy-notices/privacy-notice-for-the-eef-data-archive**.

accordance with the Data Protection Act (2018) and UK GDPR requirements. No personal information collected as part of this study will be transferred outside of the European Economic Area (EEA).

All results will be anonymised so that no setting, staff member, or individual child will be identifiable in reports or publications that use the data emerge from this study.

Responsibilities of the Delivery Team (Thinking Talking and Better Communication CIC)

- Organise a DSA with settings for the collection of administrative setting- and pupil-level data.
- Collect administrative setting- and pupil-level data directly from settings.
- Conduct training and support activities for Concept Cat implementation in settings, as described above. This includes coordinating one-to-one coaching sessions.

Responsibilities of the Evaluation Team (RAND Europe and the University of York)

- RAND Europe will act as the first point of contact for any questions about the evaluation.
- RAND Europe will conduct the random allocation of settings to the intervention and control group for the academic year 2023/2024.
- RAND Europe will inform settings whether they have been allocated to the intervention or control group.
- RAND Europe will provide information sheets, privacy notices, and withdrawal forms for settings to send to parents and carers.
- RAND Europe, through its sub-contractor Elklan, will organise baseline and endline assessments with participating children.
- University of York will conduct baseline and endline surveys with setting staff.
- University of York will conduct setting visits to a selection of settings during the study.
- Both RAND Europe and the University of York will analyse data from the study and co-author the final report disseminate research findings through the EEF report, at conferences, and through academic papers.

Responsibilities of participating settings

- Sign and return a DSA that will be provided by the Delivery Team.
- Provide the contact details of a main and secondary contact person during the study.
- Send and collect information sheets, privacy notices, and withdrawal forms to and from parents and carers of participating children.
- Inform Better Communication CIC about any children whose parents have opted to withdraw them from the study or who leave the setting before the trial has been completed.
- Ensure that designated staff attend all required training and support sessions, if randomised into the intervention group.
- If randomised into the control group, setting staff will be expected to continue with usual practice. Control settings will be eligible to receive the programme in the 2024-2025 academic year
- Facilitate the administration of CELF Preschool-2 and the Early Numeracy Assessment, which will be conducted by RAND Europe’s testing partner, Elklan, before randomisation and after the delivery of Concept Cat.
- Ensure that staff and parents complete baseline and endline surveys administered by the Delivery Team.
- For a selected sub-sample: facilitate two visits by one or two researchers from the University of York for setting observations and interview administration.
- Inform the Delivery Team, at the earliest opportunity, if there are support or operational issues that could prevent the delivery of Concept Cat.

i. Key Dates

Sept - Oct 2023	• Baseline assessment (CELF Preschool-2)
Sept 2023, Week 3	• Randomisation of settings into treatment and control groups • Inform settings of randomisation allocation
Sept 2023, Week 4	• Delivery of Concept Cat training
Sept 2023 - Jun 2024	• Concept Cat programme delivery
Jun - Jul 2024	• Endline assessment (CELF Preschool-2, Early Numeracy Assessment)

Please provide the following information about your setting:

Name of setting	
Type of setting	[ ] School-based [ ] PVI
Local authority	
County	
Unique Reference Number (composed of six digits) *Indicate only if your setting is school-based. Otherwise, write N/A.*	
Setting administrator's email address	
% of children on Early Years Pupil Premium (EYPP)	
% of children for whom English is an additional language (EAL)	
Total number of children in setting in academic year 2023/2024	
Expected number of children aged three to four enrolled in the setting for more than 15 hours per week in academic year 2023/2024	

Main contact persons during the study		
Role	**Contact name**	**Email address and contact number**
Primary contact *This will be a member of the setting's Senior Leadership Team who will assist in implementing the study.*		

Secondary contact		

CONCEPT CAT AGREEMENT

This form is to be completed by a member of the senior leadership team of the setting. Please read carefully and, if you are happy to take part, please initial beside each statement and complete the subsequent sections and complete and sign the Data Sharing Agreement in the Annex[5].

1. I confirm that I have read and understood the Memorandum of Understanding for the above project and evaluation and have had the opportunity to ask questions
   Initials
2. I agree to release staff as required above to receive the training if randomised to receive the Concept Cat programme and to support them in the implementation of the programme within the setting
   Initials
3. I have completed the setting baseline information form above.
   Initials
4. I permit the publication of anonymised data collected from this study.
   Initials
5. I agree to distribute information sheets, privacy notices, and withdrawal forms to parents and to encourage parents to allow their child's participation in the study. The setting will keep a record of requests to withdraw and inform the Evaluation Team of any such request.
   Initials
6. I agree to facilitate researchers from the Evaluation and Delivery Teams in collecting evaluation data. This includes setting staff completing two surveys during the year and facilitating setting visits and staff interview(s) for the purposes of understanding the implementation of the programme, if requested, and distributing a parent survey.
   Initials
7. I agree to facilitate visits of assessment administrators from Elklan at the start and end of the academic year 2023/2024 to allow for assessment of children participating in the evaluation.
   Initials
8. I agree to notify the Delivery Team and the Evaluation Team, at the earliest opportunity, if the setting has any issues that could affect the effective implementation of the programme, if the setting is allocated to the intervention group.
   Initials
9. I agree to notify the Delivery team and Evaluation Team immediately if the setting has to withdraw from the project for operational or other unavoidable reasons and wherever possible still allow assessment data to be collected for the evaluation.
   Initials

---

[5] Please add the date of signature, the legal name, and address of the setting and data privacy contact at the top of the Data Sharing Agreement and then sign.

10. I agree to notify, at the earliest opportunity, the Delivery Team and Evaluation Team of any changes to participating staff or to the project lead or if any participating children leave the setting. Initials

11. Members of staff involved in the project will provide valid email addresses and telephone contact numbers to the Evaluation and Delivery teams and agree to check communications regularly during the period of the research. Initials

I confirm I have read the Memorandum of Understanding for the Concept Cat study. I agree the setting will take part in the study and I accept the terms and conditions outlined in this document.

Setting Name: ____

Setting Telephone Number: ____

Setting Address: ____

Setting URN: ____

Setting Lead Name: ____

Setting Lead Signature: ______________________________________Date:____/____/______

Setting Lead Email Address: ____

Lead Practitioner Name: ____

Lead Practitioner Email Address:

____

Project Lead Job Title:____

Appendix M: Privacy notice

1. Background

The Education Endowment Foundation (EEF) supported by the Department for Education (DfE) is funding a study to evaluate a programme designed to improve early years language development.

Concept Cat is a whole class teaching methodology for early verbal concepts (part of the widely used book ‘Word Aware 2’) and delivered by authors Stephen Parsons and Anna Branagan, working with a partner organisation (Better Communication CIC).

The Concept Cat approach teaches children concepts such as ‘first’, ‘wide’ and ‘empty’. This is designed to be taught in a structured and engaging manner with explicit teaching of vocabulary followed by implicit teaching in play-based learning. The teaching process includes staff acting out a scripted story with a toy cat (that we will send to each participating setting). Small changes are made to the environment so that children have chances to experience the new word an increased number of times. For instance, if the target word for the week was ‘empty’, then sand and water trays would be out that week. Families are engaged with simple home activities which will be available in a number of community languages.

This sequence is designed to be accessible to a wide range of children, including those with limited language. One word is taught per week, allowing opportunity for deep understanding to develop. Rather than general vocabulary, Concept Cat specifically teaches early verbal concepts (such as ‘before’, ‘early’ and ‘through’) core to the curriculum of maths and science, with the ultimate aim of improving maths and science attainment at Key Stage 1.

2. About the research study

RAND Europe and the Department of Education at the University of York are the independent evaluators of the Concept Cat programme.

The aim of the evaluation is to understand whether Concept Cat has had an impact on child outcomes (the ‘impact’ evaluation), which will be achieved by comparing data from children receiving the intervention to similar children that did not receive the intervention within a given timeframe. This will help us understand if and how the Concept Cat programme makes a positive impact on children. The intended outcomes of the programme include: early language development and early numeracy.

The gathered data is therefore expected to make an important contribution to understanding what works in early language development and pupil attainment, as well as effective teaching strategies to build conceptual vocabulary.

This privacy notice is for the parents/guardians of children in Early Years settings which are participating in the Concept Cat evaluation. It sets out the ways in which RAND Europe and the University of York, use, store, and share your data and your child’s data. It also sets out how long we keep this data and what rights you have in relation to this data under the General Data Protection Regulation (GDPR).

3. Who is involved in the research study?

The evaluation is being funded by the EEF, an organisation that funds research into education, supported by the DfE. RAND Europe and the University of York are doing the evaluation research (the ‘evaluation team’). RAND Europe and the University of York will be supported by Elklan, who will collect data on child outcomes through standardised assessment tools. Elklan’s trained professionals will come in to work with children, in coordination with the classroom teacher. This means that they will not have direct contact with your child without the presence or awareness of the teacher or a staff member from within the setting. Thinking Talking have designed the programme and Better Communication CIC are responsible for its delivery (the ‘delivery team’).

4. Where do we get your data from and what data do we collect?

The evaluation team collect personal data[6] about your child for the purposes of conducting the study, to evaluate the impact and effectiveness of the Concept Cat programme, as well as to prepare a report about the study.

Your child's Early Years setting will share the following information about your child with University of York:

- Full name
- Date of birth
- Gender
- Home postcode
- Early years setting postcode (or Local Authority Establishment number and the school's Unique Reference Number, if school based)
- Attendance data (number of hours and frequency)

This information will be used for future data linkage to the National Pupil Database (NPD)

- Early Years Pupil Premium (EYPP) status
- English as an Additional Language (EAL) status
- Special Educational Needs (SEN) status

This data will be matched to the following child outcomes, which will be collected from participating children via assessments conducted by Elklan:

- Clinical Evaluation of Language Fundamentals Preschool-2 (CELF Preschool-2)
    - Basic Concepts
    - Concepts & Following Directions
- Early Years Toolbox: Early Numeracy Assessment

Elklan will administer the CELF Preschool-2 at the beginning of the study. At the end of the study, Elklan will administer the same test, along with the Early Numeracy Assessment.

We will also be asking you as a parent or guardian to participate in a short survey. We will do this at the start and end of the study.

A Data Sharing Agreement (DSA) will be put in place with your setting before any data is collected. The DSA will fully specify the data to be collected and for what purpose, who will collect it, how it will be shared, how it will be stored, who will have access to it and how long we will keep it for.

5. Who will hold your data?

During the period of the evaluation, no one outside the evaluation team will have access to your data. This project is one of several in the Stronger Practice Hub initiative, an aspect of the Department for Education (DfE) Early Years Educational Recovery Programme. As such, RAND Europe and the University of York will act as joint controllers for the data throughout the evaluation period. This means that both parties will share responsibility for ensuring your data and your child's data remain protected throughout the evaluation, including making sure that it is stored and shared securely. After the end of the evaluation, the DfE and EEF will become sole controllers of the data.

RAND Europe will lead on the processing of the pupil data – that is, they will be the main point of contact for any matters relating to the protection of all personal data and will make decisions about

---

[6] By 'personal data', we mean any information about you or your child which could be used to identify you, such as your name.

how and what personal data is used, in accordance with the purposes agreed with the DfE and EEF. The University of York will lead on the processing of all staff data.

Better Communication CIC (the delivery team, funded by the EEF) as well as Elklan (commissioned by RAND Europe) will act as data processors throughout the evaluation period and will process the data, in accordance with the ways and purposes set by the joint controllers.

RAND Europe and the University of York are both registered with the Information Commissioner's Office, with registration numbers Z6947026 and Z4855807, respectively.

6. How will we collect, share and use your data?

As part of the evaluation, **Better Communication CIC** (the delivery team) will provide RAND Europe with recruited settings' contact details (i.e., email addresses, setting's postal address). **RAND Europe** and **Elklan** will use settings' contact details to collect information from settings about all the children that take part in Concept Cat (as specified in Section 4). **RAND Europe** will also ask Elklan to collect data on assessed child outcomes (as specified in Section 4). **Settings** and **Elklan** will provide this information by using an Excel data collection form provided by the evaluation team. This form will be shared via secure file transfer (e.g., Egress). The evaluation team will use the data provided by settings and Elklan on children participating in Concept Cat and compare their outcomes to similar pupils that did not receive the programme to measure if there is any difference between the two groups.

The **University of York** will ask Better Communication CIC (the delivery team) to collect information from settings about key staff. **Settings** will provide this information by using an Excel data collection form. This form will be shared via secure file transfer (e.g. Egress). **The University of York** will use the data provided by settings on key staff to invite them to take part in a short survey and/or interview. **RAND Europe** may also use this data in its research analysis to understand the impact of the programme.

At the end of the study, **RAND Europe** will submit the data in pseudo-anonymised format to the Office for National Statistics Secure Research Service (ONS SRS) for archiving in the **EEF** data archive (managed by EEF's archive manager). [7] This data will only be identifiable to the **DfE** and may be matched to the National Pupil Database (NPD) and other administrative data during subsequent research. The EEF and DfE will act as data controllers for the archive, along with contractors appointed to manage the archive.

No individual will be named in any report or other output for this project.

7. What is our legal basis for processing your data?

The legal basis for RAND Europe to process your child's personal data is legitimate interests, as detailed in Article 6(1)(f) of the UK GDPR. The legal basis for processing your child's special category data[8] is because it is necessary for archiving purposes in the public interest, scientific, or historical research purposes as detailed in Article 9(2)(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, in addition to the completion of an application to the RAND internal review board for ethical approval (see Section 8 below). RAND Europe will process

---

[7] You can find more information about the EEF archive on the EEF's website: **https://educationendowmentfoundation.org.uk/privacy-notices/privacy-notice-for-the-eef-data-archive**.

[8] 'Special category' data is personal data that needs more protection because it is sensitive, for example health or ethnicity data.

only what is required to meet these legal bases and will ensure security and safeguards are in place to protect the information.

The legal basis for the University of York, EEF and DfE to process your child's personal data is where it is necessary for the performance of a task carried out in the public interest as set out in Article 6(1)(e) of the UK GDPR. The specific legislation which allows this is Section 10 of the Education Act 1996. The legal basis for processing special category data is for reasons of substantial public interest as detailed in Article 9(2)(g) of the UK GDPR.

8. Ethical Approval

The study has received ethical approval from RAND Europe's internal review board (HSPC ID: 2023-N0010, dated 14 March 2023).

9. How do we keep your data secure?

We take information security extremely seriously and we have appropriate technical and organisational measures to protect personal data and special category data. Access to information is restricted on a need-to-know basis and security arrangements are regularly reviewed to ensure their continued suitability. The evaluation team will collect and store all personal and special category data in accordance with the Data Protection Act (2018) and UK GDPR requirements.

No personal information collected as part of this study will be transferred outside of the European Economic Area (EEA).

For further information about how we will share your data, please refer to Section 6 above.

10. How long will we keep your data?

All individually identifiable data held by RAND Europe will be destroyed one year after the end of the study (2026). All individually identifiable data held by the University of York will be destroyed 5 years after the end of the study (2030). Data in the EEF's archive in the ONS SRS will include data only individually identifiable to the Department for Education (DfE), the government department responsible for children's services and education, and is kept indefinitely for the purposes of future research. Anonymous data will be kept indefinitely by the University of York.

11. What rights do you have in relation to your data?

Under the General Data Protection Regulation, you have rights in relation to yours and your child's data, including a right of access to the data (Article 15 of the GDPR), a right to rectification (Article 16 of the GDPR), erasure (in certain circumstances; (Article 17 of the GDPR)), restriction (Article 18 of the GDPR) and objection (Article 21 of the GDPR).

You are free to withdraw your child from the study at any time during the course of the study (information already collected about you will be retained). In this event, please tell your child's setting who will communicate this to the research team, or you can contact us directly using the details below.

If you wish to exercise any of the rights set out above in connection with this research project, please email us or write to us at the contact addresses below.

12. Right to complain

If you are unhappy with the way in which the research team has handled your personal data, we ask that you contact us in the first instance, to enable us to resolve your concerns. If you remain dissatisfied, you have the right to make a complaint to the Information Commissioner's Office (ICO), the UK supervisory authority for data protection issues, by emailing **casework@ico.org.uk**. Further information about how you can exercise your right to complain is available at **www.ico.org.uk**.

13. Changes to our privacy notice

We may change this Privacy Notice from time to time. If we make any significant changes in the way we treat your personal information, we will make this clear by contacting setting and ensuring they provide you with an updated version of this Privacy Notice.

14. How to contact us

If you have any questions about this privacy notice or concerns about how your data is being processed, you can get in touch by:

- Sending an email to the evaluation team FAO, Elena Rosa Brown (**conceptcatevaluation@randeurope.org**)
- Calling the evaluation team by ringing +44 1223 353 329 and quoting Ref: 022807.015 "Concept Cat"
- Contacting RAND Europe's Data Protection Officer, Rani Viknaraja (rviknara@randeurope.org) and quoting Ref: 022807.015 "Concept Cat" or in writing to Data Protection Officer, RAND Europe, Westbrook Centre, Milton Road, Cambridge, CB4 1YG, UK
- Contacting University of York's Data Protection Officer at **dataprotection@york.ac.uk** quoting Ref: "Concept Cat Evaluation".

Appendix N: Information sheet and withdrawal form

CONCEPT CAT

PARENT/GUARDIAN INFORMATION SHEET

Dear Parent(s)/Carer(s),

Concept Cat is a new project that will be coming into your child's nursery. **Concept Cat** involves teaching children important words through songs, stories and play. The teaching is for all children and will be done by your child's usual teacher. We think that Concept Cat helps children learn, but we want to find out if it really does. To know we need to measure children's knowledge of words and maths. We are planning for a trained professional to come into nursery and talk to children in September 2023 and again in June/July 2024.

We gained full support from your child's nursery to deliver and evaluate Concept Cat, but if you don't want us to talk to your child to find out how they are learning words, you need to tell us. You can do this by filling in the 'I don't want my child to take part' form on page 6 or by emailing us on conceptcatevaluation@randeurope.org.

The full details of the **Concept Cat** project and its evaluation are below.

What is Concept Cat?

Concept Cat has been developed by Stephen Parsons and Anna Branagan, two Speech and Language Therapists, who are working with Better Communication CIC. Concept Cat is a story and play-based approach for teaching words that we call ‘concepts’. Concepts are really important words such as ‘first, whole and least’ which can be really tricky to learn. Concept Cat is a whole-class approach and so all children will be able to get involved. Each week a short story will be acted out and then the children will play some extra games and sing a song. Families can get involved when teachers send home the words and simple activities. Children will be taught early verbal concepts such as ‘first’, ‘wide’, ‘empty’ and ‘through’ in a structured and engaging manner that will involve storytelling and play. Such concepts have been selected as they are considered core to the curriculum of maths and science. The ultimate aim is to improve maths and science attainment at Key Stage 1. Children will be taught one word per week. This is because we want to allow opportunity for deep understanding to develop.

The evaluation

We want to find out if Concept Cat helps children learn words faster and if their numeracy skills develop so we will be running an evaluation of Concept Cat as it is delivered. The evaluation is being funded by the Education Endowment Foundation (EEF), an organisation that funds research into education, and the Department for Education (DfE). RAND Europe and the University of York with support from Elklan will be working together evaluate Concept Cat.

Settings that are selected to participate in the evaluation will be randomly assigned to either receive Concept Cat in the academic year 2023/2024 (the intervention group) or to continue practice as usual (the control group). There will be an equal chance of being allocated to each group, with half the settings in the intervention group and half in the control group.

The success of the programme will be measured using well-established language and numeracy assessments with selected children in the setting. In order to understand the impact of this approach, trained professionals from Elklan will visit nurseries at the start of the academic year and work with children one to one in September and October 2023 and then again in June/July 2024. They will talk to the children and look at some pictures for about 20-25 minutes each.
We are also keen to hear from parents and/or guardians so York University will also be asking you to participate in a short survey. They will do this twice: once at the start and once again at the end of the school year. The purpose of the survey will be to understand your own experience of having your child participate in the programme, along with the types of learning activities that you and your child typically engage with at home (e.g., the simple activities that the teachers send home in order for the children to practise the new words).

What will taking part in the study involve?

Participation is voluntary

It is up to you if you want your child involved but you have to tell us if you don't want your child to take part.

If you choose not to include your child in the evaluation, this will have no impact on the usual care your child receives in their nursery, or on whether the nursery uses the Concept Cat programme or not. If you change your mind at any point during the study, you will be able to withdraw your child's participation without having to say why.

Anonymity and confidentiality

We won't tell other people about anything specific about your child.

We assure you that all information relating to you and your child will be treated in the strictest confidence and processed in accordance with General Data Protection Regulations (GDPR). Any sharing of the research findings (e.g. through reports, academic publications and presentations) will be in an anonymous format. **We will not mention your child's nursery or your child's name in any report or publication coming from the research**.

Storing and using your data

How we will look after the information we have about your child.
We will also ask your child's nursery to share some additional details about your child to help with our analysis.

All information about the nursery and your child will be stored securely by code number on a password protected computer. Any identifiable information will be stored separately from the data and will be destroyed at the end of the evaluation.

At the end of the evaluation your child's data will be shared with the Department for Education, the Education Endowment Foundation (EEF), FFT Education (EEF's data processor for the EEF data archive) in order to place it into the data into the National Pupil Database. This will allow your child's data to be used in a pseudonymised form (i.e. names are not shared) during subsequent research. We may also share anonymised data (i.e. data with no identifiable data that can be linked to your child) with other research teams.

To ensure that all processing is fair and lawful, the evaluation team have completed a Data Protection Impact Assessment and have sought and obtained ethical approval from an internal review board. Further details relating to GDPR, third parties, and confidentiality are provided in the Privacy Notice. Please read this information carefully. You will need to confirm that you have read and understood this information before proceeding.

Please note: If we gather information that raises concerns about your/your child's safety or the safety of others, or about other concerns as perceived by the researcher, the researcher may pass on this information to another person.

What do I need to do now?

If you are happy with all of this, you don't have to do anything. If you are happy for your child to take the short assessments and for your child's nursery to share your child's data with us, as described above, then you do not need to do anything.
If you aren't happy to get involved, then you need to tell us.
If you do NOT want to share your child's data as described above, please complete the attached 'Parent/guardian: I don't want my child to take part' form and hand it to your teacher who will then let us know.

You are free to withdraw your child's information for use in this evaluation at any time between now and August 2024. You can do that by either writing to the setting (who will then contact the evaluation team) or emailing the evaluation team directly at: **conceptcatevaluation@randeurope.org**.

Please note: Children whose parents have indicated that they wish to withdraw their child's data from the evaluation (i.e. the assessments) may still participate in the programme.

Questions or concerns

If you need to know anything else, please contact us.

If you have any questions about this participant information sheet or concerns about how your child's data is being processed, please feel free to contact the lead researcher, Elena Rosa Brown at **conceptcatevaluation@randeurope.org**. If you still have questions about the way your child's data will be processed, please contact our Data Protection Officer at **REdpo@randeurope.org**.

Thank you for taking the time to read this information.

CONCEPT CAT

PARENT/GUARDIAN: I don't want my child to take part

Only fill this form in if you do not want your child to take part in the assessments

If you are happy for your child to take part in the evaluation (i.e. the assessments) and for your setting to share your child's data with us, as described in the information sheet, then you do not need to complete this form.

If you do **NOT** want to your child to participate in the evaluation, please complete and return this form to your child's setting. Your child may still participate in the programme, even if you indicate that you wish to withdraw your child's data from the evaluation.

If you do not return a completed form, we will assume you are happy for your setting to share information for use in this evaluation and for your child to participate in the research.

☐ **I do NOT** want my child to participate in the research or for my child's data to be shared for the evaluation of the Concept Cat programme.

Child's Name (Please print clearly)…………………………………………………………………………………..............

Parent/Carer Name (Please print clearly)……………………………………………………………………………....

Signature ……………………………………………………………………………………

Date…………………………….

Setting Name ………………………………………………………………………………

Name of setting staff member processing the withdrawal …………………………………..

Hand in to nursery

Appendix O: Implementation dimensions

Logic Model Outcomes	Aspect of the Logic Model	Data Collection											Implementation Dimensions							
	Moderators: P = Practitioner factors S = Setting factors H = Home factors	Monitoring data			Observations of Training Delivery	Setting Observations	Embedded Setting Observations	Practitioner Interviews	Baseline Practitioner Survey and quiz	Endline Practitioner survey and quiz and manager Survey	Baseline Parent Survey	Endline Parent Survey	Fidelity/Adherence	Dosage	Quality	Reach	Responsiveness	Programme Differentiation	Monitoring of Control Condition	Adaptation
		Coach visits to settings	Group supervision	Training Logs / attendance data																
Outputs	Practitioners engage in initial training support and materials to acquire understanding of programme approaches, including not to teach opposites simultaneously (P, S)	🟦	🟦	🟦	🟦								🟩	🟩			🟩			
	Practitioner continuously deliver all elements of the approach (P, S)	🟦	🟦			🟦	🟦			🟦			🟩	🟩	🟩			🟩		
	Practitioners engage in interim support sessions with peers and Word Aware trainers (P, S)	🟦	🟦					🟦					🟩		🟩		🟩			
	Children experience increased explicit multi-sensory teaching of key conceptual vocabulary (explicit teaching) (P, S)					🟦	🟦			🟦			🟩	🟩			🟩			

	Children have daily opportunities to hear and explore target concepts within play activities (implicit teaching) (P, S)					🟦	🟦			🟦			🟩	🟩			🟩			
	In setting and at home, children have increased opportunities throughout the intervention period to review words that have been previously introduced (P, S, H)					🟦	🟦			🟦	🟦	🟦	🟩	🟩			🟩			
	Families have increased opportunities to learn how to support their child’s vocabulary							🟦		🟦	🟦	🟦			🟩		🟩		🟩	
Short-term outcomes	Teachers gain understanding of the importance of conceptual vocabulary (P, S)							🟦	🟦	🟦					🟩		🟩			🟩
	Teachers have the skills and resources to effectively teach conceptual vocabulary: repetition, opportunities for embedding learning and implicit and explicit teaching					🟦	🟦	🟦	🟦	🟦			🟩		🟩				🟩	
	Children show increased engagement with learning new words through participations in activities and adult/peer interactions					🟦		🟦		🟦						🟩	🟩		🟩	
	Children show understanding and use conceptual vocabulary at school and at home					🟦		🟦		🟦		🟦				🟩	🟩		🟩	
	Families demonstrate interest and engagement in pupil conceptual vocabulary development through regular attendance in the info sessions and watching videos									🟦		🟦				🟩	🟩		🟩	
	Families engage in word learning activities at home using the provided guidance and suggested activities									🟦		🟦				🟩	🟩		🟩	

Long-term outcomes	Quality teaching of early verbal concepts in daily practice					🟦	🟦	🟦		🟦			🟩			🟩				
	Increased conceptual receptive vocabulary							🟦		🟦		🟦				🟩				
	Increased early numeracy development							🟦		🟦		🟦				🟩				

Appendix P: Focus children selection criteria

Concept Cat: Choosing Focus Children

All children will be learning about the selected words but 'focus children' will get a bit more support.

Use the flowchart on the next page to help you pick the right children. Your Concept Cat Coach will help you choose the right children.

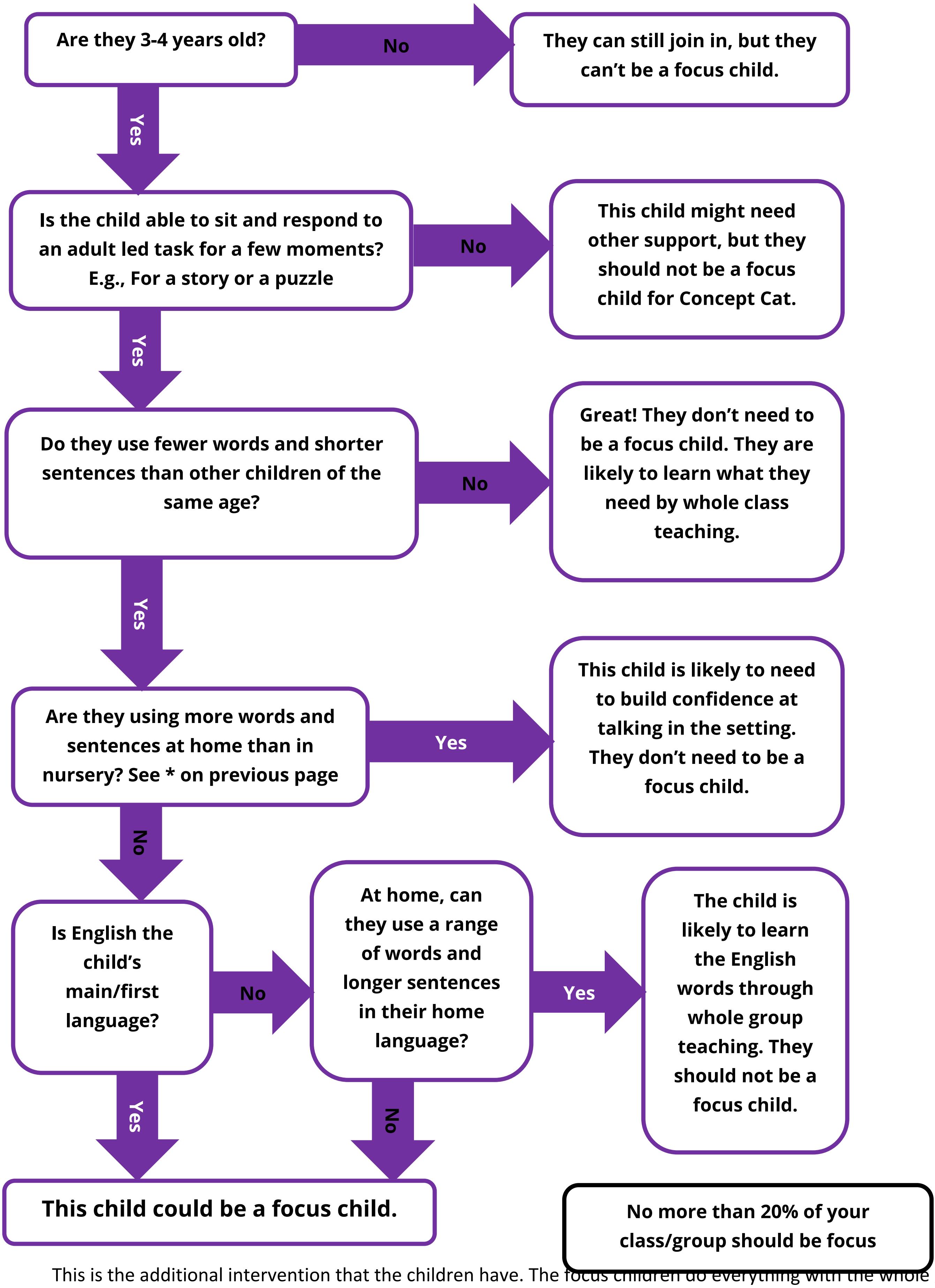
Ideally you want no more than 20% of your group to be focus children.

*** If a child isn't speaking much in your setting it is good to talk to the family to see how the child is talking at home. The sorts of questions to ask are:**

- **What is your child's longest sentence?**
- **What is their most advanced/grown up word?**
- **Do they use words such as 'big, under, more?'**

Think about how this compares with the other children of a similar age.

Focus children



Are they 3-4 years old?

No

They can still join in, but they can't be a focus child.

Yes

Is the child able to sit and respond to an adult led task for a few moments? E.g., For a story or a puzzle

No

This child might need other support, but they should not be a focus child for Concept Cat.

Yes

Do they use fewer words and shorter sentences than other children of the same age?

No

Great! They don't need to be a focus child. They are likely to learn what they need by whole class teaching.

Yes

Are they using more words and sentences at home than in nursery? See * on previous page

Yes

This child is likely to need to build confidence at talking in the setting. They don't need to be a focus child.

No

Is English the child's main/first language?

No

At home, can they use a range of words and longer sentences in their home language?

Yes

The child is likely to learn the English words through whole group teaching. They should not be a focus child.

Yes

No

This child could be a focus child.

No more than 20% of your class/group should be focus

This is the additional intervention that the children have. The focus children do everything with the whole group **as well** as this.

Monday	Repeat: gesture/sign, symbol, say it and song (1) Repeat: Concept Cat story or video (2)
Tuesday	Repeat a teach activity (3) Extra contact with families of focus children: • Face to face or by phone • Tell them about the word and how it has been taught in school/nursery • Talk about the suggested activity sent on Monday. • Encourage to send in a photo for the Concept book. Key message: repeat the word, experience the word, use the word in sentences
Wednesday	Go to the Concept Cat house, tell the Concept Cat story again (2) and model with objects
Thursday	Picky puppet with current word. Use the word with children in an activate activity (4)
Friday	Picky puppet with current word. Word bag with previous week's words.

Focus children that attend 3 days a week

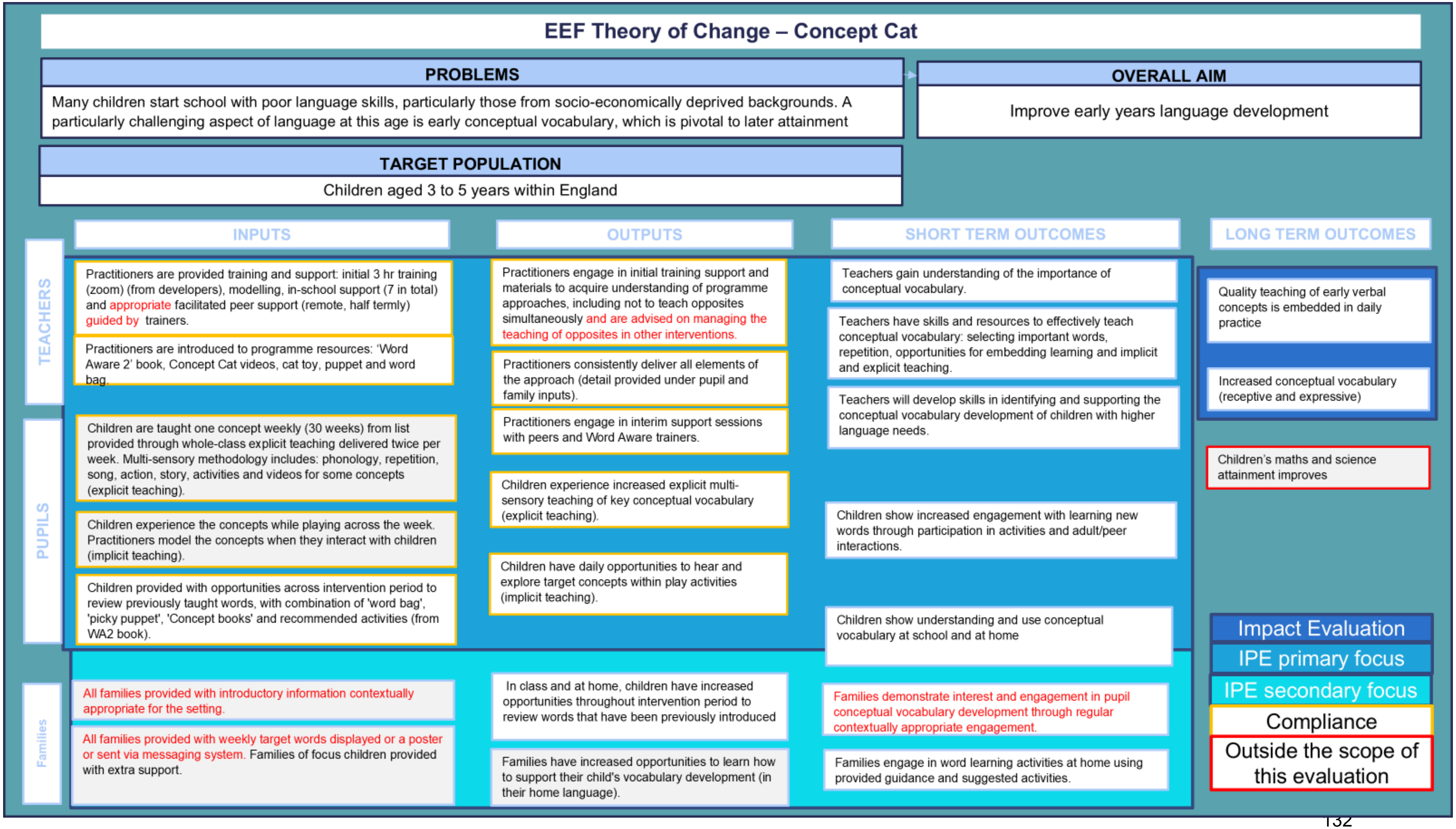
This is the additional intervention that the children have. The focus children do everything with the whole group **as well** as this.

Day 1	Repeat: gesture/sign, symbol, say it and song (1) Repeat: Concept Cat story or video (2)
Day 2	Repeat a teach activity (3) Go to the Concept Cat house, tell the Concept Cat story again (2) Extra contact with families of focus children: • Face to face or by phone • Tell them about the word and how it has been taught in school/nursery

	• Talk about the suggested activity sent on Monday. Key message: repeat the word, experience the word, use the word in sentences
Day 3	Picky puppet with current word. Take objects from the word bag. Talk about this week's concept and 2 other concepts. (5)

*Numbers relate to the steps in the teaching sequence in the WA2 book

Appendix Q: Updated Theory of Change



EEF Theory of Change – Concept Cat

PROBLEMS

Many children start school with poor language skills, particularly those from socio-economically deprived backgrounds. A particularly challenging aspect of language at this age is early conceptual vocabulary, which is pivotal to later attainment

OVERALL AIM

Improve early years language development

TARGET POPULATION

Children aged 3 to 5 years within England

INPUTS

TEACHERS

- Practitioners are provided training and support: initial 3 hr training (zoom) (from developers), modelling, in-school support (7 in total) and appropriate facilitated peer support (remote, half termly) guided by trainers.
- Practitioners are introduced to programme resources: 'Word Aware 2' book, Concept Cat videos, cat toy, puppet and word bag.

PUPILS

- Children are taught one concept weekly (30 weeks) from list provided through whole-class explicit teaching delivered twice per week. Multi-sensory methodology includes: phonology, repetition, song, action, story, activities and videos for some concepts (explicit teaching).
- Children experience the concepts while playing across the week. Practitioners model the concepts when they interact with children (implicit teaching).
- Children provided with opportunities across intervention period to review previously taught words, with combination of 'word bag', 'picky puppet', 'Concept books' and recommended activities (from WA2 book).

Families

- All families provided with introductory information contextually appropriate for the setting.
- All families provided with weekly target words displayed or a poster or sent via messaging system. Families of focus children provided with extra support.

OUTPUTS

- Practitioners engage in initial training support and materials to acquire understanding of programme approaches, including not to teach opposites simultaneously and are advised on managing the teaching of opposites in other interventions.
- Practitioners consistently deliver all elements of the approach (detail provided under pupil and family inputs).
- Practitioners engage in interim support sessions with peers and Word Aware trainers.
- Children experience increased explicit multi-sensory teaching of key conceptual vocabulary (explicit teaching).
- Children have daily opportunities to hear and explore target concepts within play activities (implicit teaching).
- In class and at home, children have increased opportunities throughout intervention period to review words that have been previously introduced
- Families have increased opportunities to learn how to support their child's vocabulary development (in their home language).

SHORT TERM OUTCOMES

- Teachers gain understanding of the importance of conceptual vocabulary.
- Teachers have skills and resources to effectively teach conceptual vocabulary: selecting important words, repetition, opportunities for embedding learning and implicit and explicit teaching.
- Teachers will develop skills in identifying and supporting the conceptual vocabulary development of children with higher language needs.
- Children show increased engagement with learning new words through participation in activities and adult/peer interactions.
- Children show understanding and use conceptual vocabulary at school and at home
- Families demonstrate interest and engagement in pupil conceptual vocabulary development through regular contextually appropriate engagement.
- Families engage in word learning activities at home using provided guidance and suggested activities.

LONG TERM OUTCOMES

- Quality teaching of early verbal concepts is embedded in daily practice
- Increased conceptual vocabulary (receptive and expressive)
- Children's maths and science attainment improves

Impact Evaluation
IPE primary focus
IPE secondary focus
Compliance
Outside the scope of this evaluation

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

https://educationendowmentfoundation.org.u

@EducEndowFoundn

Facebook.com/EducEndowFoundn