Digital media supporting literacy learning in children with communicative and cognitive disabilities

27 Oktober, 2022

(ref:abstract)

(ref:abstract) **Background**: BLALBLALALBLABLABLABLALBALBLABLABLALB **Method**: One hundred and thirty-three adolescents with ID were recruited. The participants were instructed to train in school for a total of 12 hours. **Results**: Blablablablalbalba **Conclusions**: blablalbalbalbal.

# Introduction

Children with intellectual and communicative disabilities may be among the most socially disadvantaged student categories. Failing in acquiring literacy skills is a serious and common problem. The purpose of this project is to investigate the effects of two specific digital literacy learning tools (apps) on the acquisition of literacy skills in children with intellectual disability and who are reliant on augmentative and alternative communication, AAC. This means that they need aids such as manual signing, symbols/pictures, speech-generating devices, etc. to understand and/or express themselves.

We will recruit children with intellectual disability of different etiologies from special needs schools. The age range will be broad. The children will be assigned to one of three interventions or to a comparison group that will receive teaching-as-usual (TaU). The three interventions are: 1) the ALL-Accessible Literacy Learning app (phonemic strategies); the Animega-is app (comprehension-based strategies); 3) a combination of both ALL and Animega-is. The children’s literacy development during a 12-week intervention period will be compared across the four different groups. In line with a previous study targeting children in reading difficulties (Gustafson et al., 2011), we expect a positive effect on reading and phonological awareness across all intervention groups, with the largest effect in the group receiving the combined intervention. Hypotheses

## Hypotheses

1: Training phonemic or comprehension-based reading strategies improves phonological awareness. 2: Training phonemic or comprehension-based reading strategies improves reading ability. 3: The combined training is more effective than either intervention on its own.

The hypotheses will be tested on these outcome variables: phonological awareness (1, 3), word reading (2, 3), and sentence reading (2, 3).

# Methods

## Participants

Participants were recruited from special needs schools in southern Sweden, after consent of the school’s principal. Participating children gave oral consent at the beginning of the first test session and prior to the study, caregivers signed an informed consent, regarding the confidentiality of data and group-level analysis. The study follows the Ethical principles for medical research involving human subjects from the WMA Declaration of Helsinki (World Medical Association 2013). The study was reviewed and approved by the Ethical Review Board, Sweden (*insert dr no*).

In Sweden, special schools are open for children with an ID diagnosis^1. The recruitment process resulted in 137 participants ( = 58, = 79). However, one participant was excluded from the study due to testing not being followed as per protocol, and the final sample thus included 133 participants (for demographics, see Table …). Data on diagnoses were collected using parental surveys. The diagnoses in the ID group can be seen in Table @ref(tab:diagnosis-table).

##Behavioral measures  
Non-verbal intelligence and reading measures were collected.

###Non-verbal intelligence  
The participants’ IQ were calculated based on The Raven’s 2 Progressive Matrices Clinical Edition (Raven’s 2; *Raven, Rust, Chan, & Zhou, 2018*).

###Phonological awareness  
Three subtests from MiniDUVAN (Wolff, 2013) were used to assess phonological awareness skills: A2 Rhyme identification, A3 Phoneme identification, and B4 Phoneme synthesis. The dependent measure was total number of correct answers across all subtests (max = 45).

###Letter-sound knowledge  
To be added…

###Word reading  
OS64 (Nielsen et al., 1997) and OLAF (Magnusson & Naucler, 2010) were used to assess word reading skills. However, the presentation of tasks, and response modes, were adapted to suit the participants in this project. In OS64, … matched a written word to a widgit symbol. In OLAF, written words were matched to pictures. The dependent measure was the total number of correct answers (max = 15 for OS64, and max = 20 for OLAF).

###Sentence reading  
In DLS Bas (Järpsten, 2004) the participants read short sentences and match these to their corresponding pictures. The dependent measure was the total number of correct answers (max = 20).

## Instruction materials: ALL and Animega-is

## Procedure

The training took place in the participants’ school. The participants trained in a group in schools and the teachers were instructed to allow the students to train for 300 minutes (20 sessions for 15 min, five days a week for four weeks). The teachers were asked to let the participants train by themselves and to not assist in solving the tasks for the children. No verbal or written instructions were given to the participants. A research group member attended the first training session to instruct how to operate the tablet and program. All included participants that attended the same class also trained in the same room. However, in some cases only one student in the class participated in the study. Thus, sometimes the participants trained in a group and sometimes the participant sat by themselves, but the participant always trained on their own.

### Time

The children were tested on four occasions, before the intervention, half way through the intervention, right after the intervention, and at a six week follow-up after the intervention had stopped. Due to practical and external reasons (e.g. restrictions and sick leave due to the COVID-19 pandemic), the testing could not be done with the same intervals for all children. Thus, time was coded as an interval variable rather than a categorical, using days and the first testing time set to 0.

### Training time

Training time was extracted from time spent training in the program and was measured in minutes.

## Data analysis

The -value was set to 0.05.

For all our analyses, we used R (Version 4.1.2; R Core Team 2018) and the R-packages *dplyr* (Version 1.0.7; Wickham et al. 2018), *emmeans* (Version 1.7.0; **R-emmeans?**), *FSA* (Version 0.9.1; **R-FSA?**), *ggplot2* (Version 3.3.5; Wickham 2016), *ggpubr* (Version 0.4.0; **R-ggpubr?**), *lme4* (Version 1.1.27.1; Bates et al. 2015), *lmerTest* (Version 3.1.3; **R-lmerTest?**), *Matrix* (Version 1.3.4; Bates and Maechler 2018), *papaja* (Version 0.1.0.9997; Aust and Barth 2018), *rio* (Version 0.5.27; **R-rio?**), *sjPlot* (Version 2.8.9; **R-sjPlot?**), *sjstats* (Version 0.18.1; **R-sjstats?**), *tidyr* (Version 1.1.4; Wickham and Henry 2018), and *tinylabels* (Version 0.2.1; **R-tinylabels?**). DETTA MÅSTE KOMPLETTERAS SÅ ATT RÄTT REFERENSER KOMMER MED.

### Statistical analysis

In the preregistration, it was stated that mixed ANOVAs was going to be used. However, due to the children not being tested with the same time intervals, missing data, and the groups not being matched on IQ, linear mixed-effects models with repeated measures was used to analyze the effects of the interventions. Linear mixed-effects models are superior to ANOVA when dealing with missing data, and the difference in IQ and time intervals for testing ((**kalla?**)). Models were fitted with the lme4 package ((**kalla?**)) in R using using maximum likelihood and missing data was handled under the less restrictive assumption of missing at random. The assumption of linearity was tested by plotting the model-predicted values to the observed ones, homogeneity of variance was tested by plotting the residuals vs. fitted values. To check that the residuals of the model were normally distributed using a QQ plot. For the sentence reading, the residuals were non-normally distributed, thus a Generalized linear mixed-model with a Poisson distribution was used rather than a Gaussian distribution.

### Procedure modelling building

The effects of the interventions were evaluated on the four different outcome measures separately. The outcome measures were phonological awareness (PA), word reading, sentence reading, and letter-sound recognition. To evaluate the effects of the interventions, we first build an unconditional model (Model 1) with time as a fixed effect and participants as a random intercept. Thereafter, we added time as a random intercept (Model 2) and compared it to Model 1. If Model 2 were significantly better than Model 1, we included time as a random intercept, otherwise time was only included as a fixed effect. To test Hypothesis 1 and 2 respectively, we build a conditional model, Model 3, were we added intervention as a binary variable (0 = comparison group, 1 = intervention groups). Thereafter, we added IQ as a control variable in Model 4 to see if it interacted with the effect of the intervention. Only Model 3 and Model 4 will be presented in the result section, the other models can be found in the supplements. Thereafter, we investigated Hypothesis 3, if the combined training was more effective than the ALL and Animega-is on their own. The same procedure of model building was repeated for this hypothesis testing.

### Covariance structure

Random effects were fitted with both no covariance structure (variance components) and an unstructured variance-covariance matrix. The unstructured covariance-matrix was used if it significantly improved the model, otherwise the less complex (?), no covariance structure was used.

### Model comparison

Models were compared using an ANOVA (or loglikelihoodtest och chi2, tror det är samma måste jämföra med Emil) when the same amount of parameters were estimated in the models and by using AIC scores when different parameters were used (ELLER HUR SKA VI GÖRA?).

# Results

The descriptive statistics for all variables on each assessment time can be seen in Table @ref(tab:descriptives-table).

(#tab:descriptives-table)

*Descriptive statistics of included variables presented by group*

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Age | 13.8 | 3.04 | (9, 19) | 15.3 | 3.28 | (8, 22) | 12.4 | 3.47 | (7, 20) | 13.6 | 2.84 | (8, 19) |
| IQ | 44.6 | 7.71 | (40, 70) | 44.8 | 7.89 | (40, 74) | 50.7 | 12.54 | (40, 78) | 47.4 | 12.19 | (40, 79) |
| Trained time | 0 | 0 | (0, 0) | 413 | 226 | (83, 907) | 385 | 188 | (77, 760) | 366 | 180 | (85, 860) |

*Note.* Chronological is are presented in years.

(#tab:desc-read-control-table)

*Descriptive statistics for the control group of included variables presented by time*

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Letter | 6.33 | 2.56 | (0, 8) | 6.32 | 2.50 | (0, 8) | 6.42 | 1.92 | (3, 8) | 5.78 | 3.25 | (0, 8) |
| Word | -0.4348 | 0.771 | (-1, 2) | -0.0643 | 0.987 | (-1, 2) | -0.1350 | 0.851 | (-1, 1) | -0.2063 | 0.860 | (-1, 1) |
| Sentence | 0.233 | 0.568 | (0, 2) | 0.640 | 1.350 | (0, 4) | 1.000 | 3.232 | (0, 14) | 0.630 | 2.133 | (0, 11) |
| PA | 10.8 | 8.90 | (0, 26) | 11.1 | 8.69 | (0, 25) | 11.6 | 9.51 | (0, 25) | 11.7 | 8.40 | (0, 27) |

*Note.* this is a note.

(#tab:desc-read-ALL-table)

*Descriptive statistics for the ALL group of included variables presented by time*

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Letter | 6.00 | 2.54 | (0, 8) | 6.94 | 1.93 | (0, 8) | 6.68 | 2.41 | (0, 8) | 7.41 | 1.01 | (4, 8) |
| Word | -0.3081 | 0.672 | (-1, 1) | -0.1569 | 0.882 | (-1, 2) | -0.0522 | 0.746 | (-1, 1) | 0.1675 | 0.934 | (-1, 2) |
| Sentence | 0.344 | 1.29 | (0, 7) | 1.233 | 2.92 | (0, 14) | 1.000 | 2.82 | (0, 15) | 2.129 | 3.85 | (0, 16) |
| PA | 12.8 | 8.03 | (0, 25) | 15.2 | 8.03 | (0, 26) | 14.5 | 8.26 | (0, 26) | 15.8 | 7.48 | (0, 27) |

*Note.* this is a note.

(#tab:desc-read-animega-table)

*Descriptive statistics for the Animega-is group of included variables presented by time*

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Letter | 6.03 | 2.75 | (0, 8) | 6.72 | 2.29 | (0, 8) | 6.93 | 2.00 | (0, 8) | 6.70 | 2.25 | (2, 8) |
| Word | -0.0438 | 0.949 | (-1, 2) | 0.0614 | 1.009 | (-1, 2) | 0.2478 | 1.035 | (-1, 2) | 0.3784 | 1.008 | (-1, 2) |
| Sentence | 1.26 | 3.85 | (0, 18) | 1.53 | 3.30 | (0, 15) | 1.63 | 4.09 | (0, 18) | 3.27 | 5.64 | (0, 20) |
| PA | 12.0 | 7.12 | (0, 27) | 12.7 | 8.24 | (0, 26) | 13.3 | 7.59 | (1, 27) | 14.3 | 8.28 | (0, 27) |

*Note.* this is a note.

(#tab:desc-read-combi-table)

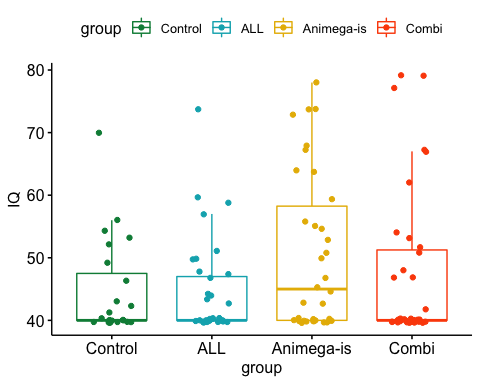
*Descriptive statistics for the Combi group of included variables presented by time*

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Letter | 6.34 | 2.36 | (0, 8) | 6.53 | 2.31 | (0, 8) | 6.76 | 1.84 | (2, 8) | 6.67 | 1.83 | (2, 8) |
| Word | -0.0758 | 0.709 | (-1, 1) | 0.2096 | 0.905 | (-1, 2) | 0.1235 | 0.875 | (-1, 2) | 0.1971 | 1.066 | (-1, 3) |
| Sentence | 0.257 | 0.886 | (0, 4) | 0.824 | 1.487 | (0, 7) | 0.625 | 1.879 | (0, 8) | 0.467 | 0.860 | (0, 3) |
| PA | 12.9 | 8.33 | (0, 24) | 14.5 | 7.88 | (0, 25) | 16.0 | 6.85 | (0, 26) | 16.5 | 6.72 | (1, 26) |

*Note.* this is a note.

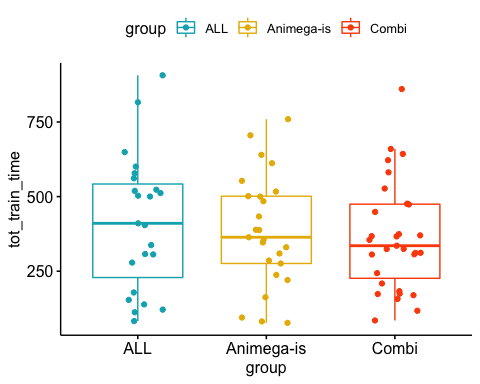
## IQ

The participants were randomly selected on school level to different intervention groups. This was done to try to match the groups on IQ. Nevertheless, the IQ in the groups differed significantly, *H*(3) = 19.26, *p* <.001. Facon, Magis, and Belmont (2011) proposed groups be equated with an α-level of *p* > 0.50. The Animega-is group had a significantly higher IQ ( *m* =50.706; *sd* =12.537) compared to the Comparison group ( *m* =44.609; *sd* =7.709), *p* =.002, the ALL group ( *m* =44.758; *sd* =7.89),  
*p* <.001, and the combination group ( *m* =47.361; *sd* =12.187), *p* =.0018. Thus, IQ was used as a control variable in all outcome variables.



## Training time

The participants were instructed to train for a total of 18 hours (36 sessions for 30 minutes, three days a week for a 12-week period). The teachers were instructed to adapt the length of the training sessions after the need of the student. That is, if the student could not sit for a full 30 minute session the teacher was encouraged to have shorter sessions but more often than three times a week. On average the participants trained for 385.609 (*sd* =195.263) minutes (6.427 (*sd* =3.254) hours). There was no difference on training time between the three intervention groups (*H*(2) = 0.632, *p* =.729 ).



## Phonological awareness

### Hypothesis 1a: Training phonemic reading strategies improves phonological awareness.

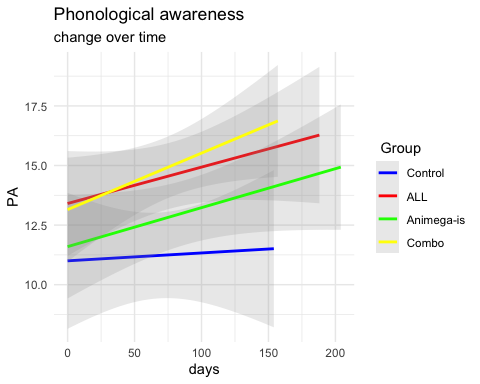
There was almost a significant interaction between time and the intervention on phonemic reading, the group training phonemic strategies improved more than the comparison group (, 95% CI , , ).

### Hypothesis 1b: Training comprehension-based reading strategies improves phonological awareness.

There was no significant interaction between time and intervention for the comprehension-based reading strategy on PA (, 95% CI , , ).

### Hypothesis 3: The combined training is more effective than either intervention on its own.

There was a significant interaction between the combined group and the other two intervention groups, (, 95% CI , , ), the combined training improved PA over time more than the other two intervention groups.



## Exploratory analyses

### Effect of IQ

There was a significant main effect on IQ (, 95% CI , , ), indicating that higher IQ results in higher PA at T1. Moreover, there was a significant interaction between time and IQ (, 95% CI , , ), indicating that the participants with the highest IQ benefited the most from the intervention. However, the benefit of the combined training over the two other intervention types remained (, 95% CI , , )

The results from the PA models can be seen in Table @ref(tab:PA-table).

PA

Predictors

Estimates

CI

p

df

(Intercept)

12.66

11.32 – 13.99

<0.001

458.00

days

0.01

0.01 – 0.02

<0.001

458.00

contrast 1vs2

1.64

-0.56 – 3.84

0.143

458.00

contrast 1vs3

-1.17

-3.41 – 1.08

0.307

458.00

contrast 4vs23

0.52

-1.68 – 2.72

0.645

458.00

IQ scale

2.52

0.91 – 4.13

0.002

458.00

days \* contrast 1vs2

0.01

0.00 – 0.02

0.046

458.00

days \* contrast 1vs3

0.00

-0.01 – 0.01

0.711

458.00

days \* contrast 4vs23

0.01

0.00 – 0.02

0.006

458.00

days \* IQ scale

0.01

-0.00 – 0.01

0.069

458.00

Random Effects

σ2

12.15

τ00 id

46.48

ICC

0.79

N id

125

Observations

470

Marginal R2 / Conditional R2

0.122 / 0.818

## Word reading

### Hypothesis 1a: Training phonemic reading strategies improves word reading.

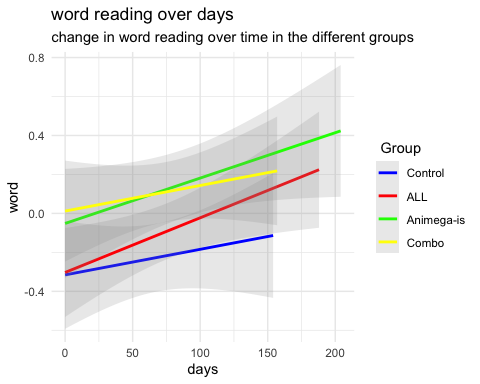
There was not a significant interaction between time and the intervention on phonemic-based reading (, 95% CI , , ).

### Hypothesis 1b: Training comprehension-based reading strategies improves word awareness.

There was no significant interaction between time and intervention for the comprehension-based reading strategy on word reading (, 95% CI , , ).

### Hypothesis 3: The combined training is more effective than either intervention on its own.

There was a not significant interaction between the combined group and the other two intervention groups, (, 95% CI , , ).



## Exploratory analyses

### Effect of IQ

There was a significant main effect on IQ (, 95% CI , , ), indicating that higher IQ results in higher word reading at T1. Moreover, there was a significant interaction between time and IQ (, 95% CI , , ), indicating that the participants with the highest IQ improved most over the training period regardless of intervention condition (including control).

The results from the word models can be seen in Table @ref(tab:word-table).

word

word

Predictors

Estimates

CI

p

df

Estimates

CI

p

df

(Intercept)

-0.16

-0.30 – -0.02

0.029

452.00

-0.12

-0.24 – -0.00

0.046

452.00

days

0.20

0.13 – 0.27

<0.001

452.00

0.13

0.07 – 0.20

<0.001

452.00

contrast 1vs2

-0.01

-0.24 – 0.22

0.920

452.00

-0.08

-0.27 – 0.12

0.443

452.00

contrast 1vs3

0.11

-0.12 – 0.35

0.347

452.00

0.16

-0.03 – 0.36

0.106

452.00

contrast 4vs23

0.16

-0.07 – 0.39

0.177

452.00

0.14

-0.05 – 0.33

0.146

452.00

IQ scale

0.25

0.08 – 0.42

0.004

452.00

0.15

0.01 – 0.30

0.035

452.00

days \* contrast 1vs2

0.07

-0.04 – 0.18

0.197

452.00

0.01

-0.09 – 0.11

0.887

452.00

days \* contrast 1vs3

-0.03

-0.13 – 0.08

0.613

452.00

-0.03

-0.13 – 0.07

0.529

452.00

days \* contrast 4vs23

-0.01

-0.12 – 0.11

0.903

452.00

-0.08

-0.19 – 0.02

0.129

452.00

days \* IQ scale

0.06

-0.02 – 0.14

0.140

452.00

0.00

-0.08 – 0.08

0.962

452.00

PA

0.29

0.21 – 0.38

<0.001

452.00

days \* PA

0.10

0.04 – 0.16

0.002

452.00

Random Effects

σ2

0.15

0.16

τ00

0.49 id

0.30 id

τ11

0.02 id.scale(days, center = FALSE)

ρ01

0.51 id

ICC

0.80

0.66

N

124 id

124 id

Observations

466

466

Marginal R2 / Conditional R2

0.108 / 0.820

0.317 / 0.765

# Sentence reading

### Hypothesis 1a: Training phonemic reading strategies improves sentence reading.

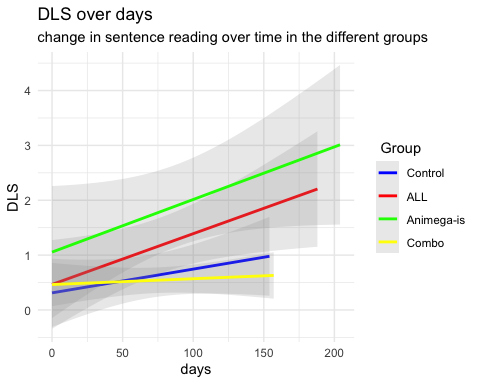
There was not a significant interaction between time and the intervention on phonemic training on sentence reading ().

### Hypothesis 1b: Training comprehension-based reading strategies improves sentence awareness.

There was no significant interaction between time and intervention for the comprehension-based reading strategy on sentence reading ().

### Hypothesis 3: The combined training is more effective than either intervention on its own.

There was a not significant interaction between the combined group and the other two intervention groups, ().



## Exploratory analyses

### Effect of IQ

There was a significant main effect on IQ (, 95% CI , , ), indicating that higher IQ results in higher sentence reading at T1. Moreover, there was a significant interaction between time and IQ (, 95% CI , , ), indicating that the participants with the highest IQ improved most over the training period regardless of intervention condition (including control).

The results from the word models can be seen in Table @ref(tab:DLS-table).

DLS

Predictors

Incidence Rate Ratios

CI

p

(Intercept)

0.04

0.02 – 0.12

<0.001

IQ scale

2.52

1.24 – 5.10

0.010

days

2.88

1.52 – 5.45

0.001

IQ scale \* days

1.11

0.82 – 1.49

0.511

Random Effects

σ2

2.50

τ00 id

7.02

τ11 id.scale(days, center = FALSE)

0.45

ρ01 id

-0.68

ICC

0.69

N id

124

Observations

465

Marginal R2 / Conditional R2

0.122 / 0.728

# Explorative analyses

## Letter-sound recognition

### Hypothesis 1a: Training phonemic reading strategies improves letter-sound recognition.

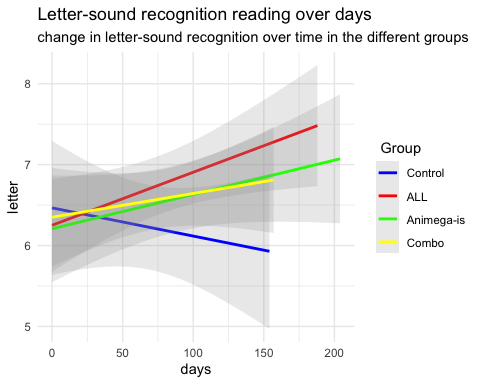
There was a significant interaction between time and the intervention on phonemic training on letter-sound recognition (, 95% CI , , ).

### Hypothesis 1b: Training comprehension-based reading strategies improves sletter-sound recognition.

There was no significant interaction between time and intervention for the comprehension-based reading strategy on letter-sound recognition (, 95% CI , , ).

### Hypothesis 3: The combined training is more effective than either intervention on its own.

There was a not significant interaction between the combined group and the other two intervention groups, (, 95% CI , , ).



### Effect of IQ

There was a significant main effect on IQ (, 95% CI , , ), indicating that higher IQ results in higher sentence reading at T1. However, there was not a significant interaction between time and IQ (, 95% CI , , ), indicating that the participants’ IQ did not affect the improvement rate over the training period.

The results from the word models can be seen in Table @ref(tab:letter-table).

letter

Predictors

Estimates

CI

p

df

(Intercept)

6.34

5.90 – 6.77

<0.001

457.00

days

0.22

0.03 – 0.42

0.025

457.00

contrast 1vs2

0.04

-0.68 – 0.75

0.918

457.00

contrast 1vs3

-0.31

-1.04 – 0.42

0.411

457.00

contrast 4vs23

0.03

-0.69 – 0.74

0.942

457.00

IQ scale

0.66

0.13 – 1.18

0.014

457.00

days \* contrast 1vs2

0.40

0.09 – 0.72

0.012

457.00

days \* contrast 1vs3

0.11

-0.21 – 0.42

0.499

457.00

days \* contrast 4vs23

0.03

-0.30 – 0.37

0.839

457.00

days \* IQ scale

-0.17

-0.40 – 0.07

0.165

457.00

Random Effects

σ2

1.20

τ00 id

5.00

τ11 id.scale(days, center = FALSE)

0.27

ρ01 id

-0.76

ICC

0.76

N id

126

Observations

471

Marginal R2 / Conditional R2

0.050 / 0.775

# marginal and conditional R2

A marginal R2 close to zero tells us that the fixed effects aren’t explaining much variation, and a conditional R2 close to 1 tells us that most of that unexplained variation is between groups (people) rather than between observations within groups (people). So, for example, if the context was a longitudinal cohort study, we wouldn’t expect to improve our model much by collecting more data on characteristics/measures that mainly vary within people, but instead should find characteristics that mainly vary between people.

# Results summary

Combination group improves more on PA compared to the other intervention groups. ALL intervention improves letter-sound recognition more than other intervention groups. IQ influences improvement over time on many outcome measures (sammanfatta vilka här). PA interacts with time on letter-sound recognition, word reading and sentence reading, suggesting that PA predicts improvement in those outcome variables.

# Discussion

Blablablablalba

##Conclusions Blablabla

##Declaration of interest Blablabla

# References

Aust, Frederik, and Marius Barth. 2018. *papaja: Create APA Manuscripts with R Markdown*. <https://github.com/crsh/papaja>.

Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker. 2015. “Fitting Linear Mixed-Effects Models Using lme4.” *Journal of Statistical Software* 67 (1): 1–48. <https://doi.org/10.18637/jss.v067.i01>.

Bates, Douglas, and Martin Maechler. 2018. *Matrix: Sparse and Dense Matrix Classes and Methods*. <https://CRAN.R-project.org/package=Matrix>.

Facon, Bruno, David Magis, and John M. Belmont. 2011. “Beyond matching on the mean in developmental disabilities research.” *Research in Developmental Disabilities* 32 (6): 2134–47. <https://doi.org/10.1016/j.ridd.2011.07.029>.

R Core Team. 2018. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <https://ggplot2.tidyverse.org>.

Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller. 2018. *Dplyr: A Grammar of Data Manipulation*. <https://CRAN.R-project.org/package=dplyr>.

Wickham, Hadley, and Lionel Henry. 2018. *Tidyr: Easily Tidy Data with ’Spread()’ and ’Gather()’ Functions*. <https://CRAN.R-project.org/package=tidyr>.

World Medical Association. 2013. “World Medical Association Declaration of Helsinki.” *JAMA* 310 (20): 2191. <https://doi.org/10.1001/jama.2013.281053>.