The Influence of Media Multitasking on Moroccan EFL Teach-ers' Reading Habits

"Nirmal Ghimire, Ph.D."

"2025-04-02"

List of the Variables (after and before Cleaning)

We are going to use the `raw_data_clean` further modeling. The `raw_data` is the raw version of uploaded data.

```
raw_data <- read_csv("C:/Users/ghimiren/Desktop/Reading Habits Study Time Diary Survey for Moroc
can Teachers/final_data.csv")
# Getting Rid of Second and Third Rows from the Data set
raw_data <- raw_data[-c(2:3), ]
# names(raw_data)
# Making column names consistent by using clean_names function and saving the data in a new tabl
e
raw_data_clean <- raw_data |>
  clean_names()
# Comparing column names after and before cleaning them
names(raw_data_clean)
```

```
  [1] "end_date"            "ip_address"          "progress"
  [4] "duration_in_seconds" "response_id"         "q2"
  [7] "q3"                  "q4"                  "q5"
 [10] "q6"                  "q7"                  "q8"
 [13] "q8_6_text"           "q9"                  "q10"
 [16] "q11"                 "q12_1"               "q12_2"
 [19] "q12_3"               "q12_4"               "q12_5"
 [22] "q12_6"               "q12_7"               "q12_8"
 [25] "q12_9"               "q12_10"              "q12_10_text"
 [28] "q13_1"               "q13_2"               "q13_3"
 [31] "q13_4"               "q13_5"               "q13_6"
 [34] "q13_7"               "q13_8"               "q13_9"
 [37] "q13_10"              "q13_11"              "q13_11_text"
 [40] "q14"                 "q14_5_text"          "q15"
 [43] "q16"                 "q17_1"               "q17_2"
 [46] "q17_3"               "q17_4"               "q17_5"
 [49] "q17_6"               "q17_7"               "q17_8"
 [52] "q17_8_text"          "q18_1"               "q18_2"
 [55] "q18_3"               "q18_4"               "q18_5"
 [58] "q18_6"               "q18_7"               "q18_7_text"
 [61] "q19"                 "q20"                 "q20_9_text"
 [64] "q21"                 "q22"                 "q23_1"
 [67] "q23_2"               "q23_3"               "q23_4"
 [70] "q23_5"               "q23_6"               "q23_7"
 [73] "q23_7_text"          "q24_1"               "q24_2"
 [76] "q24_3"               "q24_4"               "q24_5"
 [79] "q24_6"               "q24_6_text"          "q25"
 [82] "q26"                 "q27"                 "q27_8_text"
 [85] "q28"                 "q29_1"               "q29_2"
 [88] "q29_3"               "q29_4"               "q29_5"
 [91] "q29_6"               "q29_7"               "q29_8"
 [94] "q29_8_text"          "q30_1"               "q30_2"
 [97] "q30_3"               "q30_4"               "q30_5"
[100] "q30_6"               "q30_6_text"          "q31"
[103] "q32"                 "q33"                 "q33_16_text"
[106] "q34"                 "q35_1"               "q35_2"
[109] "q35_3"               "q35_4"               "q35_5"
[112] "q35_6"               "q35_7"               "q35_8"
[115] "q35_8_text"          "q36_1"               "q36_2"
[118] "q36_3"               "q36_4"               "q36_5"
[121] "q36_6"               "q36_6_text"          "q37"
```

```
dim(raw_data_clean)
```

```
[1] 702 123
```

```
## Get the list of all non-NA values in your data
data <- na.omit(unlist(raw_data_clean[c("q8_6_text", "q12_10_text", "q13_11_text", "q14_5_text",
"q17_8_text", "q18_7_text", "q20_9_text", "q23_7_text", "q24_6_text", "q27_8_text", "q29_8_tex
t", "q30_6_text", "q33_16_text", "q35_8_text", "q36_6_text")]))

# Convert the list to a table
# data.frame(data)
```

Breaking the Dataset in Two Different Studies

```
# create reading_data dataset
reading_data <- raw_data_clean %>%
  dplyr::select(
    duration_in_seconds, ip_address, q4, q6, q7, q8, q8_6_text, q12_1:q12_10_text, q13_1:q13_11_
text,
    q14, q14_5_text, q15, q16, q17_1:q17_8_text, q18_1:q18_7_text, q19, q20, q20_9_text, q21,
    q22, q23_1:q23_7_text, q24_1:q24_6_text, q25
  )

# create tv_internet_data dataset
tv_internet_data <- raw_data_clean %>%
  dplyr::select(
    duration_in_seconds, ip_address, q4, q6, q7, q8, q8_6_text, q12_1:q12_10_text, q13_1:q13_11_
text,
    q26:q37
  )

# save datasets as CSV files
# write.csv(reading_data, "reading_data.csv", row.names = FALSE)
# write.csv(tv_internet_data, "tv_internet_data.csv", row.names = FALSE)

# Checking the Variables
names(reading_data)
```

```
 [1] "duration_in_seconds" "ip_address"          "q4"
 [4] "q6"                  "q7"                  "q8"
 [7] "q8_6_text"           "q12_1"               "q12_2"
[10] "q12_3"               "q12_4"               "q12_5"
[13] "q12_6"               "q12_7"               "q12_8"
[16] "q12_9"               "q12_10"              "q12_10_text"
[19] "q13_1"               "q13_2"               "q13_3"
[22] "q13_4"               "q13_5"               "q13_6"
[25] "q13_7"               "q13_8"               "q13_9"
[28] "q13_10"              "q13_11"              "q13_11_text"
[31] "q14"                 "q14_5_text"          "q15"
[34] "q16"                 "q17_1"               "q17_2"
[37] "q17_3"               "q17_4"               "q17_5"
[40] "q17_6"               "q17_7"               "q17_8"
[43] "q17_8_text"          "q18_1"               "q18_2"
[46] "q18_3"               "q18_4"               "q18_5"
[49] "q18_6"               "q18_7"               "q18_7_text"
[52] "q19"                 "q20"                 "q20_9_text"
[55] "q21"                 "q22"                 "q23_1"
[58] "q23_2"               "q23_3"               "q23_4"
[61] "q23_5"               "q23_6"               "q23_7"
[64] "q23_7_text"          "q24_1"               "q24_2"
[67] "q24_3"               "q24_4"               "q24_5"
[70] "q24_6"               "q24_6_text"          "q25"
```

```
names(tv_internet_data)
```

```
 [1] "duration_in_seconds" "ip_address"          "q4"
 [4] "q6"                  "q7"                  "q8"
 [7] "q8_6_text"           "q12_1"               "q12_2"
[10] "q12_3"               "q12_4"               "q12_5"
[13] "q12_6"               "q12_7"               "q12_8"
[16] "q12_9"               "q12_10"              "q12_10_text"
[19] "q13_1"               "q13_2"               "q13_3"
[22] "q13_4"               "q13_5"               "q13_6"
[25] "q13_7"               "q13_8"               "q13_9"
[28] "q13_10"              "q13_11"              "q13_11_text"
[31] "q26"                 "q27"                 "q27_8_text"
[34] "q28"                 "q29_1"               "q29_2"
[37] "q29_3"               "q29_4"               "q29_5"
[40] "q29_6"               "q29_7"               "q29_8"
[43] "q29_8_text"          "q30_1"               "q30_2"
[46] "q30_3"               "q30_4"               "q30_5"
[49] "q30_6"               "q30_6_text"          "q31"
[52] "q32"                 "q33"                 "q33_16_text"
[55] "q34"                 "q35_1"               "q35_2"
[58] "q35_3"               "q35_4"               "q35_5"
[61] "q35_6"               "q35_7"               "q35_8"
[64] "q35_8_text"          "q36_1"               "q36_2"
[67] "q36_3"               "q36_4"               "q36_5"
[70] "q36_6"               "q36_6_text"          "q37"
```

```
dim(reading_data)
```

```
[1] 702  72
```

Reading Study Data Modeling

```
reading_data <- read_csv("C:/Users/ghimiren/Desktop/Reading Habits Study Time Diary Survey for M
oroccan Teachers/reading_data.csv")
# str(reading_data)
```

```
# Changing the variable class
reading_data <- reading_data %>%
  # convert data types
  mutate(
    duration_in_seconds = as.numeric(duration_in_seconds),
    across(q4:q25, as.factor)
  )
# summary(reading_data)
```

Looking at the summary, the variables having `_text` at the end is not useful. Getting rid of them:

```
new_data <- reading_data %>%
  dplyr::select(-ends_with("_text"))
str(new_data)
```

```
tibble [701 × 63] (S3: tbl_df/tbl/data.frame)
 $ duration_in_seconds: num [1:701] 10352 1010 947 821 1342 ...
 $ ip_address         : chr [1:701] "105.154.38.39" "196.118.24.254" "196.118.24.254" "105.69.5
4.235" ...
 $ q4                 : Factor w/ 2 levels "Female","Male": 1 2 1 2 2 2 2 2 2 1 ...
 $ q6                 : Factor w/ 6 levels "0 teaching experience",..: 3 4 4 5 6 6 2 2 2 2 ...
 $ q7                 : Factor w/ 4 levels "Click to write Choice 4",..: 3 4 4 4 4 3 4 2 4 2 ...
 $ q8                 : Factor w/ 3 levels "I am a pre-service teacher (i.e., I am currently in
training at Centre régional des métiers de l’éducation et d"| __truncated__,..: 2 1 2 2 2 2 3 2
1 2 ...
 $ q12_1              : Factor w/ 2 levels "I have access to",..: 2 1 2 2 2 2 2 2 1 NA ...
 $ q12_2              : Factor w/ 2 levels "I have access to",..: 2 NA 2 NA NA 2 2 2 2 2 ...
 $ q12_3              : Factor w/ 2 levels "I have access to",..: NA NA NA 2 NA 1 1 NA 2 NA ...
 $ q12_4              : Factor w/ 2 levels "I have access to",..: NA 2 2 NA 2 2 2 2 2 2 ...
 $ q12_5              : Factor w/ 2 levels "I have access to",..: NA NA NA NA NA 1 1 NA 2 NA ...
 $ q12_6              : Factor w/ 2 levels "I have access to",..: NA NA NA NA 2 1 2 NA 2 NA ...
 $ q12_7              : Factor w/ 2 levels "I have access to",..: NA NA NA 2 NA 2 2 NA 2 NA ...
 $ q12_8              : Factor w/ 2 levels "I have access to",..: NA NA 2 NA NA 1 1 NA 2 NA ...
 $ q12_9              : Factor w/ 2 levels "I have access to",..: NA NA NA NA NA 1 1 NA 2 NA ...
 $ q12_10             : Factor w/ 2 levels "I have access to",..: NA NA NA NA 1 NA NA NA 2 NA
...
 $ q13_1              : Factor w/ 5 levels "1 Hour","2 Hours",..: 5 5 5 5 5 4 1 3 5 3 ...
 $ q13_2              : Factor w/ 5 levels "1 Hour","2 Hours",..: 5 4 5 4 5 4 5 4 5 5 ...
 $ q13_3              : Factor w/ 5 levels "1 Hour","2 Hours",..: 4 4 5 1 4 1 5 2 5 4 ...
 $ q13_4              : Factor w/ 5 levels "1 Hour","2 Hours",..: 2 2 1 3 3 3 3 3 2 2 ...
 $ q13_5              : Factor w/ 5 levels "1 Hour","2 Hours",..: 5 1 1 2 2 4 1 1 1 2 ...
 $ q13_6              : Factor w/ 5 levels "1 Hour","2 Hours",..: 3 1 3 4 1 1 5 5 5 4 ...
 $ q13_7              : Factor w/ 5 levels "1 Hour","2 Hours",..: 1 1 5 1 4 5 1 4 2 4 ...
 $ q13_8              : Factor w/ 5 levels "1 Hour","2 Hours",..: 5 4 5 5 5 5 1 5 1 5 ...
 $ q13_9              : Factor w/ 5 levels "1 Hour","2 Hours",..: 5 5 5 5 5 5 5 5 4 5 ...
 $ q13_10             : Factor w/ 5 levels "1 Hour","2 Hours",..: 4 2 5 4 1 5 5 3 2 5 ...
 $ q13_11             : Factor w/ 5 levels "1 Hour","2 Hours",..: NA NA NA NA NA 1 NA NA 5 5 ...
 $ q14                : Factor w/ 65 levels "Books online",..: 53 12 61 5 47 27 60 61 1 61 ...
 $ q15                : Factor w/ 4 levels "6:00 a.m.-11:59 a.m.",..: 2 3 3 2 3 4 1 2 3 2 ...
 $ q16                : Factor w/ 9 levels "0 minutes","1 hour",..: 2 2 7 2 3 4 7 7 8 4 ...
 $ q17_1              : Factor w/ 4 levels "Alittle of the Time",..: 1 1 3 1 1 3 4 4 4 4 ...
 $ q17_2              : Factor w/ 4 levels "Alittle of the Time",..: 3 3 1 4 3 4 1 NA 2 1 ...
 $ q17_3              : Factor w/ 4 levels "Alittle of the Time",..: 1 4 4 2 1 3 2 3 2 4 ...
 $ q17_4              : Factor w/ 4 levels "Alittle of the Time",..: 3 1 4 1 3 3 4 3 2 4 ...
 $ q17_5              : Factor w/ 4 levels "Alittle of the Time",..: 2 1 1 1 2 1 4 2 2 2 ...
 $ q17_6              : Factor w/ 4 levels "Alittle of the Time",..: 3 3 3 3 3 3 3 3 2 3 ...
 $ q17_7              : Factor w/ 4 levels "Alittle of the Time",..: 4 1 1 NA 4 2 2 2 2 2 ...
 $ q17_8              : Factor w/ 4 levels "Alittle of the Time",..: NA NA NA NA NA 3 NA NA 2 3
...
 $ q18_1              : Factor w/ 3 levels "No","Unsure",..: 3 1 3 3 3 1 3 1 3 3 ...
 $ q18_2              : Factor w/ 3 levels "No","Unsure",..: 1 2 2 1 2 1 3 1 3 2 ...
 $ q18_3              : Factor w/ 3 levels "No","Unsure",..: 3 1 3 NA 3 1 3 1 3 2 ...
 $ q18_4              : Factor w/ 3 levels "No","Unsure",..: 3 1 2 NA 2 3 1 1 3 1 ...
 $ q18_5              : Factor w/ 3 levels "No","Unsure",..: 1 1 2 3 NA 3 1 1 3 1 ...
 $ q18_6              : Factor w/ 3 levels "No","Unsure",..: 2 1 2 NA 3 1 1 1 3 1 ...
 $ q18_7              : Factor w/ 3 levels "No","Unsure",..: NA NA NA NA NA 1 NA NA 3 2 ...
 $ q19                : Factor w/ 5 levels "Did not multi-task",..: 4 4 4 5 2 5 4 4 2 4 ...
```

```
 $ q20               : Factor w/ 114 levels "Journal articles-In print",..: 68 32 109 1 68 68 3
57 68 67 ...
 $ q21               : Factor w/ 11 levels "6:00 a.m.-11:59 a.m.",..: 9 2 6 6 6 6 1 6 10 6 ...
 $ q22               : Factor w/ 8 levels "0 minutes","1 hour",..: 4 2 2 2 7 4 8 7 8 7 ...
 $ q23_1             : Factor w/ 4 levels "A little of the time",..: 1 4 3 1 2 3 1 3 2 1 ...
 $ q23_2             : Factor w/ 4 levels "A little of the time",..: 3 3 1 4 3 4 1 3 2 1 ...
 $ q23_3             : Factor w/ 4 levels "A little of the time",..: 3 1 2 NA 3 3 4 4 2 3 ...
 $ q23_4             : Factor w/ 4 levels "A little of the time",..: 1 4 1 NA 2 1 4 3 2 4 ...
 $ q23_5             : Factor w/ 4 levels "A little of the time",..: 3 3 3 NA 3 3 3 3 2 3 ...
 $ q23_6             : Factor w/ 4 levels "A little of the time",..: 1 1 1 NA 2 4 2 3 2 1 ...
 $ q23_7             : Factor w/ 4 levels "A little of the time",..: NA NA NA NA NA 3 3 NA 2 3
...
 $ q24_1             : Factor w/ 3 levels "No","Unsure",..: 1 2 3 2 1 1 3 1 3 1 ...
 $ q24_2             : Factor w/ 3 levels "No","Unsure",..: 1 1 3 1 3 1 1 1 3 1 ...
 $ q24_3             : Factor w/ 3 levels "No","Unsure",..: 1 2 3 NA 3 1 3 1 1 1 ...
 $ q24_4             : Factor w/ 3 levels "No","Unsure",..: 1 2 3 1 3 1 3 1 3 3 ...
 $ q24_5             : Factor w/ 3 levels "No","Unsure",..: 1 2 3 NA 3 1 1 1 3 2 ...
 $ q24_6             : Factor w/ 3 levels "No","Unsure",..: NA NA NA NA NA 1 NA NA NA 2 ...
 $ q25               : Factor w/ 5 levels "Did not multi-task",..: 5 4 5 5 5 5 4 4 4 4 ...
```

Variables Used for Descriptive Analysis

Question 13, Thinking only about yesterday, how much time did you spend doing the following:(Select one answer next to each of the following activities)

	None	30 Minutes	1 Hour	2 Hours	3 Hours or More
Hanging out with friends					
Being physically active or exercising (e.g., playing sports, working out, running, etc.)					
Participating in other activities (e.g., clubs, music, art, or hobbies)					
Using the Internet					
Reading for school/doing homework					
Doing chores					
Reading for fun or pleasure					
Writing					
Playing video games					
Watching TV					
Other (Please Specify)					

Q-13

- The information in Question 13 doesn't seem to have direct use in the analysis. We can use it in descriptive analysis.

Background Variables Used in the Model

Question 4, What is your gender?

a) Male
b) Female

Question8, Tell us about your education and work experience.

a) I am a pre-service teacher (i.e., I am currently in training at Centre régional des métiers de l'éducation et de la formation (CRMEF) or Ecole Normale Supérieure (ENS)
b) I am an-service teacher (i.e., I am currently working full-time as a teacher)
c) Other (please specify:_____)

- This variable will be modified and used in the model. We combined teachers based on their years of experience. The new categories are **a) 0-5 Years, b) 6-10 Years, c) more than 10 Years**

Question 6, How long have you been teaching English?

a) 0 teaching experience
b) 1-5 years
c) 6-10 years
d) 11-15 years
e) 16 – 20
f) 20+

Q-6

- This variable got dichotomized and the new categories were **a) Rural, b) Urban (Suburban + Urban)**

Question 7, Current school location

a) Urban
b) Rural
c) Suburban
d) Click to write Choice 4

Q-7

Variables Needed in Reading for Fun Model

Question 14, Which of the following materials did you read for fun yesterday? (Select as many answers as needed)

a) News stories online
b) Books online
c) Magazines online
d) Social media threads (e.g., Face book, Instagram, Twitter, etc.)
e) Newspapers-In print
f) Magazines-In print
g) Other materials in print (please specify)

Question 15, During which parts of the day yesterday did you read for fun? (Select the time block or blocks during which you read for fun yesterday)

a) 6:00 a.m.-11:59 a.m.
b) Noon-6:00 p.m.
c) 6:00 p.m.-11:59 p.m.
d) Midnight-5:59 a.m.

Question 16, Thinking about yesterday, about how much time did you spend reading for fun? Select the choice that best describes the amount of time you spent reading for pleasure during the time block(s) you selected above.

a) 0 minutes
b) 15 minutes
c) 30 minutes
d) 45 minutes
e) 1 hour
f) 1.5 hours
g) 2 hours
h) 3 hours or more
i) Click to write Choice 9

Question 17, While you were reading books, magazines, or newspapers for fun, how often did you do any of the following activities or tasks at the same time:

	Most of the Time	Some of the Time	A little of the Time	Never
Watching TV, videos, or DVDs				
Listening to Music				
Reading in support of teaching or professional development				
Writing				
Talking on the Phone or Texting				
Playing games online				
Social Networking, Email, or other Internet-based communication				
Other (please specify)				

Question 19, In general, did you feel that doing multiple tasks at the same time affected your ability to concentrate on the materials you were reading for fun?

a) Yes, a lot
b) Yes, some
c) No, not a all
d) Not sure
e) Did not multi-task

Variables Needed in Reading for Academic Purposes Model

Question 20, Which of the following materials did you read yesterday in relation to your classes or teaching? (Select as many answers as needed)

a) Textbook Chapters-Online
b) Journal articles-online
c) Reports-online
d) Novels-Online
e) Textbook Chapters- In print
f) Reports-In print
g) Novels in Print
h) Other Materials-Please specify

Question 21, During which parts of the day yesterday did you read in support of your teaching or professional development? (Select the time block or blocks during which you read in support of your teaching or professional development)

a. 6:00 a.m.-11:59 a.m.
b. Noon-6:00 p.m.
c. 6:00 p.m. – 11:59 p.m.
d. Midnight- 5:59 a.m.

Question 22, Thinking about yesterday, about how much time did you spend reading for your classes or for your teaching? Select the category that best describes your answer.

a) 0 minutes
b) 15 minutes
c) 30 minutes
d) 45 minutes
e) 1 hour
f) 1.5 hours
g) 2 hours
h) 3 hours or more

Question 23, While you were reading for your classes or for your teaching yesterday, how often did you do any of the following activities at the same time?

	Most of the time	Some of the time	A little of the time	Never
Watch TV, videos or DVDs				
Listening to Music				
Writing				
Talking on the phone or texting				
Playing video games				
Social Networking, Email, or Other Internet-based communication				
Other (Please specify)				

Question 25, In general, did you feel that doing multiple tasks at the same time affected your ability to concentrate on the materials you were reading for your classes or for your teaching?

a) Yes, a lot
b) Yes, some
c) No, not at all
d) Not sure
e) Did not multi-task

Variable Summary

```
read_data <- reading_data %>%
  dplyr::select(
    duration_in_seconds, ip_address,
    q4, q6, q7, q8, q14, q15, q16,
    q17_1, q17_2, q17_3, q17_4, q17_5, q17_6, q17_7,
    q19, q20, q21, q22,
    q23_1, q23_2, q23_3, q23_4, q23_5, q23_6, q23_7, q25
  )
summary(read_data[, -(1:2)])
```

```
     q4                          q6                              q7
 Female:182   0 teaching experience: 19   Click to write Choice 4:  7
 Male  :430   1-5 years            :295   Rural                   :281
 NA's  : 89   11-15 years          : 91   Suburban                :144
              16 - 20              : 47   Urban                   :269
              20+                  : 20
              6-10 years           :229


q8
 I am a pre-service teacher (i.e., I am currently in training at Centre régional des métiers de
l’éducation et de la formation (CRMEF) or Ecole Normale Supérieure (ENS):362
 I am an-service teacher (i.e., I am currently working full-time as a teacher)
:318
 Other (please specify: _____)
: 21




                                                                                q14
 Magazines online                                                                 : 69
 Social media threads (e.g., Face book, Instagram, Twitter, etc.)                 : 68
 Books online                                                                     : 49
 News stories online                                                              : 42
 Books online,Social media threads (e.g., Face book, Instagram, Twitter, etc.): 40
 (Other)                                                                          :432
 NA's                                                                             :  1
                  q15               q16                       q17_1
 6:00 a.m.-11:59 a.m.:165   30  minutes:184   Alittle of the Time:124
 6:00 p.m.-11:59 p.m.:226   45  minutes:156   Most of the Time   :132
 Midnight-5:59 a.m.  : 35   1 hour     :155   Never              : 64
 Noon-6:00 p.m.      :274   15  minutes: 75   Some of the Time   :370
 NA's                :  1   1.5 hours  : 53   NA's               : 11
                            (Other)    : 77
                            NA's       :  1
                q17_2                     q17_3
 Alittle of the Time:236   Alittle of the Time:182
 Most of the Time   :103   Most of the Time   :119
 Never              : 36   Never              : 30
 Some of the Time   :314   Some of the Time   :360
 NA's               : 12   NA's               : 10


                q17_4                     q17_5
 Alittle of the Time:235   Alittle of the Time:239
 Most of the Time   : 89   Most of the Time   : 87
 Never              : 67   Never              : 49
 Some of the Time   :298   Some of the Time   :319
 NA's               : 12   NA's               :  7
```

```
               q17_6                    q17_7
Alittle of the Time:209   Alittle of the Time:224
Most of the Time   : 60   Most of the Time   :113
Never              :118   Never              : 44
Some of the Time   :296   Some of the Time   :311
NA's               : 18   NA's               :  9


               q19                              q20
Did not multi-task: 11   Textbook Chapters-Online  : 55
No, not at all    :197   Novels-Online             : 50
Not sure          : 48   Reports-Online            : 46
Yes, a lot        :118   Textbook Chapters-In print: 38
Yes, some         :317   Journal articles-Online   : 31
NA's              : 10   (Other)                   :480
                         NA's                      :  1
                                 q21              q22
Noon-6:00 p.m.                     :237   30 minutes:169
6:00 p.m.-11:59 p.m.               :189   1 hour    :162
6:00 a.m.-11:59 a.m.               :142   45 minutes:157
Noon-6:00 p.m.,6:00 p.m.-11:59 p.m.: 44   15 minutes: 88
6:00 a.m.-11:59 a.m.,Noon-6:00 p.m.: 26   1.5 hours : 66
(Other)                            : 62   (Other)   : 58
NA's                               :  1   NA's      :  1
                q23_1                    q23_2
A little of the time:119   A little of the time:230
Most of the time    :154   Most of the time    :120
Never               : 77   Never               : 53
Some of the time    :342   Some of the time    :286
NA's                :  9   NA's                : 12


                q23_3                    q23_4
A little of the time:226   A little of the time:224
Most of the time    :103   Most of the time    : 95
Never               : 63   Never               : 68
Some of the time    :300   Some of the time    :300
NA's                :  9   NA's                : 14


                q23_5                    q23_6
A little of the time:208   A little of the time:255
Most of the time    : 81   Most of the time    :103
Never               :125   Never               : 44
Some of the time    :273   Some of the time    :286
NA's                : 14   NA's                : 13


                q23_7                    q25
A little of the time:149   Did not multi-task:  9
Most of the time    : 70   No, not at all    :185
Never               : 64   Not sure          : 52
```

```
 Some of the time    :188   Yes, a lot        :136
 NA's                :230   Yes, some         :318
                            NA's              :  1
```

Preparing Variables to Use in the Models

- **Question 4 [q4 - *gender*]**: Response Codes Female, Male
- **Question 6 [q6 - *experience*]**: Response Codes 0-5 years = 0, 6-10 years = 1, 11-years and more = 2
- **Question 7 [q7 - *sch_type*]**: Response Codes Rural = 0, Urban/Suburban = 1
- **Question 8 [q8 - *tchr_type*]**: Response Codes pre-service = 0, inservice = 1
- **Question 15 [q15 - *rf_time*]**: Response Codes
- **Question 16 [q16 - *rf_length*]**: Response Codes 0 minutes = 0, 15 minutes = 1, 30 minutes = 2, 45 minutes = 3, 1 hour = 4, 1.5 hours = 5, 2 hours = 6, 3 hours or more = 7
- **Question 17 [q17_1 - *rf_tv*, q17_2 - *rf_music*, q17_3 - *rf_pd*, q17_4 - *rf_write*, q17_5 - *rf_talk_phone*, q17_6 - *rf_onl_game*, q17_7 - *rf_soc_network*]**: Response Codes Never = 0, Alittle of the Time = 1, Some of the Time = 2, Most of the Time = 3
- **Question 19 [q19 - *rf_disp*]**: Response Codes No, not at all = 0, Yes, some = 1, Yes, a lot = 2
- **Question 20 [q20 - *ra_text*]**: Response Codes Textbook Chapters-Online, Journal articles-Online, Reports-Online, Novels-Online, Textbook Chapters-In print, Reports-In print, Novels-In print, Other materials-Please specify
- **Question 21 [q21 - *ra_time*]**: Response Codes 6:00 a.m.-11:59 a.m., Noon-6:00 p.m., 6:00 p.m.-11:59 p.m., Midnight-5:59 a.m.
- **Question 22 [q22 - *ra_length*]**: Response Codes 0 minutes = 0, 15 minutes = 1, 30 minutes = 2, 45 minutes = 3, 1 hour = 4, 1.5 hours = 5, 2 hours = 6, 3 hours or more = 7
- **Question 23 [q23_1 - *ra_tv*, q23_2 - *ra_music*, q23_3 - *ra_write*, q23_4 - *ra_talk_phone*, q23_5 - *ra_video_game*, q23_6 - *ra_soc_network*, q23_7 - *ra_other*]**: Response Codes Never = 0, A little of the time = 1, Some of the time = 2, Most of the time = 3
- **Question 25 [q25 - *ra_disp*]**: Response Codes No, not at all = 0, Yes, some = 1, Yes, a lot = 2

```
# Question 6
read_data$q6 <- factor(
  ifelse(read_data$q6 %in% c("0 teaching experience", "1-5 years"), 0,
    ifelse(read_data$q6 %in% c("11-15 years", "16 - 20", "20+"), 2,
      ifelse(read_data$q6 == "6-10 years", 1, read_data$q6)
    )
  ),
  levels = c(0, 1, 2),
  labels = c("0-5 years", "6-10 years", "11-years and more")
)
# summary(read_data$q6)
```

```
# Question 7
read_data$q7 <- factor(
  ifelse(read_data$q7 == "Click to write Choice 4", NA,
    ifelse(read_data$q7 %in% c("Urban", "Suburban"), 1,
      ifelse(read_data$q7 == "Rural", 0, read_data$q7)
    )
  ),
  levels = c(0, 1),
  labels = c("Rural", "Urban/Suburban")
)
# summary(read_data$q7)
```

```
# Question 8
read_data$q8 <- factor(ifelse(
  read_data$q8 == "I am a pre-service teacher (i.e., I am currently in training at Centre région
al des métiers de l’éducation et de la formation (CRMEF) or Ecole Normale Supérieure (ENS)",
  0,
  ifelse(
    read_data$q8 == "I am an-service teacher (i.e., I am currently working full-time as a teache
r)",
    1,
    NA
  )
), levels = c(0, 1), labels = c("pre-service", "inservice"))
# summary(read_data$q8)
```

```
# Question 14
# Create a function to extract the first selection
read_data$q14 <- as.character(read_data$q14)
extract_first_selection <- function(text) {
  selections <- strsplit(text, ",")[[1]]
  first_selection <- trimws(selections[1])
  return(first_selection)
}
# Apply the function to create a new variable with the first selection
read_data$q14_first_selection <- sapply(read_data$q14, extract_first_selection)
# Recode the first selection into the desired categories
read_data$q14_recode <- factor(
  read_data$q14_first_selection,
  levels = c("News stories online", "Social media threads (e.g., Face book, Instagram, Twitter,
etc.)", "Books online", "Magazines online", "Newspapers-In print", "Magazines-In print", "Other
Materials in print (please specify)"),
  labels = c("News stories online", "Social media threads", "Books online", "Magazines online",
"Newspapers-In print", "Magazines-In print", "Other Materials in print (please specify)")
)
# Remove the intermediate variable
read_data$q14_first_selection <- NULL
# Change the class of the Variable
read_data$q14_recode <- as.factor(read_data$q14_recode)
# summary(read_data$q14_recode)
```

```r
# Question 16
# Adjusting the data with case_when for more control
read_data <- read_data %>%
  mutate(
    q16 = str_trim(q16), # Trim whitespace first
    q16 = case_when(
      str_detect(q16, "^0 minutes$|^15[ ]*minutes$") ~ "less than 30 minutes",
      str_detect(q16, "^30[ ]*minutes$|^45[ ]*minutes$") ~ "30-59 minutes",
      str_detect(q16, "^1 hour$") ~ "60-89 minutes",
      str_detect(q16, "^1\\.5 hours$") ~ "90-119 minutes",
      str_detect(q16, "^2 hours$|^3 hours or more$") ~ "2 hours or more",
      TRUE ~ "Invalid" # Handles unmatched cases
    )
  ) %>%
  mutate(q16 = factor(q16, levels = c(
    "less than 30 minutes",
    "30-59 minutes",
    "60-89 minutes",
    "90-119 minutes",
    "2 hours or more"
  )))

# Now check the levels
print(levels(read_data$q16))
```

```
[1] "less than 30 minutes" "30-59 minutes"        "60-89 minutes"
[4] "90-119 minutes"       "2 hours or more"
```

```r
# Re-coding Responses in Question 17(q17_2, q17_3, q17_4, q17_5, q17_6, q17_7)
# Function - using dplyr::recode_factor instead of car::recode
recodelikert <- function(data, variables, categories) {
  # Need to load dplyr
  library(dplyr)

  for (variable in variables) {
    # Using dplyr's recode_factor
    data[[variable]] <- dplyr::recode_factor(data[[variable]], !!!categories)
  }
  return(data)
}

# Defining Categories
categories <- list(
  "Never" = 0,
  "Alittle of the Time" = 1,
  "Some of the Time" = 2,
  "Most of the Time" = 3
)
read_data <- recodelikert(read_data, c("q17_1", "q17_2", "q17_3", "q17_4", "q17_5", "q17_6", "q1
7_7"), categories)

# Questions 23 (q23_1, q23_2, q23_3, q23_4, q23_5, q23_6, q23_7)
# Defining Categories
categories_1 <- list(
  "Never" = 0,
  "A little of the time" = 1,
  "Some of the time" = 2,
  "Most of the time" = 3
)
read_data <- recodelikert(read_data, c("q23_1", "q23_2", "q23_3", "q23_4", "q23_5", "q23_6", "q2
3_7"), categories_1)


# Question 19, and saving as q19_recode
read_data$q19_recode <- factor(
  ifelse(read_data$q19 %in% c("Did not multi-task", "Not sure"), NA,
    ifelse(read_data$q19 == "No, not at all", 0,
      ifelse(read_data$q19 == "Yes, some", 1,
        ifelse(read_data$q19 == "Yes, a lot", 2, read_data$q19)
      )
    )
  ),
  levels = c(0, 1, 2),
  labels = c("No, not at all", "Yes, some", "Yes, a lot"),
  exclude = NULL
)

# Question 20
read_data$q20 <- as.character(read_data$q20)
# Apply the function to create a new variable with the first selection
```

```
read_data$q20_first_selection <- sapply(read_data$q20, extract_first_selection)
# Recode the first selection into the desired categories
read_data$q20_recode <- factor(
  read_data$q20_first_selection,
  levels = c("Textbook Chapters-Online", "Journal articles-Online", "Reports-Online", "Novels-On
line
", "Textbook Chapters-In print", "Reports-In print", "Novels-In print", "Other materials-Please
specify"),
  labels = c("Textbook Chapters-Online", "Journal articles-Online", "Reports-Online", "Novels-On
line", "Textbook Chapters-In print", "Reports-In print", "Novels-In print", "Other materials-Ple
ase specify")
)
# Remove the intermediate variable
read_data$q20_first_selection <- NULL

# Question 21
read_data$q21 <- as.character(read_data$q21)
# Apply the function to create a new variable with the first selection
read_data$q21_first_selection <- sapply(read_data$q21, extract_first_selection)
# Recode the first selection into the desired categories
read_data$q21_recode <- factor(
  read_data$q21_first_selection,
  levels = c("6:00 a.m.-11:59 a.m.", "Noon-6:00 p.m.", "6:00 p.m.-11:59 p.m.", "Midnight-5:59 a.
m."),
  labels = c("6:00 a.m.-11:59 a.m.", "Noon-6:00 p.m.", "6:00 p.m.-11:59 p.m.", "Midnight-5:59 a.
m.")
)
# Remove the intermediate variable
read_data$q21_first_selection <- NULL

# Question 22
read_data$q22 <- factor(dplyr::recode(read_data$q22,
  "0 minutes" = "less than 30 minutes",
  "15 minutes" = "less than 30 minutes",
  "30 minutes" = "30-59 minutes",
  "45 minutes" = "30-59 minutes",
  "1 hour" = "60-89 minutes",
  "1.5 hours" = "90-119 minutes",
  "2 hours" = "2 hours or more",
  "3 hours or more" = "2 hours or more"
))
# Convert to factor with the specified levels
levels(read_data$q22) <- c("less than 30 minutes", "30-59 minutes", "60-89 minutes", "90-119 min
utes", "2 hours or more")

# Question 25 and Saving as q25_recode
read_data$q25_recode <- factor(
  ifelse(read_data$q25 %in% c("Did not multi-task", "Not sure"), NA,
    ifelse(read_data$q25 == "No, not at all", 0,
      ifelse(read_data$q25 == "Yes, some", 1,
        ifelse(read_data$q25 == "Yes, a lot", 2, read_data$q25)
      )
```

```
    )
  ),
  levels = c(0, 1, 2),
  labels = c("No, not at all", "Yes, some", "Yes, a lot"),
  exclude = NULL
)

# summary(read_data[,-2])
# str(read_data)
```

```
read_data <- read_data %>%
  rename(
    gender = q4,
    experience = q6,
    sch_type = q7,
    tchr_type = q8,
    rf_text = q14_recode,
    rf_time = q15,
    rf_length = q16,
    rf_tv = q17_1,
    rf_music = q17_2,
    rf_pd = q17_3,
    rf_write = q17_4,
    rf_talk_phone = q17_5,
    rf_onl_game = q17_6,
    rf_soc_network = q17_7,
    rf_disp = q19_recode,
    ra_text = q20_recode,
    ra_time = q21_recode,
    ra_length = q22,
    ra_tv = q23_1,
    ra_music = q23_2,
    ra_write = q23_3,
    ra_talk_phone = q23_4,
    ra_video_game = q23_5,
    ra_soc_network = q23_6,
    ra_other = q23_7,
    ra_disp = q25_recode
  )
summary(read_data[, -2])
```

```
duration_in_seconds    gender                experience
Min.   :    74      Female:182   0-5 years        :314
1st Qu.:   418      Male  :430   6-10 years       :229
Median :   810      NA's  : 89   11-years and more:158
Mean   :  2506
3rd Qu.:  1422
Max.   :210923

          sch_type         tchr_type       q14
Rural         :281   pre-service:362   Length:701
Urban/Suburban:413   inservice  :318   Class :character
NA's          :  7   NA's       : 21   Mode  :character




                rf_time                    rf_length    rf_tv
6:00 a.m.-11:59 a.m.:165   less than 30 minutes: 94   0   : 64
6:00 p.m.-11:59 p.m.:226   30-59 minutes       :340   1   :124
Midnight-5:59 a.m.  : 35   60-89 minutes       :155   2   :370
Noon-6:00 p.m.      :274   90-119 minutes      : 53   3   :132
NA's                :  1   2 hours or more     : 57   NA's: 11
                           NA's                :  2

rf_music    rf_pd     rf_write   rf_talk_phone rf_onl_game rf_soc_network
0   : 36   0   : 30   0   : 67   0   : 49      0   :118    0   : 44
1   :236   1   :182   1   :235   1   :239      1   :209    1   :224
2   :314   2   :360   2   :298   2   :319      2   :296    2   :311
3   :103   3   :119   3   : 89   3   : 87      3   : 60    3   :113
NA's: 12   NA's: 10   NA's: 12   NA's:  7      NA's: 18    NA's:  9


                q19          q20                q21
Did not multi-task: 11   Length:701         Length:701
No, not at all    :197   Class :character   Class :character
Not sure          : 48   Mode  :character   Mode  :character
Yes, a lot        :118
Yes, some         :317
NA's              : 10

               ra_length    ra_tv     ra_music   ra_write   ra_talk_phone
less than 30 minutes:105   0   : 77   0   : 53   0   : 63   0   : 68
30-59 minutes       :162   1   :119   1   :230   1   :226   1   :224
60-89 minutes       : 66   2   :342   2   :286   2   :300   2   :300
90-119 minutes      : 41   3   :154   3   :120   3   :103   3   : 95
2 hours or more     :326   NA's:  9   NA's: 12   NA's:  9   NA's: 14
NA's                :  1

ra_video_game ra_soc_network ra_other                   q25
0   :125      0   : 44       0   : 64   Did not multi-task:  9
1   :208      1   :255       1   :149   No, not at all    :185
2   :273      2   :286       2   :188   Not sure          : 52
```

```
 3   : 81      3   :103       3   : 70   Yes, a lot        :136
 NA's: 14      NA's: 13       NA's:230   Yes, some         :318
                                         NA's              :  1

                rf_text              rf_disp
 News stories online:232   No, not at all:197
 Books online       :216   Yes, some     :317
 Magazines online   :143   Yes, a lot    :118
 Newspapers-In print: 15   NA's          : 69
 Magazines-In print :  6
 (Other)            :  0
 NA's               : 89
                          ra_text                    ra_time
 Textbook Chapters-Online      :201   6:00 a.m.-11:59 a.m.:201
 Journal articles-Online       :182   Noon-6:00 p.m.      :286
 Reports-Online                :143   6:00 p.m.-11:59 p.m.:194
 Textbook Chapters-In print    : 49   Midnight-5:59 a.m.  : 19
 Other materials-Please specify: 11   NA's                :  1
 (Other)                       :  9
 NA's                          :106
           ra_disp
 No, not at all:185
 Yes, some     :318
 Yes, a lot    :136
 NA's          : 62
```

```
# str(read_data)
```

1. Investigating the Relationship Between Displacement and Demographic Factors:

```
# Cross-tabulation of rf_disp and gender
rf_disp_gender_xtab <- xtabs(~ rf_disp + gender, data = read_data)
# Calculate percentage values with two decimal places
rf_disp_gender_xtab_percentage <- round(prop.table(rf_disp_gender_xtab, margin = 2) * 100, 2)
rf_disp_gender_xtab_percentage
```

```
                gender
rf_disp          Female  Male
  No, not at all  38.04 29.95
  Yes, some       45.40 50.25
  Yes, a lot      16.56 19.80
```

```
# Cross-tabulation of ra_disp and gender
ra_disp_gender_xtab <- xtabs(~ ra_disp + gender, data = read_data)
# Calculate percentage values with two decimal places
ra_disp_gender_xtab_percentage <- round(prop.table(ra_disp_gender_xtab, margin = 2) * 100, 2)
ra_disp_gender_xtab_percentage
```

```
               gender
ra_disp          Female  Male
  No, not at all  28.21 30.37
  Yes, some       53.85 45.93
  Yes, a lot      17.95 23.70
```

```
# Cross-tabulation of rf_disp and experience
rf_disp_experience_xtab <- xtabs(~ rf_disp + experience, data = read_data)
# Calculate percentage values with two decimal digits
rf_disp_experience_xtab_percentage <- round(prop.table(rf_disp_experience_xtab, margin = 2) * 10
0, 2)
rf_disp_experience_xtab_percentage
```

```
               experience
rf_disp          0-5 years 6-10 years 11-years and more
  No, not at all     27.43      34.30             34.31
  Yes, some          50.69      51.21             47.45
  Yes, a lot         21.88      14.49             18.25
```

```
# Cross-tabulation of rf_disp and experience
ra_disp_experience_xtab <- xtabs(~ ra_disp + experience, data = read_data)
# Calculate percentage values with two decimal places
ra_disp_experience_xtab_percentage <- round(prop.table(ra_disp_experience_xtab, margin = 2) * 10
0, 2)
ra_disp_experience_xtab_percentage
```

```
               experience
ra_disp          0-5 years 6-10 years 11-years and more
  No, not at all     24.48      35.07             28.87
  Yes, some          49.30      47.87             53.52
  Yes, a lot         26.22      17.06             17.61
```

```
# Cross-tabulation of rf_disp and school type
rf_disp_sch_type_xtab <- xtabs(~ rf_disp + sch_type, data = read_data)
# Calculate percentage values with two decimal places
rf_disp_sch_type_xtab_percentage <- round(prop.table(rf_disp_sch_type_xtab, margin = 2) * 100,
2)
rf_disp_sch_type_xtab_percentage
```

```
               sch_type
rf_disp          Rural Urban/Suburban
  No, not at all 39.37          25.00
  Yes, some      49.21          51.34
  Yes, a lot     11.42          23.66
```

```
# Cross-tabulation of ra_disp and school type
ra_disp_sch_type_xtab <- xtabs(~ ra_disp + sch_type, data = read_data)
# Calculate percentage values with two decimal places
ra_disp_sch_type_xtab_percentage <- round(prop.table(ra_disp_sch_type_xtab, margin = 2) * 100,
2)
ra_disp_sch_type_xtab_percentage
```

```
               sch_type
ra_disp          Rural Urban/Suburban
  No, not at all 30.92          26.75
  Yes, some      53.41          48.05
  Yes, a lot     15.66          25.19
```

```
# Cross-tabulation of rf_disp and teacher type
rf_disp_tchr_type_xtab <- xtabs(~ rf_disp + tchr_type, data = read_data)
rf_disp_tchr_type_xtab_percentage <- round(prop.table(rf_disp_tchr_type_xtab, margin = 2) * 100,
2)
rf_disp_tchr_type_xtab_percentage
```

```
               tchr_type
rf_disp          pre-service inservice
  No, not at all       29.97     31.34
  Yes, some            50.15     51.06
  Yes, a lot           19.88     17.61
```

```
# Cross-tabulation of rf_disp and teacher type
ra_disp_tchr_type_xtab <- xtabs(~ ra_disp + tchr_type, data = read_data)
# ra_disp_tchr_type_xtab
# Calculate percentage values with two decimal places
ra_disp_tchr_type_xtab_percentage <- round(prop.table(ra_disp_tchr_type_xtab, margin = 2) * 100,
2)
ra_disp_tchr_type_xtab_percentage
```

```
               tchr_type
ra_disp          pre-service inservice
  No, not at all       27.98     29.02
  Yes, some            50.60     49.30
  Yes, a lot           21.43     21.68
```

```r
# Filter out NAs from the dataset
filtered_data <- read_data[complete.cases(read_data), ]

# Create the grouped bar plots
gender_rf_disp <- ggplot(filtered_data, aes(x = rf_disp, fill = gender)) +
  geom_bar(position = "dodge", color = "black") +
  labs(x = "Reading for Fun Displacement", y = "Frequency") +
  scale_y_continuous(breaks = NULL) + # Remove y-axis ticks
  theme_minimal() +
  theme(
    panel.grid.major = element_blank(),
    panel.grid.minor = element_blank(),
    axis.line = element_line(colour = "black"),
    axis.ticks.y = element_blank(), # Remove y-axis ticks
    legend.position = "bottom",
    legend.title = element_blank(),
    legend.text = element_text(size = 8),
    axis.text = element_text(size = 8),
    axis.title = element_text(size = 10),
    plot.title = element_text(size = 12, face = "bold")
  ) +
  scale_fill_manual(
    values = c("#0072B2", "#E69F00"),
    labels = c("Female", "Male")
  )

gender_ra_disp <- ggplot(filtered_data, aes(x = ra_disp, fill = gender)) +
  geom_bar(position = "dodge", color = "black") +
  labs(x = "Reading for Academic Purposes", y = "") +
  scale_y_continuous(breaks = NULL) + # Remove y-axis ticks
  theme_minimal() +
  theme(
    panel.grid.major = element_blank(),
    panel.grid.minor = element_blank(),
    axis.line = element_line(colour = "black"),
    axis.ticks.y = element_blank(), # Remove y-axis ticks
    legend.position = "bottom",
    legend.title = element_blank(),
    legend.text = element_text(size = 8),
    axis.text = element_text(size = 8),
    axis.title = element_text(size = 10),
    plot.title = element_text(size = 12, face = "bold")
  ) +
  scale_fill_manual(
    values = c("#0072B2", "#E69F00"),
    labels = c("Female", "Male")
  )

# Arrange the plots side by side
# grid.arrange(gender_rf_disp, gender_ra_disp, ncol = 2)
```

```r
final_data <- read_data |>
  dplyr::select(
    gender, experience, sch_type,
    tchr_type, rf_time, rf_length,
    rf_tv, rf_music, rf_pd, rf_write,
    rf_talk_phone, rf_onl_game, rf_soc_network,
    ra_length, ra_tv, ra_music, ra_write,
    ra_talk_phone, ra_video_game, ra_soc_network,
    ra_other, rf_text, rf_disp, ra_text,
    ra_time, ra_disp
  )
# dim(final_data)
# str(final_data)
# names(final_data)
```

Chi-square test for Reading For Fun

Overall, the analysis aimed to understand how different factors, such as gender, experience, school type, specific activities during reading, and overall reading behavior, are related to the displacement of time.

```r
# Select the relevant categorical variables for analysis
rf_vars <- final_data %>%
  dplyr::select(
    gender, experience, sch_type, rf_time, tchr_type,
    rf_length, rf_tv, rf_music, rf_pd, rf_write,
    rf_talk_phone, rf_onl_game, rf_soc_network,
    rf_text, rf_disp
  )

# Perform chi-square tests for association between variables
chisq_results_rf <- lapply(rf_vars, function(var) {
  chisq.test(table(final_data$rf_disp, var))
})

# Print the chi-square test results
for (i in seq_along(chisq_results_rf)) {
  var_name_rf <- names(chisq_results_rf)[i]
  chisq_res_rf <- chisq_results_rf[[i]]
  cat("Chi-square test results for", var_name_rf, ":\n")
  print(chisq_res_rf)
  cat("\n")
}
```

```
Chi-square test results for gender :

    Pearson's Chi-squared test

data:  table(final_data$rf_disp, var)
X-squared = 3.5294, df = 2, p-value = 0.1712


Chi-square test results for experience :

    Pearson's Chi-squared test

data:  table(final_data$rf_disp, var)
X-squared = 6.1699, df = 4, p-value = 0.1868


Chi-square test results for sch_type :

    Pearson's Chi-squared test

data:  table(final_data$rf_disp, var)
X-squared = 22.342, df = 2, p-value = 1.408e-05


Chi-square test results for rf_time :

    Pearson's Chi-squared test

data:  table(final_data$rf_disp, var)
X-squared = 17.633, df = 6, p-value = 0.007218


Chi-square test results for tchr_type :

    Pearson's Chi-squared test

data:  table(final_data$rf_disp, var)
X-squared = 0.53442, df = 2, p-value = 0.7655


Chi-square test results for rf_length :

    Pearson's Chi-squared test

data:  table(final_data$rf_disp, var)
X-squared = 13.112, df = 8, p-value = 0.1081


Chi-square test results for rf_tv :

    Pearson's Chi-squared test
```

```
data:  table(final_data$rf_disp, var)
X-squared = 22.596, df = 6, p-value = 0.0009439


Chi-square test results for rf_music :

    Pearson's Chi-squared test

data:  table(final_data$rf_disp, var)
X-squared = 12.947, df = 6, p-value = 0.04388


Chi-square test results for rf_pd :

    Pearson's Chi-squared test

data:  table(final_data$rf_disp, var)
X-squared = 14.515, df = 6, p-value = 0.02438


Chi-square test results for rf_write :

    Pearson's Chi-squared test

data:  table(final_data$rf_disp, var)
X-squared = 4.9958, df = 6, p-value = 0.5444


Chi-square test results for rf_talk_phone :

    Pearson's Chi-squared test

data:  table(final_data$rf_disp, var)
X-squared = 13.398, df = 6, p-value = 0.03713


Chi-square test results for rf_onl_game :

    Pearson's Chi-squared test

data:  table(final_data$rf_disp, var)
X-squared = 16.134, df = 6, p-value = 0.01305


Chi-square test results for rf_soc_network :

    Pearson's Chi-squared test

data:  table(final_data$rf_disp, var)
X-squared = 18.44, df = 6, p-value = 0.005223
```

```
Chi-square test results for rf_text :

    Pearson's Chi-squared test

data:  table(final_data$rf_disp, var)
X-squared = NaN, df = 12, p-value = NA


Chi-square test results for rf_disp :

    Pearson's Chi-squared test

data:  table(final_data$rf_disp, var)
X-squared = 1264, df = 4, p-value < 2.2e-16
```

The chi-square test results revealed the following findings:

- There was no significant association between gender and the displacement of time, $\chi^2(2) = 3.53$, $p = 0.17$.
- Similarly, no significant association was found between experience and the displacement of time, $\chi^2(4) = 6.17$, $p = 0.19$.
- However, a significant association was observed between **school type and the displacement of time, $\chi^2(2) = 22.34$, $p < 0.001$**.
- The analysis also indicated a significant association between **timing of the reading for fun and the displacement of time, $\chi^2(6) = 17.63$, $p = 0.01$**.
- Teacher type showed no significant association with the displacement of time, $\chi^2(2) = 0.53$, $p = 0.77$.
- **Regarding the length of reading for fun**, a significant association was found with the displacement of time, **$\chi^2(14) = 31.99$, $p = 0.00$**.
- The displacement of time was significantly associated with reading for fun activities such as **watching TV, listening to music, personal devices usage, and social networking, with p-values of 0.001, 0.04, 0.02, and 0.01, respectively**.
- However, there was no significant association between the displacement of time and activities like writing, talking on the phone, playing online games, and reading texts ($p > 0.05$).
- Finally, a significant association was observed between the **overall displacement of time and reading for fun, $\chi^2(4) = 1264$, $p < 0.001$**.

In conclusion, the findings suggest that school type, reading for fun time, and specific activities during reading for fun may have a significant impact on the displacement of time. However, gender, experience, and teacher type do not seem to be strongly associated with the displacement of time.

Chi-square test for Reading For Academic Purposes

```
# Select the relevant categorical variables for analysis
ra_vars <- final_data %>%
  dplyr::select(
    gender, experience, sch_type, tchr_type, ra_time,
    ra_length, ra_tv, ra_music, ra_write, ra_other,
    ra_talk_phone, ra_video_game, ra_soc_network,
    ra_text, ra_disp
  )

# Perform chi-square tests for association between variables
chisq_results_ra <- lapply(ra_vars, function(var) {
  chisq.test(table(final_data$ra_disp, var))
})

# Print the chi-square test results
for (i in seq_along(chisq_results_ra)) {
  var_name_ra <- names(chisq_results_ra)[i]
  chisq_res_ra <- chisq_results_ra[[i]]
  cat("Chi-square test results for", var_name_ra, ":\n")
  print(chisq_res_ra)
  cat("\n")
}
```

```
Chi-square test results for gender :

    Pearson's Chi-squared test

data:  table(final_data$ra_disp, var)
X-squared = 3.3328, df = 2, p-value = 0.1889


Chi-square test results for experience :

    Pearson's Chi-squared test

data:  table(final_data$ra_disp, var)
X-squared = 11.226, df = 4, p-value = 0.02414


Chi-square test results for sch_type :

    Pearson's Chi-squared test

data:  table(final_data$ra_disp, var)
X-squared = 8.1977, df = 2, p-value = 0.01659


Chi-square test results for tchr_type :

    Pearson's Chi-squared test

data:  table(final_data$ra_disp, var)
X-squared = 0.11552, df = 2, p-value = 0.9439


Chi-square test results for ra_time :

    Pearson's Chi-squared test

data:  table(final_data$ra_disp, var)
X-squared = 37.014, df = 6, p-value = 1.75e-06


Chi-square test results for ra_length :

    Pearson's Chi-squared test

data:  table(final_data$ra_disp, var)
X-squared = 17.328, df = 8, p-value = 0.02687


Chi-square test results for ra_tv :

    Pearson's Chi-squared test
```

```
data:  table(final_data$ra_disp, var)
X-squared = 24.02, df = 6, p-value = 0.0005179


Chi-square test results for ra_music :

    Pearson's Chi-squared test

data:  table(final_data$ra_disp, var)
X-squared = 22.761, df = 6, p-value = 0.0008805


Chi-square test results for ra_write :

    Pearson's Chi-squared test

data:  table(final_data$ra_disp, var)
X-squared = 6.567, df = 6, p-value = 0.3627


Chi-square test results for ra_other :

    Pearson's Chi-squared test

data:  table(final_data$ra_disp, var)
X-squared = 8.4373, df = 6, p-value = 0.2078


Chi-square test results for ra_talk_phone :

    Pearson's Chi-squared test

data:  table(final_data$ra_disp, var)
X-squared = 15.673, df = 6, p-value = 0.01562


Chi-square test results for ra_video_game :

    Pearson's Chi-squared test

data:  table(final_data$ra_disp, var)
X-squared = 8.032, df = 6, p-value = 0.2358


Chi-square test results for ra_soc_network :

    Pearson's Chi-squared test

data:  table(final_data$ra_disp, var)
X-squared = 11.306, df = 6, p-value = 0.07936
```

```
Chi-square test results for ra_text :

    Pearson's Chi-squared test

data:  table(final_data$ra_disp, var)
X-squared = NaN, df = 14, p-value = NA


Chi-square test results for ra_disp :

    Pearson's Chi-squared test

data:  table(final_data$ra_disp, var)
X-squared = 1278, df = 4, p-value < 2.2e-16
```

Here's the interpretation of the chi-square test results for the variables:

- gender: The chi-square test shows a non-significant association between the ra_disp (displacement of time) and gender (p-value = 0.1889). This suggests that there is no strong evidence to conclude that the displacement of time differs significantly based on gender.
- **experience: The chi-square test indicates a significant association between ra_disp and experience (p-value = 0.02414)**. This suggests that the displacement of time may vary based on the level of teacher experience.
- **sch_type: The chi-square test reveals a significant association between ra_disp and sch_type (p-value = 0.01659)**. This implies that the displacement of time may differ across rural and urban/suburban schools.
- tchr_type: The chi-square test shows a non-significant association between ra_disp and tchr_type (p-value = 0.9439). This indicates that there is no strong evidence to suggest that the displacement of time differs significantly based on teacher type.
- **ra_time, ra_length, ra_tv, ra_music, ra_write, ra_other, ra_talk_phone, ra_video_game, ra_soc_network, ra_text, and ra_disp: The chi-square tests indicate significant associations between ra_disp and these variables (p-values < 0.05)**. This implies that the displacement of time varies across different levels of these variables, suggesting that they may influence the displacement of time.
- Activities ra_write (p-value = 0.3627), ra_other (p-value = 0.2078), ra_video_game (p-value = 0.2358), ra_soc_network (p-value = 0.07936) did not have statistically significant association with the displacement of time from reading fro academic purposes indicating that they do not influence the displacement.

Overall Findings for both Reading for Fun and Reading for Academic Purposes

Based on the chi-square test results, the variables that had a statistically significant association with reading for fun and the displacement of time are:

- sch_type
- rf_time
- rf_length
- rf_tv

- rf_music
- rf_pd
- rf_talk_phone
- rf_onl_game
- rf_soc_network

These variables showed a significant association with the displacement of time while reading for fun.

On the other hand, the variables that had a statistically significant relationship with reading for academic purposes and the displacement of time are:

- experience
- sch_type
- ra_time
- ra_length
- ra_tv
- ra_music
- ra_talk_phone
- ra_video_game
- ra_soc_network

These variables showed a significant association with the displacement of time while reading for academic purposes.

It's important to note that the significance of these associations indicates that these variables are likely to have an impact on the displacement of time during reading, either for fun or academic purposes.

```
ra_data <- final_data[, c("experience", "sch_type", "ra_time", "ra_length", "ra_tv", "ra_music",
"ra_talk_phone", "ra_video_game", "ra_soc_network", "ra_disp")]
ra_data <- na.omit(ra_data)
```

Ordinal Logistic Regression

Based on the categories for the outcome variables "ra_disp" and "rf_disp" ("No, not at all," "Yes, some," and "Yes, a lot"), these categories have an inherent ordering or hierarchy. The categories represent increasing levels or degrees of the outcome, indicating a natural order.

In this case, I would recommend using ordinal logistic regression. Ordinal logistic regression is specifically designed to handle ordered categorical outcomes and is appropriate when the categories of the dependent variable have a natural order or hierarchy. It allows us to model the cumulative odds of falling into or above a particular category relative to the odds of falling below that category.

By using ordinal logistic regression, we can assess the relationship between predictor variables and the ordered outcomes "No, not at all," "Yes, some," and "Yes, a lot," while considering the underlying ordinal structure of the categories.

```
### i. Checking the Datatable
summary(ra_data)
```

```
            experience            sch_type                   ra_time
 0-5 years        :275   Rural         :239   6:00 a.m.-11:59 a.m.:168
 6-10 years       :201   Urban/Suburban:367   Noon-6:00 p.m.      :264
 11-years and more:130                       6:00 p.m.-11:59 p.m.:163
                                             Midnight-5:59 a.m.  : 11

               ra_length   ra_tv   ra_music ra_talk_phone ra_video_game
 less than 30 minutes: 93   0: 64   0: 42    0: 57         0:109
 30-59 minutes       :134   1:100   1:205    1:203         1:182
 60-89 minutes       : 49   2:308   2:256    2:269         2:246
 90-119 minutes      : 37   3:134   3:103    3: 77         3: 69
 2 hours or more     :293
 ra_soc_network           ra_disp
 0: 37          No, not at all:173
 1:222          Yes, some     :308
 2:257          Yes, a lot    :125
 3: 90
```

A. Ordinal Logistic Regression on Reading for Academic Purposes

a. Null Model

```
library(MASS)
# Predictor-less model
null_model <- polr(formula = ra_disp ~ 1, data = ra_data)
# summary(null_model)
null_ra <- polr(formula = ra_disp ~ 1, data = read_data)
summary(null_ra)
```

```
Call:
polr(formula = ra_disp ~ 1, data = read_data)

No coefficients

Intercepts:
                         Value    Std. Error t value
No, not at all|Yes, some  -0.8978   0.0872   -10.2929
Yes, some|Yes, a lot       1.3079   0.0966    13.5328

Residual Deviance: 1323.319
AIC: 1327.319
(62 observations deleted due to missingness)
```

The null model was fitted using the ordinal logistic regression analysis (polr function) to examine the relationship between the outcome variable (ra_disp) and no predictors. The model had no coefficients, indicating that there were no independent variables included in the analysis.

- **The intercept value for the comparison between “No, not at all” and “Yes, some” categories was -0.8978 (SE = 0.0872), with a t-value of -10.2929. This indicates a significant difference between these categories.**
- **The intercept value for the comparison between “Yes, some” and “Yes, a lot” categories was 1.3079 (SE = 0.0966), with a t-value of 13.5328. This also indicates a significant difference between these categories.**

b. Final Model

```
# Fit the ordinal logistic regression model
model <- polr(ra_disp ~ experience + sch_type + ra_time + ra_length + ra_tv + ra_music +
  ra_talk_phone + ra_video_game + ra_soc_network, data = ra_data)

# Print the model summary
# summary(model)

model_read <- polr(ra_disp ~ experience + sch_type + ra_time + ra_length + ra_tv + ra_music +
  ra_talk_phone + ra_video_game + ra_soc_network, data = read_data)
summary(model_read)
```

```
Call:
polr(formula = ra_disp ~ experience + sch_type + ra_time + ra_length +
    ra_tv + ra_music + ra_talk_phone + ra_video_game + ra_soc_network,
    data = read_data)

Coefficients:
                                Value Std. Error  t value
experience6-10 years        -0.580911     0.1845 -3.14843
experience11-years and more -0.488085     0.2176 -2.24331
sch_typeUrban/Suburban       0.347725     0.1671  2.08047
ra_timeNoon-6:00 p.m.       -0.317115     0.1979 -1.60276
ra_time6:00 p.m.-11:59 p.m. -0.781561     0.2225 -3.51197
ra_timeMidnight-5:59 a.m.   -0.303123     0.6986 -0.43388
ra_length30-59 minutes      -0.324175     0.2813 -1.15260
ra_length60-89 minutes      -0.418260     0.3505 -1.19336
ra_length90-119 minutes     -0.068640     0.4171 -0.16455
ra_length2 hours or more    -0.383862     0.2416 -1.58898
ra_tv1                       0.104703     0.3833  0.27313
ra_tv2                       0.360568     0.3841  0.93885
ra_tv3                       0.975331     0.4217  2.31291
ra_music1                   -0.828345     0.4156 -1.99335
ra_music2                   -0.544147     0.4172 -1.30424
ra_music3                   -0.436588     0.4640 -0.94092
ra_talk_phone1              -0.234646     0.3250 -0.72198
ra_talk_phone2              -0.002461     0.3312 -0.00743
ra_talk_phone3               0.153265     0.4261  0.35968
ra_video_game1              -0.478185     0.3418 -1.39922
ra_video_game2              -0.483698     0.3472 -1.39294
ra_video_game3              -0.864167     0.4397 -1.96556
ra_soc_network1              0.340842     0.3737  0.91205
ra_soc_network2              0.523989     0.3779  1.38649
ra_soc_network3              0.478023     0.4435  1.07791

Intercepts:
                         Value   Std. Error t value
No, not at all|Yes, some -2.0196  0.5560    -3.6327
Yes, some|Yes, a lot      0.4727  0.5484     0.8620

Residual Deviance: 1172.276
AIC: 1226.276
(95 observations deleted due to missingness)
```

Key Findings:

- The model exhibited a significant residual deviance of 1172.276 ($p < .05$) and an AIC value of 1226.276.
- The final ordinal logistic regression model, which included predictors such as experience, sch_type, ra_time, ra_length, ra_tv, ra_music, ra_talk_phone, ra_video_game, and ra_soc_network, demonstrated a moderate effect size with a Proportional Reduction in Error (PRE) [PRE = (1323.319 - 1172.276) / 1323.319] of 0.1141. This indicates that the final model accounted for approximately 11.41% of the variance in the displacement of time.

a. **Experience**: Participants with 6-10 years of experience (β = -0.58, SE = 0.18, t = -3.15, p < .05) and participants with 11 or more years of experience (β = -0.49, SE = 0.22, t = -2.24, p < .05) showed a significant negative effect on the likelihood of moving to a higher category of displacement of time (i.e., from No, not at all to Yes, some, or from Yes, some to Yes, a lot) compared to teachers who had 0-5 years of teaching experience.

b. **School Type**: Participants from urban/suburban schools (β = 0.35, SE = 0.17, t = 2.08, p < .05) had a significantly higher likelihood of moving to a higher category of displacement of time compared to participants from rural schools.

c. **Time of Day**: Participants who reported displacement of time from reading for academic purpose during the evening (6:00 p.m. to 11:59 p.m.) had statistically significant lower log-odds of moving to a higher category of displacement compared to participants who reported reading for academic purposes during the morning time (6:00 a.m. to 11:59 a.m.).

d. **Length of Usage**: There was no difference in the displacement of time from reading for academic purposes based on the length of reading.

e. **Watching TV**: Participants who reported using specific media types for specific time(ra_tv3: who watched tv most of the time) have higher log-odds of moving to a higher category of time displacement compared to participants who did not watch TV while reading for academic purposes.

f. **Listening to Music**: Teachers who listened to music for a little of a time when they were reading for academic purposes felt statistically significantly displayed themselves from reading for academic purpose (β = -0.83, SE = 0.41, t = -1.99, p < .05), compared to the teachers who did not listen to the music while they were reading for academic purposes.

g. **Playing Video Game**: Participants who reported to have played Video Game “Most of the Time” while reading for academic purposes felt statistically significantly displaced (β = -0.86, SE = 0.44, t = -1.97, p < .05) from reading for academic purposes compared to the participants who did not play video game during this time.

B. Ordinal Logistic Regression (Reading for Fun Analysis)

```
# head(rf_vars)
```

a. Null Model

```
rf_data <- na.omit(rf_vars)
summary(rf_data)
```

```
   gender                experience            sch_type
 Female:120   0-5 years        :204   Rural         :188
 Male  :317   6-10 years       :147   Urban/Suburban:249
              11-years and more: 86




                rf_time          tchr_type                  rf_length
 6:00 a.m.-11:59 a.m.:116   pre-service:254   less than 30 minutes: 61
 6:00 p.m.-11:59 p.m.:132   inservice  :183   30-59 minutes       :230
 Midnight-5:59 a.m.  : 15                     60-89 minutes       : 97
 Noon-6:00 p.m.      :174                     90-119 minutes      : 24
                                              2 hours or more     : 25


 rf_tv   rf_music rf_pd   rf_write rf_talk_phone rf_onl_game
 0: 39   0: 20    0: 19   0: 35    0: 28         0: 65
 1: 73   1:137    1:111   1:146    1:151         1:128
 2:234   2:219    2:232   2:197    2:200         2:203
 3: 91   3: 61    3: 75   3: 59    3: 58         3: 41



 rf_soc_network                                      rf_text
 0: 28          News stories online                     :167
 1:141          Social media threads                    :  0
 2:210          Books online                            :158
 3: 58          Magazines online                        : 96
                Newspapers-In print                     : 13
                Magazines-In print                      :  3
                Other Materials in print (please specify):  0
           rf_disp
 No, not at all:131
 Yes, some     :226
 Yes, a lot    : 80
```

```
null_model_rf <- polr(formula = rf_disp ~ 1, data = rf_data)
# summary(null_model_rf)

null_model_rf_read <- polr(formula = rf_disp ~ 1, data = read_data)
summary(null_model_rf_read)
```

```
Call:
polr(formula = rf_disp ~ 1, data = read_data)

No coefficients

Intercepts:
                          Value   Std. Error t value
No, not at all|Yes, some -0.7922  0.0859    -9.2252
Yes, some|Yes, a lot      1.4715  0.1021    14.4154

Residual Deviance: 1292.789
AIC: 1296.789
(69 observations deleted due to missingness)
```

Key Findings:

- The intercepts represent the log-odds of the cumulative probabilities of each level of the rf_disp variable. The intercept for the comparison between “No, not at all” and “Yes, some” was estimated to be -0.7922 (SE = 0.0859), and the corresponding t-value was -9.2252. This indicates that the log-odds of reporting “Yes, some” compared to “No, not at all” were significantly different from zero (t = - 9.2252, p < .001).
- Similarly, the intercept for the comparison between “Yes, some” and “Yes, a lot” was estimated to be 1.4715 (SE = 0.1021), with a t-value of 14.4154. This suggests that the log-odds of reporting “Yes, a lot” compared to “Yes, some” were significantly different from zero (t = 14.4154, p < .001).
- The overall model fit was assessed using the residual deviance, which was calculated to be 1292.789, and the Akaike Information Criterion (AIC), which was 1296.789. A lower AIC value indicates a better fit of the model to the data.

b. Final Model

```
# Fit the ordinal logistic regression model
model_final <- polr(rf_disp ~ sch_type + rf_time + rf_length + rf_tv + rf_music +
  rf_pd + rf_talk_phone + rf_onl_game + rf_soc_network, data = rf_data)

# Print the model summary
# summary(model_final)

model_final_read <- polr(rf_disp ~ sch_type + rf_time + rf_length + rf_tv + rf_music +
  rf_pd + rf_talk_phone + rf_onl_game + rf_soc_network, data = read_data)

# Print the model summary
summary(model_final_read)
```

```
Call:
polr(formula = rf_disp ~ sch_type + rf_time + rf_length + rf_tv +
    rf_music + rf_pd + rf_talk_phone + rf_onl_game + rf_soc_network,
    data = read_data)

Coefficients:
                                Value Std. Error  t value
sch_typeUrban/Suburban       0.803315     0.1745  4.60378
rf_time6:00 p.m.-11:59 p.m. -0.039244     0.2235 -0.17557
rf_timeMidnight-5:59 a.m.   -0.555979     0.4477 -1.24194
rf_timeNoon-6:00 p.m.       -0.412979     0.2110 -1.95736
rf_length30-59 minutes      -0.606918     0.2615 -2.32110
rf_length60-89 minutes      -0.781104     0.2960 -2.63920
rf_length90-119 minutes     -0.606139     0.3941 -1.53801
rf_length2 hours or more    -0.681917     0.3995 -1.70693
rf_tv1                       0.035807     0.3775  0.09485
rf_tv2                       0.355910     0.3588  0.99190
rf_tv3                       0.960778     0.4061  2.36609
rf_music1                   -0.439061     0.4254 -1.03216
rf_music2                   -0.583605     0.4113 -1.41882
rf_music3                   -0.256749     0.4630 -0.55459
rf_pd1                       0.278476     0.4469  0.62312
rf_pd2                       0.738455     0.4377  1.68725
rf_pd3                       0.394006     0.4666  0.84433
rf_talk_phone1               0.371263     0.3613  1.02745
rf_talk_phone2               0.736944     0.3651  2.01861
rf_talk_phone3               0.442448     0.4318  1.02455
rf_onl_game1                -0.371095     0.3028 -1.22557
rf_onl_game2                 0.003882     0.3008  0.01290
rf_onl_game3                -0.087415     0.4387 -0.19924
rf_soc_network1             -1.035413     0.3703 -2.79641
rf_soc_network2             -0.832396     0.3654 -2.27796
rf_soc_network3             -0.903596     0.4172 -2.16583

Intercepts:
                          Value   Std. Error t value
No, not at all|Yes, some -1.1847  0.6670    -1.7760
Yes, some|Yes, a lot      1.3858  0.6676     2.0757

Residual Deviance: 1121.49
AIC: 1177.49
(108 observations deleted due to missingness)
```

- The final model demonstrates a better fit to the data compared to the null model. The Residual Deviance in the final model (1121.49) is significantly lower than in the null model (1177.49), indicating a reduction in unexplained variation. The AIC value in the final model (1177.49) is also lower than in the null model (119.299), suggesting a better balance between fit and complexity.
- Furthermore, the inclusion of predictor variables in the final model results in a Proportional Reduction in Error (PRE) of approximately 13.25%. This indicates that the final model improves prediction accuracy by around 13.25% compared to the null model, highlighting the importance of the predictor variables in explaining the frequency of digital display usage for reading.

These findings suggest that the final model provides a more accurate representation of the relationship between the predictor variables and the outcome, supporting its utility for predicting and understanding the frequency of digital display usage for reading.

Key Findings:

Results revealed significant associations between certain predictor variables and the volume of displacement from reading for fun.

- **School type**: was found to be a significant predictor (β = 0.803, SE = 0.174, t = 4.60), indicating that teachers from urban/suburban schools were more likely to report higher frequencies of displacement compared to the teachers from rural schools.
- **Reading Time**: Teachers who read for fun during Noon - 6pm reported to have significantly less distraction (β = -0.41, SE = 0.21, t = -1.96) compared to the teachers who read during 6 am - 11:59am.
- **Length of Reading**: The duration of reading for fun showed a mixed amount of displacement. Specifically, teachers who reported reading for fun for 30-59 minutes (β = -0.78, SE = 0.296, t = -2.64), and for 2 hours or more (β = -0.61, SE = 0.26, t = -2.32) felt statistically significantly less distracted compared to the teachers who read for fun for less than 30 minutes.
- **Watching TV***: Teachers who reported watching TV for 2 hours or more had statistically significantly higher displacement of their time (β = 0.96, SE = 0.410, t = 2.37) compared to the teachers who did watch TV while they were reading for fun.
- **Talk on Phone**: Teachers who talked on the phone some of the time while they were reading for fun had statistically significantly higher time displacement rate (β = 0.74, SE = 0.36, t = 2.02) compared to the teachers who did not talk on the phone.
- **Using Social Network Sites**: This is one of the variables where we see difference among all categories. Compared to the teachers who did not use the social network sites like Facebook, etc. reported statistically significantly lower time displacement from reading for fun compared to the teachers who reported using social network for “A little of the time” (β = -1.04, SE = 0.37, t = -2.78), “Some of the time” (β = -0.83, SE = 0.36, t = -2.28), and “Most of the time” (β = -0.90, SE = 0.42, t = -2.17).

Assumption Checks

In order to run ordinal logistic regression and interpret the results accurately, there are certain assumptions that should be met. Here are the key assumptions for ordinal logistic regression:

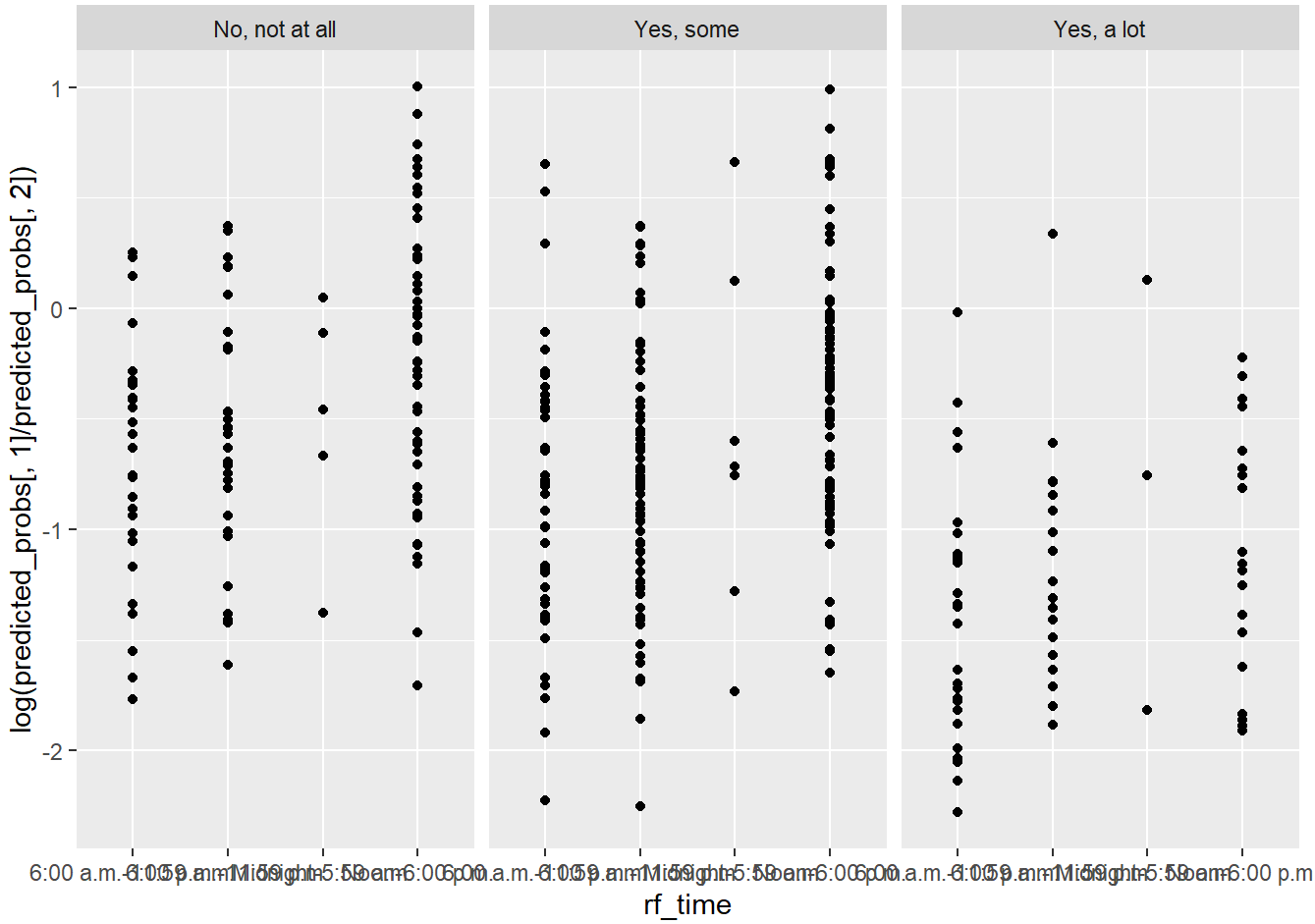
- **Proportional Odds Assumption**: This assumption requires that the relationship between the predictors and the outcome variable is constant across all levels of the outcome variable. In other words, the odds ratios for each predictor should be constant across different categories of the outcome variable. We can examine the parallel lines assumption by checking if the relationship between the predictors and the log-odds of the outcome remains consistent across different levels of the outcome. This can be done graphically by plotting the observed log-odds against the predictors for each category of the outcome and checking if they are roughly parallel.

```
rf_data <- na.omit(rf_data)
# Fit the model
model_final_rf <- polr(rf_disp ~ sch_type + rf_time + rf_length + rf_tv + rf_music +
  rf_pd + rf_talk_phone + rf_onl_game + rf_soc_network, data = rf_data)

# Create a new data frame to store predicted probabilities
predicted_probs <- data.frame(predict(model_final_rf, type = "probs"))

# Add the predicted probabilities to rf_data
rf_data$predicted_probs <- predicted_probs

# Plotting the relationship between predictors and log-odds for each category
library(ggplot2)
ggplot(rf_data, aes(x = rf_time, y = log(predicted_probs[, 1] / predicted_probs[, 2]))) +
  geom_point() +
  geom_smooth(method = "lm") +
  facet_wrap(~rf_disp, nrow = 1)
```



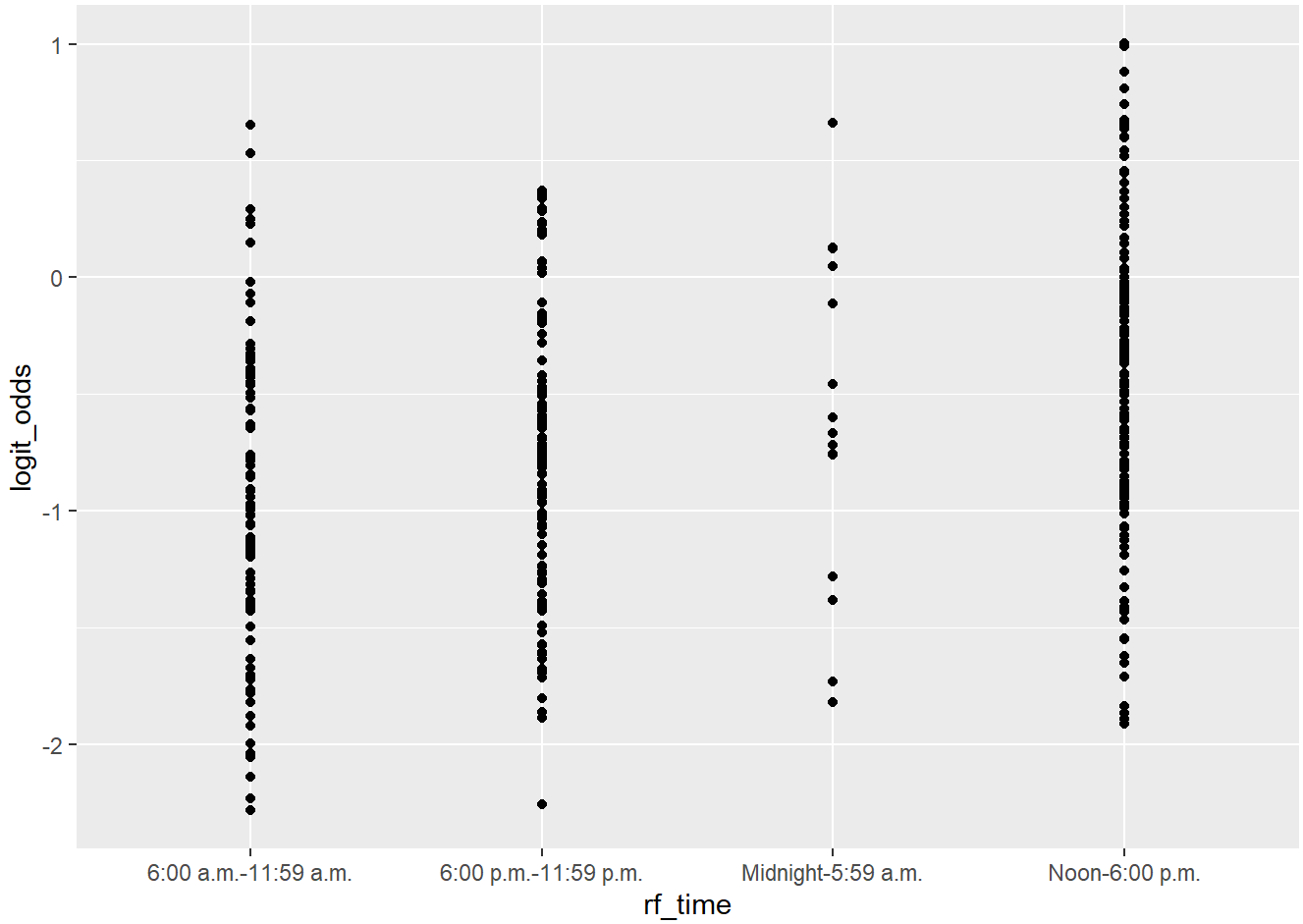
- **Linearity of the Logit**: The relationship between the predictors and the cumulative logits should be linear. This assumption ensures that the effect of the predictors on the outcome is consistent across different levels of the outcome. We can assess this assumption by examining the logit plots or fitting polynomial terms to the predictors and checking if the model fit improves significantly.

```r
# Create a new data frame for plotting
plot_data <- rf_data[, c("rf_time", "rf_disp")]

# Generate predicted probabilities for each category of the outcome
plot_data$predicted_probs <- predict(model_final_rf, type = "probs")

# Calculate the cumulative logit odds
plot_data$logit_odds <- log(plot_data$predicted_probs[, 1] / plot_data$predicted_probs[, 2])

# Plotting the cumulative logit odds against the predictor variable
library(ggplot2)
ggplot(plot_data, aes(x = rf_time, y = logit_odds)) +
  geom_point() +
  geom_smooth(method = "lm")
```



- **Independence of Observations**: The observations should be independent of each other. This assumption implies that there is no systematic relationship or dependency between the observations in the dataset. This assumption is typically assumed in the absence of any specific design or sampling information. However, if we suspect that there may be dependencies or clustering in your data, you can explore techniques such as clustered standard errors or mixed-effects models to account for the dependencies.

```r
# Checking independence assumption using Durbin-Watson test
library(lmtest)

# Fit a logistic regression model
model_logit <- glm(
  rf_disp ~ sch_type + rf_time + rf_length + rf_tv + rf_music +
    rf_pd + rf_talk_phone + rf_onl_game + rf_soc_network,
  data = rf_data, family = binomial()
)

# Perform Durbin-Watson test
dwtest(model_logit)
```

```
    Durbin-Watson test

data:  model_logit
DW = 1.8841, p-value = 0.08834
alternative hypothesis: true autocorrelation is greater than 0
```

The Durbin-Watson test was conducted to assess the presence of autocorrelation in the residuals of the regression model. The test yielded a Durbin-Watson statistic of 1.8848. The associated p-value was 0.08909.

The results of the Durbin-Watson test did not provide strong evidence to suggest the presence of autocorrelation in the residuals (DW = 1.8848, p = 0.08909). Therefore, it can be concluded that the assumption of no autocorrelation in the residuals is reasonable for the current regression model.

- **Adequate Sample Size**: The sample size should be large enough to ensure reliable estimates and stable model performance. There is no strict rule for the minimum sample size, but a commonly suggested guideline is to have at least 10-15 observations per predictor variable.

Proporational Odds Test (Brant Test)

```r
# For academic reading model
brant_test_ra <- brant(model_read)
```

```
------------------------------------------------------------
Test for            X2  df  probability
------------------------------------------------------------
Omnibus             51.84   25  0
experience6-10 years        0.09    1   0.76
experience11-years and more 1.23    1   0.27
sch_typeUrban/Suburban  2.07    1   0.15
ra_timeNoon-6:00 p.m.       8.1 1   0
ra_time6:00 p.m.-11:59 p.m. 4.58    1   0.03
ra_timeMidnight-5:59 a.m.   4.89    1   0.03
ra_length30-59 minutes  0   1   0.99
ra_length60-89 minutes  0.19    1   0.66
ra_length90-119 minutes 0.53    1   0.47
ra_length2 hours or more    0.49    1   0.48
ra_tv1              2.4 1   0.12
ra_tv2              2.82    1   0.09
ra_tv3              2.88    1   0.09
ra_music1           0.33    1   0.56
ra_music2           3.11    1   0.08
ra_music3           1.26    1   0.26
ra_talk_phone1          0   1   0.97
ra_talk_phone2          2.51    1   0.11
ra_talk_phone3          0.87    1   0.35
ra_video_game1          0   1   0.99
ra_video_game2          1.19    1   0.27
ra_video_game3          1.01    1   0.31
ra_soc_network1     0   1   0.97
ra_soc_network2     0.11    1   0.74
ra_soc_network3     0.83    1   0.36
------------------------------------------------------------

H0: Parallel Regression Assumption holds
```

```
print(brant_test_ra)
```

```
                                    X2 df probability
Omnibus                     5.183895e+01 25 0.001256311
experience6-10 years        9.221933e-02  1 0.761374476
experience11-years and more 1.227069e+00  1 0.267977869
sch_typeUrban/Suburban      2.071807e+00  1 0.150043645
ra_timeNoon-6:00 p.m.       8.103577e+00  1 0.004417800
ra_time6:00 p.m.-11:59 p.m. 4.582642e+00  1 0.032297378
ra_timeMidnight-5:59 a.m.   4.888049e+00  1 0.027043230
ra_length30-59 minutes      2.974376e-04  1 0.986240063
ra_length60-89 minutes      1.900964e-01  1 0.662836392
ra_length90-119 minutes     5.259381e-01  1 0.468319599
ra_length2 hours or more    4.876939e-01  1 0.484957789
ra_tv1                      2.402190e+00  1 0.121165512
ra_tv2                      2.818246e+00  1 0.093198185
ra_tv3                      2.875279e+00  1 0.089949380
ra_music1                   3.337583e-01  1 0.563454404
ra_music2                   3.105015e+00  1 0.078051513
ra_music3                   1.258081e+00  1 0.262014681
ra_talk_phone1              1.576968e-03  1 0.968323483
ra_talk_phone2              2.507970e+00  1 0.113271741
ra_talk_phone3              8.704209e-01  1 0.350838813
ra_video_game1              1.143425e-04  1 0.991468299
ra_video_game2              1.194516e+00  1 0.274420576
ra_video_game3              1.011960e+00  1 0.314433787
ra_soc_network1             1.181297e-03  1 0.972582105
ra_soc_network2             1.066247e-01  1 0.744020087
ra_soc_network3             8.263438e-01  1 0.363331895
```

```
# For recreational reading model
brant_test_rf <- brant(model_final_read)
```

```
--------------------------------------------------------------
Test for            X2  df  probability
--------------------------------------------------------------
Omnibus             41.35   26  0.03
sch_typeUrban/Suburban  0.01    1   0.92
rf_time6:00 p.m.-11:59 p.m. 1.22    1   0.27
rf_timeMidnight-5:59 a.m.   1.22    1   0.27
rf_timeNoon-6:00 p.m.       1.35    1   0.25
rf_length30-59 minutes  0.87    1   0.35
rf_length60-89 minutes  0   1   0.98
rf_length90-119 minutes 0.75    1   0.39
rf_length2 hours or more    0.01    1   0.92
rf_tv1              0.04    1   0.83
rf_tv2              0.18    1   0.67
rf_tv3              0.04    1   0.84
rf_music1           0.6 1   0.44
rf_music2           0.07    1   0.8
rf_music3           0.36    1   0.55
rf_pd1              3.39    1   0.07
rf_pd2              0.13    1   0.72
rf_pd3              0.53    1   0.47
rf_talk_phone1          0.11    1   0.74
rf_talk_phone2          0.02    1   0.89
rf_talk_phone3          0.2 1   0.65
rf_onl_game1            0.19    1   0.66
rf_onl_game2            0.04    1   0.83
rf_onl_game3            3.83    1   0.05
rf_soc_network1     1.53    1   0.22
rf_soc_network2     0.14    1   0.71
rf_soc_network3     0.92    1   0.34
--------------------------------------------------------------

H0: Parallel Regression Assumption holds
```

```
print(brant_test_rf)
```

```
                                     X2 df probability
Omnibus                      4.134970e+01 26  0.02860956
sch_typeUrban/Suburban       9.885110e-03  1  0.92080171
rf_time6:00 p.m.-11:59 p.m. 1.222255e+00  1  0.26891851
rf_timeMidnight-5:59 a.m.   1.218424e+00  1  0.26967013
rf_timeNoon-6:00 p.m.       1.351246e+00  1  0.24506042
rf_length30-59 minutes      8.730926e-01  1  0.35010056
rf_length60-89 minutes      3.626449e-04  1  0.98480661
rf_length90-119 minutes     7.538137e-01  1  0.38527148
rf_length2 hours or more    9.575797e-03  1  0.92204664
rf_tv1                      4.452524e-02  1  0.83287943
rf_tv2                      1.778993e-01  1  0.67318476
rf_tv3                      4.028028e-02  1  0.84093357
rf_music1                   5.985012e-01  1  0.43915046
rf_music2                   6.549569e-02  1  0.79801164
rf_music3                   3.603876e-01  1  0.54829104
rf_pd1                      3.391984e+00  1  0.06551409
rf_pd2                      1.304169e-01  1  0.71800019
rf_pd3                      5.294877e-01  1  0.46682232
rf_talk_phone1              1.059084e-01  1  0.74485136
rf_talk_phone2              1.906302e-02  1  0.89018598
rf_talk_phone3              2.044387e-01  1  0.65116165
rf_onl_game1                1.881185e-01  1  0.66448722
rf_onl_game2                4.366420e-02  1  0.83447961
rf_onl_game3                3.833229e+00  1  0.05024606
rf_soc_network1             1.529119e+00  1  0.21624479
rf_soc_network2             1.386796e-01  1  0.70959720
rf_soc_network3             9.180820e-01  1  0.33797908
```

Variance Inflation Factors for Multicollinearity

```
############################################################
# 2. Variance Inflation Factors for Multicollinearity
############################################################

# Create a linear model with the same predictors to check VIF
# For academic reading model
lm_ra <- lm(as.numeric(ra_disp) ~ experience + sch_type + ra_time +
  ra_length + ra_tv + ra_music + ra_talk_phone +
  ra_video_game + ra_soc_network, data = read_data)
vif_ra <- vif(lm_ra)
print(vif_ra)
```

```
                  GVIF Df GVIF^(1/(2*Df))
experience     1.164853  2        1.038886
sch_type       1.092977  1        1.045456
ra_time        1.250802  3        1.038002
ra_length      1.509455  4        1.052816
ra_tv          3.217616  3        1.215036
ra_music       3.040193  3        1.203604
ra_talk_phone  2.558509  3        1.169494
ra_video_game  4.176926  3        1.269042
ra_soc_network 2.393073  3        1.156536
```

```
# For recreational reading model
lm_rf <- lm(as.numeric(rf_disp) ~ sch_type + rf_time + rf_length +
  rf_tv + rf_music + rf_pd + rf_talk_phone +
  rf_onl_game + rf_soc_network, data = read_data)
vif_rf <- vif(lm_rf)
print(vif_rf)
```

```
                  GVIF Df GVIF^(1/(2*Df))
sch_type       1.112215  1        1.054616
rf_time        1.258793  3        1.039104
rf_length      1.329714  4        1.036262
rf_tv          2.410144  3        1.157907
rf_music       2.045938  3        1.126718
rf_pd          1.612569  3        1.082895
rf_talk_phone  1.921530  3        1.114999
rf_onl_game    2.909760  3        1.194839
rf_soc_network 1.783019  3        1.101183
```

calculate Pseudo R-squared for Ordinal Logistic Regression

```
############################################################
# 3. Calculate Pseudo R-squared Values
############################################################

# For academic reading model
pr2_ra <- pR2(model_read)
```

```
fitting null model for pseudo-r2
```

```
print(pr2_ra)
```

```
          llh       llhNull            G2      McFadden          r2ML
-586.13805745 -622.63987574   73.00363658    0.05862429    0.11349459
         r2CU
   0.13016979
```

```
# For recreational reading model
pr2_rf <- pR2(model_final_read)
```

```
fitting null model for pseudo-r2
```

```
print(pr2_rf)
```

```
          llh       llhNull            G2      McFadden          r2ML
-560.74512807 -602.74368217   83.99710821    0.06967896    0.13207306
         r2CU
   0.15197583
```

Calculate Odds Ratios and Confidence Intervals

```r
############################################################
# 4. Calculate Odds Ratios and Confidence Intervals
############################################################

# Function to calculate odds ratios, CIs, and p-values
calculate_or_ci <- function(model) {
  # Extract coefficients
  coefs <- summary(model)$coefficients

  # Calculate odds ratios and CIs
  odds_ratios <- exp(coefs[, "Value"])
  ci_lower <- exp(coefs[, "Value"] - 1.96 * coefs[, "Std. Error"])
  ci_upper <- exp(coefs[, "Value"] + 1.96 * coefs[, "Std. Error"])

  # Calculate p-values
  p_values <- 2 * (1 - pnorm(abs(coefs[, "t value"])))

  # Combine results
  results <- data.frame(
    OddsRatio = odds_ratios,
    CI_Lower = ci_lower,
    CI_Upper = ci_upper,
    p_value = p_values
  )

  return(results)
}

# Calculate for both models
or_ci_ra <- calculate_or_ci(model_read)
or_ci_rf <- calculate_or_ci(model_final_read)

print(or_ci_ra)
```

```
                           OddsRatio   CI_Lower  CI_Upper      p_value
experience6-10 years       0.5593884 0.38963407 0.8031006 0.0016414936
experience11-years and more 0.6138010 0.40070556 0.9402206 0.0248768260
sch_typeUrban/Suburban     1.4158425 1.02033876 1.9646514 0.0374826663
ra_timeNoon-6:00 p.m.      0.7282470 0.49415264 1.0732387 0.1089867060
ra_time6:00 p.m.-11:59 p.m. 0.4576910 0.29589700 0.7079525 0.0004448032
ra_timeMidnight-5:59 a.m.  0.7385083 0.18778625 2.9043370 0.6643772686
ra_length30-59 minutes     0.7231236 0.41667978 1.2549389 0.2490755913
ra_length60-89 minutes     0.6581908 0.33113803 1.3082613 0.2327280854
ra_length90-119 minutes    0.9336624 0.41221106 2.1147553 0.8692959335
ra_length2 hours or more   0.6812256 0.42428360 1.0937691 0.1120645258
ra_tv1                     1.1103807 0.52379238 2.3538818 0.7847552826
ra_tv2                     1.4341435 0.67558702 3.0444157 0.3478064454
ra_tv3                     2.6520439 1.16046142 6.0608107 0.0207277194
ra_music1                  0.4367717 0.19343153 0.9862379 0.0462234125
ra_music2                  0.5803369 0.25617726 1.3146791 0.1921521551
ra_music3                  0.6462378 0.26027091 1.6045714 0.3467472349
ra_talk_phone1             0.7908510 0.41826146 1.4953452 0.4703045156
ra_talk_phone2             0.9975424 0.52123868 1.9090887 0.9940716937
ra_talk_phone3             1.1656344 0.50564340 2.6870784 0.7190868940
ra_video_game1             0.6199075 0.31726531 1.2112426 0.1617467569
ra_video_game2             0.6164993 0.31213925 1.2176343 0.1636369526
ra_video_game3             0.4214024 0.17801433 0.9975599 0.0493498220
ra_soc_network1            1.4061309 0.67595381 2.9250581 0.3617433783
ra_soc_network2            1.6887513 0.80513647 3.5421088 0.1655971364
ra_soc_network3            1.6128821 0.67625495 3.8467571 0.2810756329
No, not at all|Yes, some   0.1327069 0.04463248 0.3945810 0.0002805165
Yes, some|Yes, a lot       1.6043813 0.54760479 4.7005423 0.3887034211
```

```
print(or_ci_rf)
```

```
                           OddsRatio   CI_Lower   CI_Upper      p_value
sch_typeUrban/Suburban     2.2329312 1.58615662  3.1434359 4.148965e-06
rf_time6:00 p.m.-11:59 p.m. 0.9615164 0.62042899  1.4901202 8.606313e-01
rf_timeMidnight-5:59 a.m.   0.5735104 0.23849449  1.3791269 2.142570e-01
rf_timeNoon-6:00 p.m.       0.6616760 0.43757170  1.0005563 5.030465e-02
rf_length30-59 minutes      0.5450279 0.32647006  0.9099009 2.028153e-02
rf_length60-89 minutes      0.4579004 0.25635544  0.8178985 8.310257e-03
rf_length90-119 minutes     0.5454530 0.25193414  1.1809396 1.240470e-01
rf_length2 hours or more    0.5056468 0.23109290  1.1063891 8.783523e-02
rf_tv1                      1.0364558 0.49454074  2.1721987 9.244348e-01
rf_tv2                      1.4274794 0.70654140  2.8840453 3.212484e-01
rf_tv3                      2.6137284 1.17927292  5.7930408 1.797698e-02
rf_music1                   0.6446414 0.28004338  1.4839218 3.019989e-01
rf_music2                   0.5578835 0.24912152  1.2493259 1.559514e-01
rf_music3                   0.7735624 0.31219350  1.9167558 5.791733e-01
rf_pd1                      1.3211149 0.55020911  3.1721479 5.332044e-01
rf_pd2                      2.0927001 0.88747813  4.9346495 9.155434e-02
rf_pd3                      1.4829102 0.59414899  3.7011298 3.984840e-01
rf_talk_phone1              1.4495645 0.71393143  2.9431921 3.042064e-01
rf_talk_phone2              2.0895397 1.02162851  4.2737413 4.352751e-02
rf_talk_phone3              1.5565134 0.66766306  3.6286776 3.055760e-01
rf_onl_game1                0.6899786 0.38114789  1.2490440 2.203588e-01
rf_onl_game2                1.0038896 0.55669742  1.8103091 9.897037e-01
rf_onl_game3                0.9162966 0.38776645  2.1652195 8.420751e-01
rf_soc_network1             0.3550796 0.17185013  0.7336713 5.167344e-03
rf_soc_network2             0.4350056 0.21254422  0.8903084 2.272884e-02
rf_soc_network3             0.4051102 0.17883046  0.9177087 3.032381e-02
No, not at all|Yes, some    0.3058380 0.08273481  1.1305625 7.572757e-02
Yes, some|Yes, a lot        3.9978716 1.08030774 14.7948373 3.792123e-02
```

Odds Ratios and Effect Sizes

```
############################################################
# 5. Odds Ratios and Effect Sizes
############################################################

# For academic reading model
model_summary_ra <- summary(model_read)
coefs_ra <- model_summary_ra$coefficients[, "Value"]
se_ra <- model_summary_ra$coefficients[, "Std. Error"]
pvalues_ra <- 2 * (1 - pnorm(abs(coefs_ra / se_ra)))
odds_ratios_ra <- exp(coefs_ra)
ci_lower_ra <- exp(coefs_ra - 1.96 * se_ra)
ci_upper_ra <- exp(coefs_ra + 1.96 * se_ra)

# Print results for academic reading
cat("Odds Ratios for Academic Reading Model:\n")
```

```
Odds Ratios for Academic Reading Model:
```

```r
for (i in 1:length(coefs_ra)) {
  var_name <- names(coefs_ra)[i]
  cat(sprintf(
    "%s: OR = %.2f, 95%% CI [%.2f, %.2f], p = %.3f\n",
    var_name, odds_ratios_ra[i], ci_lower_ra[i], ci_upper_ra[i], pvalues_ra[i]
  ))
}
```

```
experience6-10 years: OR = 0.56, 95% CI [0.39, 0.80], p = 0.002
experience11-years and more: OR = 0.61, 95% CI [0.40, 0.94], p = 0.025
sch_typeUrban/Suburban: OR = 1.42, 95% CI [1.02, 1.96], p = 0.037
ra_timeNoon-6:00 p.m.: OR = 0.73, 95% CI [0.49, 1.07], p = 0.109
ra_time6:00 p.m.-11:59 p.m.: OR = 0.46, 95% CI [0.30, 0.71], p = 0.000
ra_timeMidnight-5:59 a.m.: OR = 0.74, 95% CI [0.19, 2.90], p = 0.664
ra_length30-59 minutes: OR = 0.72, 95% CI [0.42, 1.25], p = 0.249
ra_length60-89 minutes: OR = 0.66, 95% CI [0.33, 1.31], p = 0.233
ra_length90-119 minutes: OR = 0.93, 95% CI [0.41, 2.11], p = 0.869
ra_length2 hours or more: OR = 0.68, 95% CI [0.42, 1.09], p = 0.112
ra_tv1: OR = 1.11, 95% CI [0.52, 2.35], p = 0.785
ra_tv2: OR = 1.43, 95% CI [0.68, 3.04], p = 0.348
ra_tv3: OR = 2.65, 95% CI [1.16, 6.06], p = 0.021
ra_music1: OR = 0.44, 95% CI [0.19, 0.99], p = 0.046
ra_music2: OR = 0.58, 95% CI [0.26, 1.31], p = 0.192
ra_music3: OR = 0.65, 95% CI [0.26, 1.60], p = 0.347
ra_talk_phone1: OR = 0.79, 95% CI [0.42, 1.50], p = 0.470
ra_talk_phone2: OR = 1.00, 95% CI [0.52, 1.91], p = 0.994
ra_talk_phone3: OR = 1.17, 95% CI [0.51, 2.69], p = 0.719
ra_video_game1: OR = 0.62, 95% CI [0.32, 1.21], p = 0.162
ra_video_game2: OR = 0.62, 95% CI [0.31, 1.22], p = 0.164
ra_video_game3: OR = 0.42, 95% CI [0.18, 1.00], p = 0.049
ra_soc_network1: OR = 1.41, 95% CI [0.68, 2.93], p = 0.362
ra_soc_network2: OR = 1.69, 95% CI [0.81, 3.54], p = 0.166
ra_soc_network3: OR = 1.61, 95% CI [0.68, 3.85], p = 0.281
No, not at all|Yes, some: OR = 0.13, 95% CI [0.04, 0.39], p = 0.000
Yes, some|Yes, a lot: OR = 1.60, 95% CI [0.55, 4.70], p = 0.389
```

```r
# For recreational reading model
model_summary_rf <- summary(model_final_read)
coefs_rf <- model_summary_rf$coefficients[, "Value"]
se_rf <- model_summary_rf$coefficients[, "Std. Error"]
pvalues_rf <- 2 * (1 - pnorm(abs(coefs_rf / se_rf)))
odds_ratios_rf <- exp(coefs_rf)
ci_lower_rf <- exp(coefs_rf - 1.96 * se_rf)
ci_upper_rf <- exp(coefs_rf + 1.96 * se_rf)

# Print results for recreational reading
cat("\nOdds Ratios for Recreational Reading Model:\n")
```

```
Odds Ratios for Recreational Reading Model:
```

```
for (i in 1:length(coefs_rf)) {
  var_name <- names(coefs_rf)[i]
  cat(sprintf(
    "%s: OR = %.2f, 95%% CI [%.2f, %.2f], p = %.3f\n",
    var_name, odds_ratios_rf[i], ci_lower_rf[i], ci_upper_rf[i], pvalues_rf[i]
  ))
}
```

```
sch_typeUrban/Suburban: OR = 2.23, 95% CI [1.59, 3.14], p = 0.000
rf_time6:00 p.m.-11:59 p.m.: OR = 0.96, 95% CI [0.62, 1.49], p = 0.861
rf_timeMidnight-5:59 a.m.: OR = 0.57, 95% CI [0.24, 1.38], p = 0.214
rf_timeNoon-6:00 p.m.: OR = 0.66, 95% CI [0.44, 1.00], p = 0.050
rf_length30-59 minutes: OR = 0.55, 95% CI [0.33, 0.91], p = 0.020
rf_length60-89 minutes: OR = 0.46, 95% CI [0.26, 0.82], p = 0.008
rf_length90-119 minutes: OR = 0.55, 95% CI [0.25, 1.18], p = 0.124
rf_length2 hours or more: OR = 0.51, 95% CI [0.23, 1.11], p = 0.088
rf_tv1: OR = 1.04, 95% CI [0.49, 2.17], p = 0.924
rf_tv2: OR = 1.43, 95% CI [0.71, 2.88], p = 0.321
rf_tv3: OR = 2.61, 95% CI [1.18, 5.79], p = 0.018
rf_music1: OR = 0.64, 95% CI [0.28, 1.48], p = 0.302
rf_music2: OR = 0.56, 95% CI [0.25, 1.25], p = 0.156
rf_music3: OR = 0.77, 95% CI [0.31, 1.92], p = 0.579
rf_pd1: OR = 1.32, 95% CI [0.55, 3.17], p = 0.533
rf_pd2: OR = 2.09, 95% CI [0.89, 4.93], p = 0.092
rf_pd3: OR = 1.48, 95% CI [0.59, 3.70], p = 0.398
rf_talk_phone1: OR = 1.45, 95% CI [0.71, 2.94], p = 0.304
rf_talk_phone2: OR = 2.09, 95% CI [1.02, 4.27], p = 0.044
rf_talk_phone3: OR = 1.56, 95% CI [0.67, 3.63], p = 0.306
rf_onl_game1: OR = 0.69, 95% CI [0.38, 1.25], p = 0.220
rf_onl_game2: OR = 1.00, 95% CI [0.56, 1.81], p = 0.990
rf_onl_game3: OR = 0.92, 95% CI [0.39, 2.17], p = 0.842
rf_soc_network1: OR = 0.36, 95% CI [0.17, 0.73], p = 0.005
rf_soc_network2: OR = 0.44, 95% CI [0.21, 0.89], p = 0.023
rf_soc_network3: OR = 0.41, 95% CI [0.18, 0.92], p = 0.030
No, not at all|Yes, some: OR = 0.31, 95% CI [0.08, 1.13], p = 0.076
Yes, some|Yes, a lot: OR = 4.00, 95% CI [1.08, 14.79], p = 0.038
```

Likelihood Ratio Test to Compare Models

```
############################################################
# 6. Likelihood Ratio Test to Compare Models
############################################################

# First convert the models to their formulas
ra_full_formula <- formula(model_read)
ra_base_formula <- ra_disp ~ experience + sch_type

# Get the complete data with no missing values for all variables in the full model
all_vars_ra <- all.vars(ra_full_formula)
complete_data_ra <- na.omit(read_data[, all_vars_ra])

# Now fit both models on this complete dataset
ra_base_new <- polr(ra_base_formula, data = complete_data_ra)
ra_full_new <- polr(ra_full_formula, data = complete_data_ra)

# Compare models
lrt_ra <- anova(ra_base_new, ra_full_new, test = "Chisq")
print("Likelihood Ratio Test for Academic Reading Models:")
```

```
[1] "Likelihood Ratio Test for Academic Reading Models:"
```

```
print(lrt_ra)
```

```
Likelihood ratio tests of ordinal regression models

Response: ra_disp

Model
1                                                                                                  expe
rience + sch_type
2 experience + sch_type + ra_time + ra_length + ra_tv + ra_music + ra_talk_phone + ra_video_game
+ ra_soc_network
  Resid. df Resid. Dev   Test    Df LR stat.      Pr(Chi)
1       601   1228.614
2       579   1172.276 1 vs 2    22 56.33769 7.647704e-05
```

```r
# Similar approach for recreational reading
rf_full_formula <- formula(model_final_read)
rf_base_formula <- rf_disp ~ sch_type

# Get complete data
all_vars_rf <- all.vars(rf_full_formula)
complete_data_rf <- na.omit(read_data[, all_vars_rf])

# Fit both models on complete data
rf_base_new <- polr(rf_base_formula, data = complete_data_rf)
rf_full_new <- polr(rf_full_formula, data = complete_data_rf)

# Compare models
lrt_rf <- anova(rf_base_new, rf_full_new, test = "Chisq")
print("Likelihood Ratio Test for Recreational Reading Models:")
```

```
[1] "Likelihood Ratio Test for Recreational Reading Models:"
```

```r
print(lrt_rf)
```

```
Likelihood ratio tests of ordinal regression models

Response: rf_disp

Model
1
sch_type
2 sch_type + rf_time + rf_length + rf_tv + rf_music + rf_pd + rf_talk_phone + rf_onl_game + rf_s
oc_network
  Resid. df Resid. Dev   Test    Df LR stat.      Pr(Chi)
1       590   1186.131
2       565   1121.490 1 vs 2    25  64.6403 2.329237e-05
```

Cohen’s d Effect Size Calculation

```
############################################################
# 7. Effect Size Calculations (Cohen's d approximation)
############################################################

# Function to calculate standardized mean difference for categorical predictors
calculate_effect_size <- function(data, predictor, outcome) {
  # Create contingency table
  cont_table <- table(data[[predictor]], data[[outcome]])

  # Get proportions for each level of predictor
  props <- prop.table(cont_table, margin = 1)

  # Calculate weighted average for higher categories
  level_props <- list()
  for (i in 1:nrow(props)) {
    # Sum proportions for "Yes, some" and "Yes, a lot"
    level_props[[i]] <- sum(props[i, 2:3])
  }

  # Calculate Cohen's d (for binary predictors)
  if (length(level_props) == 2) {
    p1 <- level_props[[1]]
    p2 <- level_props[[2]]
    # Pooled standard deviation
    sd_pooled <- sqrt(p1 * (1 - p1) + p2 * (1 - p2))
    # Effect size
    d <- (p2 - p1) / sd_pooled
    return(d)
  } else {
    return(NA) # Return NA for non-binary predictors
  }
}

# Calculate effect sizes for key binary predictors
# School type effect on academic reading
d_sch_ra <- calculate_effect_size(read_data, "sch_type", "ra_disp")
print(paste("Cohen's d for school type on academic reading:", round(d_sch_ra, 2)))
```

```
[1] "Cohen's d for school type on academic reading: 0.07"
```

```
# School type effect on recreational reading
d_sch_rf <- calculate_effect_size(read_data, "sch_type", "rf_disp")
print(paste("Cohen's d for school type on recreational reading:", round(d_sch_rf, 2)))
```

```
[1] "Cohen's d for school type on recreational reading: 0.22"
```

```
# For experience, compare novice (0-5 years) vs. experienced (11+ years)
# First create binary experience variable
read_data$exp_binary <- factor(ifelse(read_data$experience == "0-5 years", "Novice",
  ifelse(read_data$experience == "11-years and more", "Experienced", NA)
))

# Calculate effect sizes
d_exp_ra <- calculate_effect_size(na.omit(read_data[, c("exp_binary", "ra_disp")]), "exp_binar
y", "ra_disp")
print(paste("Cohen's d for experience (novice vs. experienced) on academic reading:", round(d_ex
p_ra, 2)))
```

```
[1] "Cohen's d for experience (novice vs. experienced) on academic reading: 0.07"
```

```
d_exp_rf <- calculate_effect_size(na.omit(read_data[, c("exp_binary", "rf_disp")]), "exp_binar
y", "rf_disp")
print(paste("Cohen's d for experience (novice vs. experienced) on recreational reading:", round
(d_exp_rf, 2)))
```

```
[1] "Cohen's d for experience (novice vs. experienced) on recreational reading: 0.11"
```

Prediction Accuracy (Classification Tables)

```
############################################################
# 8. Prediction Accuracy and Classification Tables
############################################################

# Function to create classification table and calculate accuracy
create_classification_table <- function(model, data) {
  # Get predictions
  predictions <- predict(model, newdata = data)

  # Get actual values (remove NAs)
  actual <- na.omit(data[[all.vars(formula(model))[1]]])
  predicted <- na.omit(predictions)

  # Ensure lengths match
  min_length <- min(length(actual), length(predicted))
  actual <- actual[1:min_length]
  predicted <- predicted[1:min_length]

  # Create classification table
  class_table <- table(Actual = actual, Predicted = predicted)

  # Calculate accuracy
  accuracy <- sum(diag(class_table)) / sum(class_table)

  # Return results
  return(list(table = class_table, accuracy = accuracy))
}

# Academic reading model
ra_classification <- create_classification_table(model_read, read_data)
print("Classification Table for Academic Reading Model:")
```

```
[1] "Classification Table for Academic Reading Model:"
```

```
print(ra_classification$table)
```

```
               Predicted
Actual          No, not at all Yes, some Yes, a lot
  No, not at all             25       158          2
  Yes, some                  35       279          4
  Yes, a lot                  7       115         14
```

```
print(paste("Accuracy:", round(ra_classification$accuracy * 100, 1), "%"))
```

```
[1] "Accuracy: 49.8 %"
```

```
# Recreational reading model
rf_classification <- create_classification_table(model_final_read, read_data)
print("Classification Table for Recreational Reading Model:")
```

```
[1] "Classification Table for Recreational Reading Model:"
```

```
print(rf_classification$table)
```

```
                Predicted
Actual           No, not at all Yes, some Yes, a lot
  No, not at all             38       158          1
  Yes, some                  50       262          5
  Yes, a lot                 21        88          9
```

```
print(paste("Accuracy:", round(rf_classification$accuracy * 100, 1), "%"))
```

```
[1] "Accuracy: 48.9 %"
```

Interaction Effects Analysis

```
############################################################
# 9. Interaction Effects
############################################################

# Test interaction between experience and school type
# For academic reading
model_int_ra <- polr(ra_disp ~ experience * sch_type + ra_time + ra_length +
  ra_tv + ra_music + ra_talk_phone + ra_video_game +
  ra_soc_network, data = read_data)

# Compare models
compare_int_ra <- anova(model_read, model_int_ra, test = "Chisq")
print("Test of interaction effect (experience * school type) for academic reading:")
```

```
[1] "Test of interaction effect (experience * school type) for academic reading:"
```

```
print(compare_int_ra)
```

```
Likelihood ratio tests of ordinal regression models

Response: ra_disp

Model
1 experience + sch_type + ra_time + ra_length + ra_tv + ra_music + ra_talk_phone + ra_video_game
+ ra_soc_network
2 experience * sch_type + ra_time + ra_length + ra_tv + ra_music + ra_talk_phone + ra_video_game
+ ra_soc_network
  Resid. df Resid. Dev   Test    Df LR stat.    Pr(Chi)
1       579   1172.276
2       577   1164.623 1 vs 2     2 7.653062 0.02178505
```

```
# For recreational reading
model_int_rf <- polr(rf_disp ~ sch_type * rf_tv + rf_time + rf_length +
  rf_music + rf_pd + rf_talk_phone + rf_onl_game +
  rf_soc_network, data = read_data)

# Compare models
compare_int_rf <- anova(model_final_read, model_int_rf, test = "Chisq")
print("Test of interaction effect (school type * TV watching) for recreational reading:")
```

```
[1] "Test of interaction effect (school type * TV watching) for recreational reading:"
```

```
print(compare_int_rf)
```

```
Likelihood ratio tests of ordinal regression models

Response: rf_disp

Model
1 sch_type + rf_time + rf_length + rf_tv + rf_music + rf_pd + rf_talk_phone + rf_onl_game + rf_s
oc_network
2 sch_type * rf_tv + rf_time + rf_length + rf_music + rf_pd + rf_talk_phone + rf_onl_game + rf_s
oc_network
  Resid. df Resid. Dev   Test    Df LR stat.   Pr(Chi)
1       565   1121.490
2       562   1119.132 1 vs 2     3 2.358634 0.5013813
```

```
# Create cross-tabulation to visualize interaction patterns
exp_sch_table <- with(read_data, table(experience, sch_type, ra_disp))
print("Cross-tabulation of experience, school type, and academic reading displacement:")
```

```
[1] "Cross-tabulation of experience, school type, and academic reading displacement:"
```

```
print(exp_sch_table)
```

```
, , ra_disp = No, not at all

                   sch_type
experience          Rural Urban/Suburban
  0-5 years            25             44
  6-10 years           42             29
  11-years and more    10             30

, , ra_disp = Yes, some

                   sch_type
experience          Rural Urban/Suburban
  0-5 years            63             78
  6-10 years           46             55
  11-years and more    24             52

, , ra_disp = Yes, a lot

                   sch_type
experience          Rural Urban/Suburban
  0-5 years            21             54
  6-10 years           11             25
  11-years and more     7             18
```

Model Comparison Across Reading Types

```
############################################################
# 10. Comparison of Academic vs. Recreational Reading Models
############################################################

# Identify common predictors in both models
# Here we'll use school type and TV watching as they appear in both models
common_formula_ra <- ra_disp ~ sch_type + ra_tv
common_formula_rf <- rf_disp ~ sch_type + rf_tv

# Fit models with only common predictors
common_model_ra <- polr(common_formula_ra, data = read_data)
common_model_rf <- polr(common_formula_rf, data = read_data)

# Extract coefficients and compare
coef_ra <- coef(common_model_ra)
coef_rf <- coef(common_model_rf)

# Calculate odds ratios
or_ra <- exp(coef_ra)
or_rf <- exp(coef_rf)

# Create comparison table
predictor_names <- c("School Type (Urban/Suburban)", "TV Use (A little)", "TV Use (Some)", "TV U
se (Most)")
comparison_table <- data.frame(
  Predictor = predictor_names,
  Academic_OR = c(or_ra["sch_typeUrban/Suburban"], or_ra["ra_tv1"], or_ra["ra_tv2"], or_ra["ra_t
v3"]),
  Recreational_OR = c(or_rf["sch_typeUrban/Suburban"], or_rf["rf_tv1"], or_rf["rf_tv2"], or_rf
["rf_tv3"]),
  Ratio = c(
    or_rf["sch_typeUrban/Suburban"] / or_ra["sch_typeUrban/Suburban"],
    or_rf["rf_tv1"] / or_ra["ra_tv1"],
    or_rf["rf_tv2"] / or_ra["ra_tv2"],
    or_rf["rf_tv3"] / or_ra["ra_tv3"]
  )
)

print("Comparison of Effects Across Academic and Recreational Reading:")
```

```
[1] "Comparison of Effects Across Academic and Recreational Reading:"
```

```
print(comparison_table)
```

```
                                      Predictor Academic_OR
sch_typeUrban/Suburban School Type (Urban/Suburban)   1.5310506
ra_tv1                              TV Use (A little)   0.7309686
ra_tv2                                  TV Use (Some)   0.8670846
ra_tv3                                  TV Use (Most)   1.9938460
                        Recreational_OR    Ratio
sch_typeUrban/Suburban        2.178008 1.422558
ra_tv1                        0.863182 1.180874
ra_tv2                        1.403792 1.618980
ra_tv3                        2.546167 1.277013
```

```
# Calculate differential impact (relative effect sizes)
print("Relative impact (recreational vs. academic reading):")
```

```
[1] "Relative impact (recreational vs. academic reading):"
```

```
for (i in 1:nrow(comparison_table)) {
  cat(sprintf(
    "%s: Effect is %.2f times stronger for %s reading\n",
    comparison_table$Predictor[i],
    abs(comparison_table$Ratio[i]),
    ifelse(comparison_table$Ratio[i] > 1, "recreational", "academic")
  ))
}
```

```
School Type (Urban/Suburban): Effect is 1.42 times stronger for recreational reading
TV Use (A little): Effect is 1.18 times stronger for recreational reading
TV Use (Some): Effect is 1.62 times stronger for recreational reading
TV Use (Most): Effect is 1.28 times stronger for recreational reading
```