A Randomized Trial of the Digital Down Syndrome LanguagePlus (DSL+) Vocabulary Intervention Program

Kari-Anne B. Næss¹. Silje Hokstad¹, Liv Inger Engevik¹,², Arne Lervåg³ and Elizabeth Smith⁴,⁵

¹University of Oslo
²Western Norway University of Applied Sciences
³University of Oslo
⁴University of Bristol
⁵Cardiff University

Author Note

Kari-Anne B. Næss, Department of Special Needs Education, University of Oslo, Mailbox 1140 Blindern, 0318 Oslo, Norway; Silje Hokstad, Department of Special Needs Education, University of Oslo, Mailbox 1140 Blindern, 0318 Oslo, Norway; Liv Inger Engevik, Department of Special Needs Education, University of Oslo, Mailbox 1140 Blindern, 0318 Oslo, Norway, Department of Pedagogy, Religion and Social Studies, Western Norway University of Applied Sciences, Mailbox 7030, 5020 Bergen, Norway; Arne Lervåg, Department of Education, University of Oslo, Norway; Elizabeth Smith, School of Psychology, Cardiff University, 70 Park Place, CF10 3AT, Cardiff, UK, School of Psychological Science, University of Bristol, BS8 1QU, Bristol, UK.

Acknowledgments: We thank the participating children and parents, schools, implementers, research assistants and people who have been involved in the development of the
DSL+ material, as well as the Department for Research and Dissemination Support, EngageLab, and University Centre for Information Technology at the University of Oslo. Additionally, we express great thanks to our colleagues Professor Christopher Jarrold, PhD, Kelly Burgoyne and Professor Charles Hulme for the important discussions related to the intervention development.

**Funding:** Grant number 238030 from the Finnut program of the Research Council of Norway.

**Conflict of Interest:** We have no conflicts of interest to disclose.

Correspondence to this article should be addressed to Kari-Anne B. Næss, Department of Special Needs Education, University of Oslo, Mailbox 1140 Blindern, 0318 Oslo, Norway. Telephone: 0047 922 40 741.
A Randomized Trial of the Digital Down Syndrome LanguagePlus (DSL+) Vocabulary Intervention Program

Abstract
This study investigated the effect of the digital Down Syndrome LanguagePlus (DSL+) intervention on vocabulary outcomes through a school-delivered randomized controlled trial. A national sample of first graders with Down syndrome from 91 schools was allocated to an intervention group (n=50), which received daily intervention for 15 weeks, or a business-as-usual control group (n=53). The intervention involved using picture book sharing and structured tasks and was organized as one-to-one, small-group and full-class lessons. Children in the intervention group made greater gains than children in the control group in expressive vocabulary breadth ($d = .429, CI [.160, .699]$) and receptive vocabulary breadth ($d = .447, CI [.193, .700]$). The outcomes indicate that the novel DSL+ intervention is an effective intervention to increase trained vocabulary among first graders with Down syndrome, and it takes only 15 minutes of effort five days a week.

Keywords: Down syndrome, vocabulary, language, intervention, RCT
A Randomized Trial of the Digital Down Syndrome LanguagePlus (DSL+) Vocabulary Intervention Program

Vocabulary is the heart of social communication and a catalyst for learning and interaction with peers (Biemiller, 2003; Bleses et al., 2016). Studies have shown that children with low vocabulary skills know fewer words and have less detailed semantic knowledge (Lawrence et al., 2012). These limitations may result in difficulties understanding and/or expressing themselves. Low vocabulary is suggested to be a bottleneck for the learning and development of basic skills such as language in general (Lewis, 1993), reading comprehension skills (e.g., Lervåg et al., 2018; Ricketts et al., 2019) and social functioning (Coplan & Armer, 2005; Dionne et al., 2003; Whitehouse et al., 2011). Therefore, long-term low vocabulary skills may result in social inequality, reduced occupational opportunities, and mental and physical health problems (Snow et al., 2016; Toppelberg et al., 2002). These negative developmental, psychological and behavioral consequences may critically affect both individuals and society.

Typically, children rapidly develop vocabulary from their first year of life through natural stimulation and interaction at home and in kindergarten, and subsequently, their vocabulary expands as they gradually increase their reading skills in school (Verhoeven et al., 2011). In contrast, children with Down syndrome usually have intellectual disability affecting their rate of development; their vocabulary learning is slow, and their skills appear to be stable over time (Næss et al., 2015).

This developmental profile of vocabulary weakness among children with Down syndrome leads to a vital need for effective vocabulary interventions. However, research concerning school practices (Næss et al., 2017) and a review by Smith et al. (2020) of earlier intervention research have shown a lack of research-based vocabulary interventions for children
with Down syndrome. Furthermore, according to Smith et al. (2020), no research-based digital vocabulary intervention for this group of children exists. However, a series of multiple baseline studies, e.g., Rivera et al. (2017), has shown promising results for a multimedia, shared-story intervention for teaching vocabulary to children with intellectual disability. The use of technology in teaching may be effective because the mixed presentation of information (e.g., visual and auditory modes) may reduce the cognitive load of knowledge construction (e.g., the multimedia learning model; Mayer & Moreno, 2003). Because children with Down syndrome have reduced cognitive capacity (Lukowski et al., 2019; Silverman, 2007), technology-based interventions may be particularly suitable to support their learning. This project aims to investigate the effect of the Down Syndrome LanguagePlus (DSL+) intervention, which is a novel digital vocabulary intervention developed for children with Down syndrome.

**Vocabulary Development**

On average, children with Down syndrome produce their first words at the end of their second year of life (Berglund et al., 2001), and they reach the 50-word milestone after three years of age (Berglund et al., 2001; Gillham, 1990). In contrast, typically developing children reach these milestones at approximately 12 months and 17-18 months, respectively (Kristoffersen et al., 2012). The delay in vocabulary development in children with Down syndrome extends to the syntactic domain (Berglund et al., 2001; Chapman & Hesketh, 2001). It is hypothesized that they need a corpus of 100 words before they can start to combine words into sentences (Lewis, 2003). In contrast, approximately 50 words are required for typically developing children (Kristoffersen et al., 2012). Notably, individuals with Down syndrome reach developmental milestones at significantly later ages than their typically developing peers.
(Cuskelly et al., 2016; Næss et al., 2015, 2021), but there is substantial variability across individuals (Berglund et al., 2001).

**Previous Vocabulary Interventions for Children with Down Syndrome**

To date, few intervention studies have attempted to increase vocabulary in children with Down syndrome. The lack of research-based vocabulary intervention programs has been confirmed by several systematic reviews of interventions aiming to promote communication and language development in children with Down syndrome (e.g., O'Toole et al., 2018; Smith et al., 2020). For example, according to Smith et al. (2020), only three of the studies included in their review trained vocabulary. Overall, a moderate-to-high risk of bias was found across the included studies, and the methodological challenges were related to age spread and sample size. According to this systematic review, the study by Burgoyne et al. (2012) is the largest randomized control trial (RCT) to date. This study compared a combined language and reading intervention to business as usual (N=57) and is the only intervention study in the Smith et al. (2020) review that included vocabulary training in school-age children. This study was a one-to-one paper-based intervention. While the intervention group showed significantly greater progress than the business-as-usual group in word reading, letter-sound knowledge and phoneme blending, there were no significant effects on receptive vocabulary and only modest effects on trained expressive vocabulary. There were no transfer effects to untrained language. Therefore, there is an urgent need for effect studies targeting vocabulary in children with Down syndrome specifically.

**Purpose of the Study**

In this study, we developed the DSL+ intervention, which is a vocabulary intervention program tailored for children with Down syndrome. The main purpose of the study was to
investigate the effects of the intervention compared to those of business as usual among first
graders with Down syndrome to contribute to knowledge regarding how to tailor vocabulary
interventions to the specific needs of children with Down syndrome. The following sub-
questions were investigated:

1. Is the DSL+ intervention more effective than business as usual on measures of trained
   expressive vocabulary, receptive vocabulary, and vocabulary breadth and depth?

2. Is the DSL+ intervention more effective than business as usual on standardized measures
   of expressive vocabulary, receptive vocabulary, receptive grammar, listening
   comprehension and narrative skills?

3. Is there an interaction effect between post-test performance and the following child
   factors: a) nonverbal mental ability, b) attention, and c) pre-test score?

**Methods**

This study was a stratified cluster RCT testing the effect of the DSL+ intervention on two
waves of first graders with Down syndrome (level 1) nested within schools (level 2). Ethical
approval was granted by the Norwegian Centre for Research Data. Informed parental consent
was obtained for all children. Consent was also obtained from all the participating schools and
the implementers (the educator responsible for the children’s daily education and the mainstream
schoolteachers) before randomization. Due to the inclusion of peers in the teaching activities, the
parents of all children in the participating children’s class were also informed of the study and
given the opportunity to actively opt out of participating in the DSL+-teaching activities.

**Sample**

Through the national habilitation service, we invited a two-wave complete national
sample of first graders with Down syndrome to participate in the study. The participants had to
1) be first graders with Down syndrome, 2) have no comorbid diagnosis of autism, and 3) have at least one parent with Norwegian as his/her first language. The details of the participant recruitment, allocation and flow throughout the study are summarized in the flow chart in Figure 1.

**Figure 1**

*Flow Chart Showing Participant Recruitment, Allocation and Flow through the Study*

[Insert Figure 1 here]

*Note. The figure is based on CONSORT (Schulz et al., 2010).*

Two clusters in the intervention group withdrew after the implementer training because the implementers determined that the intervention was not suitable for the children (n=2). However, based on the intention-to-treat (ITT) principle, these participants continued to be included in the analysis. The final sample included in the analysis consisted of 103 participants, including 22 females/29 males in the intervention group (mean age in months=76.61, SD=5.14) and 27 females/26 males in the control group (mean age in months=77.47, SD=5.66).

**Randomization**

The schools were randomly allocated to either the DSL+ intervention or control group (“business as usual”) by an independent researcher (in two rounds) using MATLAB script with the constraint that an equal number of students were placed in each group in each stratum. The randomization was conducted prior to the pre-test assessments and prior to the start of the school year to allow the schools to prepare accordingly.

**DSL+ Intervention**

The aim of the DSL+ intervention is to increase children’s expressive and receptive vocabulary, including their vocabulary breadth and depth, and improve their language skills in
general. The program is based on the findings of a longitudinal study of language and reading
development in a national age cohort of first- to third-grade children with Down syndrome
(Næss, 2012; Næss et al., 2015, 2021); systematic reviews and meta-analyses of previous
research concerning language and auditory memory skills, phonological awareness and
motivation; and previous language interventions targeting this group of children (Næss, 2016;
Næss et al., 2011; Smith et al., 2017, 2020). In the following, the DSL+ intervention is briefly
described based on the Template for Intervention Description and Replication (TIDieR) checklist
(Hoffmann et al., 2014).

The following three acknowledged theoretical approaches for vocabulary development
and learning inspired the program development: 1) The lexical quality hypothesis (Perfetti, 2007)
is an explicit didactic approach according to which repeated exposure to a target word and
multiple connections between the child’s experiences and the new word are considered necessary
for word acquisition and the development of a nuanced and flexible understanding of the word.
2) Dialogic interaction through picture book-sharing is an approach using narratives with the
rich semantic support of illustrations to engage children in dialogic exchanges found to stimulate
vocabulary acquisition and inferential language skills (Van Kleeck, 2008). This approach
highlights the value of re-exposure, active involvement and opportunities to participate in shared
sense-making and decontextualized talk (Damhuis et al., 2015; Van Kleeck, 2008). 3) Statistical
learning is an experimental, implicit approach focusing on the importance of variations in the
input and extraction of information regarding structure to support word learning (Arciuli et al.,
2014). Statistical learning also highlights the facilitation of learning through multiple encounters
with the different surface forms of an item (e.g., pictures representing the word, voices
expressing the word or morphological elements included in words, e.g., different words with the
same prefix), and high variability of input has been shown to be more effective than low variability (e.g., Plante et al., 2014). None of these theoretical approaches have been tested in children with Down syndrome.

The current program utilizes the following dual approach: dialogic interaction through picture book-sharing (c.f. Damhuis et al., 2015; Van Kleeck, 2008) and a systematic approach using a set framework of training tasks inspired by the lexical quality hypothesis (Perfetti, 2007) and statistical learning theory (Arciuli et al., 2014; Plante et al., 2014). The digital material was developed as a multimedia presentation using pictures, video, animation, sounds, and voices to reduce the cognitive load of the children. The possibility of cognitive overload (i.e., redundancy effects; Greer et al., 2013) was considered by developing an innovative framework to limit the number of features and maintaining awareness of this challenge. The intervention was developed for children with Down syndrome in mainstream schools and was implemented by special education teachers/teacher assistants (depending on the assigned resources for individual children at the school) in collaboration with the mainstream schoolteacher. The intervention was delivered in one-to-one settings on the first three days of the week, in small groups with children from the mainstream classroom on the fourth day, and in the full, mainstream class on the fifth day.

The intervention program included daily sessions of approximately 15 minutes over 15 weeks (75 sessions). The training was organized with a reoccurring task structure and divided into 4-week periods. During each period, the first three weeks involved picture book sharing and corresponding tasks, which were designed for the project following a generic structure to introduce new words each week. The fourth week had a different structure, and the task types
emphasized the consolidation and in-depth learning of the vocabulary from the previous three weeks.

All participating schools were provided the following materials necessary to implement the intervention: an iPad with the DSL+ app installed; teaching materials, including picture cards, toys and games; written manuals and scripts for the picture book dialogues; and online versions of the intervention program for implementer preparation purposes. All implementers participated in a six-hour obligatory interactive online course covering the theoretical and empirical background of the intervention, guidelines for intervention delivery and practical information. An essential component of the guidelines for the implementers was their supportive role as mediators and motivators during the sessions. Members of the research team provided implementer support throughout the intervention period via telephone and e-mail. For an overview of the intervention structure and content, see Figure 2.

**Figure 2**

*Overview of the Intervention Structure and Content*

[Insert Figure 2 here]

*Note.* For a more detailed description of the intervention content, see the DSL+ teacher guidance (Næss et al., 2016). Different visual formats (drawings, pictures, cartoon characters, etc.) that illustrate the same word are used throughout the program to train generalization. Illustrations: Beate Erikke Johansen, Gunn Berit Langeland, Marielle Laastad, and Elisabeth Smith.

**Outcomes**

The pre-test (t1) and post-test (t2; after 15 weeks) assessments were carried out individually. All measures were collected at both time points. Table 1 provides an overview of the measures used.
Data Collection

The assessments were carried out in two separate sessions in a separate room at the school, and the child was accompanied by his or her special education teacher or teacher assistant. Trained assessors (DSL+team members and research assistants) conducted the assessments. In total, 63.1% of the children at t1 and 73.3% of the children at t2 were tested by assessors who were blind to the children’s allocation.

Fidelity

The participants’ activity on the iPad was tracked by collecting log data and audio files from the individual sessions to allow for the assessment of the delivered dosage (average length of the individual intervention sessions and percentage of the individual intervention sessions attended) and script adherence (the implementer’s loyalty to the picture book scripts). Information regarding program differentiation (how the content of the intervention differs from that of other pedagogical methods used) was collected via digital implementer reports. To assess the dosage, we first calculated the average length of the picture book dialogues and structured tasks from a randomly selected week of the intervention (week 5). The average length of the picture book dialogues was 7:21 minutes (range: 2:35-13:16), and the average length of the structured tasks was 7:36 minutes (range: 3:43-11:43). Second, we calculated the average percentage of attended individual intervention sessions (maximum: 45 sessions). The average attendance in the intervention group (n=43) was 84.96% (SD: 11.98; range: 57.78-100%).

For the adherence variable, one audio recording (randomly selected from one of the individual intervention sessions) from one picture book (randomly selected from 11 picture books) was analyzed. The implementer’s script usage (questions, hints and model sentences) was
coded as 1 for script adherence or 0 for non-adherence when the implementer missed a part of the script that should have been included. The average script adherence in the intervention group (n=41) was 88.68% (SD: 12.95; range: 36.00–100%).

The intervention was not planned to be personalized, tailored or adapted beyond the stages in the scripts. However, 5 children did not complete the weekly program in full compliance with the original intervention program (see Figure 2 for a detailed description of the intervention program). The following adaptions were made in collaboration with the DSL+ team: 1 child consistently omitted Day 1 (an individual session) and 1 child consistently omitted Day 5 (a classroom activity) of the intervention because the rural areas of Norway where these children live have a 4-day school week. Furthermore, 1 child omitted the classroom activity and 2 children omitted both the group activity and classroom activity because these teaching activities were considered too difficult for the children as evaluated by their educator.

**Program Differentiation and Business as Usual**

To gain insight into the additional instruction received by the intervention group and the nature of the “business as usual” (BAU) instruction received by the control group, a questionnaire was sent to the educator responsible for the child’s daily education. The results showed that 95% of the intervention group (n=41) received additional vocabulary instruction at school for an average of 2 hours weekly. All but one (98%) of the participants in the control group (n=41) received vocabulary instruction during the control period for an average of 3 hours weekly. Therefore, the total amount of time spent on language instruction was similar across the groups. Additionally, most educators in the control group (90%) described their vocabulary instruction as systematic and detailed that it included book activities (95%), problem solving (70%), play and/or game activities (100%) and the use of technical tools/iPads (89%/84%). This
indicates that the types of activities included in the DSL+ intervention were also commonly used in the language instruction of the BAU group.

**Statistical Analysis**

Regression modeling (ANCOVA) and structural equation modeling (SEM) were used to test the effect of the DSL+ vocabulary intervention. First, the intervention effects on each language outcome were investigated with regression models, and the post-test outcomes were regressed on a group dummy variable and the corresponding pre-test outcome. The effect sizes are reported as standardized on $y$ and can be interpreted as Cohen’s $d$, i.e., the difference between the groups in standard deviation (SD) units.

Then, to investigate the overall effect of the intervention, the following two SEM models were estimated: one model for language and one model for vocabulary. The language factor was constructed by testing the common variance in the observed standardized language measures (the BPVS-II, WPPSI-III and TROG-2). The vocabulary factor was constructed by testing the common variance in the researcher-developed measures of trained vocabulary (receptive vocabulary breadth, expressive vocabulary breadth and vocabulary depth). Measurement invariance over time was tested with a confirmatory analysis (CFA).

All analyses were conducted using Mplus version 8 software (Muthén & Muthén, 1998-2017). The average cluster size was 1.13 participants, and the analysis was conducted at the participant level. Although the design effects were minimal due to the small average cluster size, we performed Huber-White correction of the standard errors. The analyses were conducted according to the intention-to-treat (ITT) principle, and all 103 participants who completed the pre-test assessments were included. Missing data were handled by the full information maximum likelihood (FIML) estimation strategy.
Results

Background Measures

The means and SDs of the background measures at the pre-test, the reliability values and the effect sizes with 95% confidence intervals (CIs) for differences between the groups are shown in Table 2. As expected, given the random assignment of the intervention group and the control group, there were no significant differences in backgrounds measures between the groups at the pre-test.

[Insert Table 2 here]

Effects of the Intervention on Trained Vocabulary

Table 2 also shows the mean, SD, reliability, and effect size (with the 95% CI) of each trained vocabulary measure. The results show that the intervention was successful, as the children learned the trained vocabulary. A significant effect was observed in both the receptive measure (words the child understands) and expressive measure (words the child can say), with an effect size of $d = .429$ (95% CI [.160, .699]) for trained expressive vocabulary and $d = .447$ (95% CI [.193, .700]) for trained receptive vocabulary. Regarding the vocabulary depth measure (how well the words are known, i.e., the ability to categorize words or connect the trained words to related words), the participants answered few items correctly on the pre-test, and the results were nonsignificant, with an effect size of $d = .209$ (95% CI [-.108, .526]).

The following two additional measures were used to assess the effect of the intervention on trained vocabulary: expressive vocabulary depth (ability to explain words) and expressive grammar (singular/plural and present/past). However, because the results showed substantial floor effects for these measures, they were removed from further analysis. No interaction effects on trained vocabulary were found between the treatment and the covariates attention and
nonverbal mental ability. In addition, no interaction was found between the pre-test and the treatment, which confirms equal regression slopes for the two groups.

The constructed latent variable model of trained vocabulary showed a significant general effect of $d = .197$ (95% CI [.060, .333]) (see Figure 3A). The model fit was excellent ($X^2 (9) = 8.710, p = .464$, root mean square error of approximation (RMSEA) = .000 (90% CI [.000, 1.08]); Tucker–Lewis index [TLI] = 1.002). However, this model did not have metric invariance (noninvariant factor loadings) over time, indicating that the trained vocabulary factor changed between the pre-test and post-test.

**Figure 3**

*Latent Variable Model of Trained Vocabulary and Standardized Language*

[Insert Figure 3 here]

*Note.* Vocab = vocabulary.

**Generalization to Standardized Language**

Table 2 shows the mean, SD, reliability and effect size (with the 95% CI) of each standardized language measure. The results indicate no generalization effects to expressive vocabulary ($d = -.061, 95\% \text{ CI } [-.246, .125])$, receptive vocabulary ($d = .154, 95\% \text{ CI } [-.104, .411])$, or grammar ($d = -.089, 95\% \text{ CI } [-.368, .190])$.

The following three additional measures were used to assess the effect of the intervention on standardized language: the Bus Story (narrative skills; Renfrew, 2010), Listening Comprehension (Hagen et al., 2017) and Grammatic Closure (Kirk et al., 1967). However, because the results showed substantial floor effects for these measures, they were not included in further analysis. No interaction effects on standardized language were found between the treatment and the pre-test performance or the covariates attention and nonverbal mental ability.
The latent variable model of language (standardized) (Figure 3B) subsequently showed no significant effect on generalization to standardized language ($d=-.006$, 95% CI [-.077, .090]). The model fit was excellent ($X^2 (14)=19.772; p=.138; \text{RMSEA}=0.063$ (90% CI [0.000, 0.123]; TLI=.980). In this model, there was full metric invariance and partial scalar invariance (the intercept of the Picture Naming measure had to be freely estimated), indicating that the standardized language factor could be interpreted similarly over time.

**Discussion**

To date, this study is the largest RCT of a digital vocabulary intervention for an age cohort of children with Down syndrome in a school setting (c.f. review by Smith et al., 2020). The findings of this RCT suggest that compared to business as usual, the DSL+ intervention with the small dosage of 15 minutes of training a day five days a week for 15 weeks is in fact effective in improving trained expressive and receptive vocabulary breadth. However, the results show no significant transfer effect to standardized measures of vocabulary or grammar.

**Effect of the DSL+ Intervention**

The total amount of time spent on language instruction was similar between the intervention group and the control group, and the types of teaching activities included in the DSL+ intervention were commonly used in the language instruction of the control group. The results showed a significant direct effect of the intervention; the children in the intervention group had significantly larger growth in their understanding and expression of words than the children in the control group. This finding is consistent with the results of a systematic review carried out by Smith et al. (2020) showing that children with Down syndrome can experience gains in their language skills as a result of interventions. The significant effect on trained expressive vocabulary is consistent with the results of the broad-based reading and language
intervention investigated in the RCT by Burgoyne et al. (2012). However, to the best of our knowledge, no previous RCT involving systematic training in children with Down syndrome in this age group has shown a significant effect of an intervention on trained receptive vocabulary compared to that achieved with a business-as-usual approach.

In the current RCT, four key components may have contributed to the children’s understanding of words. First, the children may have benefited from cross-modal statistical learning, repetition and exposure to a large number of variations in the vocabulary items, including both visual and vocal referents, which is known to increase implicit learning and consolidation (Hennies et al., 2014). Second, the holistic approach combining semantic, phonological, morphological, syntactical and pragmatic content may have supported vocabulary learning, as these components are essential for a full understanding of the words in one’s vocabulary (Bloom & Lahey, 1978). Third, the modality effects from the technology mixing the presentation forms of pictures, video, animation, sounds, and voices may have reduced the cognitive load of the children (Mayer & Moreno, 2003) and promoted the development of semantic knowledge through a more naturalistic representation of the target words.

Fourth, encountering and learning about words in different contexts (e.g., digital and physical material/practical activities organized in a combination of individual, small-group, and classroom sessions) may have increased the children’s understanding of the words (see, e.g., Nagy, 1995) and provided varied opportunities to practice the words. Repetition and practice have been found to support language learning among typically developing children (Biemiller & Boote, 2006). The fact that the intervention also represents a novel approach with a digital and interactive format, in addition to the possible increased knowledge and awareness of the
implementers after participation in the educational course, may also have contributed to the positive gains.

Based on previous cross-sectional and longitudinal research, children with Down syndrome are usually described as having a specific weakness in expressive vocabulary skills (i.e., Martin et al., 2009) compared to their receptive skills. However, the current intervention study provides no indication that children with Down syndrome have more problems with improving their expressive vocabulary breadth than their receptive vocabulary breadth skills. Instead, the effect sizes for these two measures comparing the growth in the intervention and control groups are quite similar, indicating that these two modalities can equally be improved with systematic training.

The effect found on vocabulary breadth was present only with specifically trained words and did not generalize to standardized tests, which was somewhat unsurprising given the length of the intervention and the rarity of transfer effects in language interventions for children with Down syndrome. No transfer effects have been found in previous group design language intervention studies involving children with Down syndrome (c.f. studies included in Smith et al., 2020), except for two studies training grammar (Baxter et al., 2018; Sepúlveda et al., 2013). In the present study, one singular/plural task and one present/past tense task were included in the training each day; however, these grammatical components were not the primary target of training but rather provided context to support the overall understanding of the specific target words. Therefore, it is highly possible that in the present study, the grammatical aspects were not covered to the extent needed for the children to abstract the rules (i.e., as suggested by Plante et al., 2014) and transfer them to a standardized measure. Statistical learning theory and the model of multimedia learning inspired the development of the DSL+ program to help children
understand that different visual and auditory referents represent the same word. Therefore, the intervention was used in this study with the intention of broadening and generalizing the children’s understanding of one word rather than helping them build abstract systems to generalize their understanding to new words. Thus, the only way this program could have resulted in transfer to standardized tests would have been through an accumulative learning effect of words in general; this effect did not occur, and such generalization is uncommon in children with intellectual disability in general (Kaiser & Roberts, 2013; Kaiser & Trent, 2007). However, by using a test including items with visual referents of target words that were not included in the training sessions, we could have explored the children’s near transfer effects. The time frame of the present project did not allow for the development, piloting, and revision of such tests, but future research should prioritize such an examination.

The lack of invariance over time in the trained words might suggest that the vocabulary construct varied over time because the intervention affected various aspects of this construct differently, with greater growth in vocabulary breadth than depth. Thus, we cannot make strong claims that the intervention had effects on the same unitary underlying vocabulary construct over time. However, this finding may not be surprising since the same result has been found in intervention studies involving children with poor vocabulary skills without Down syndrome (Hagen et al., 2017). At least three potential explanations related to the limited growth in vocabulary depth exist. 1) The focus on vocabulary depth is premature considering the children’s limited vocabulary breadth skills, and improving one’s vocabulary depth requires deeper processing, which is a particular weakness in other groups with intellectual disability (Purser et al., 2011). 2) The measures may not be sensitive enough to measure vocabulary depth in the mastery level of this group of children at this age. Binder et al. (2017) found that different tests
intended to capture vocabulary depth were not correlated. Therefore, the use of both sensitive
tests and an extensive battery of assessments may be necessary. 3) The DSL+ program may not
be effective in training vocabulary depth because the provided dosage was insufficient (c.f.
Kaiser & Roberts, 2013) or the depth tasks did not sufficiently accommodate the specific
learning needs of this group of children (c.f. Laws et al., 2015).

The lack of interaction between the treatment and nonverbal mental ability was somewhat
unsurprising because there are uncertainties about the influence of nonverbal mental ability in
relation to children’s language outcomes and response to interventions (Norbury et al., 2016). In
contrast, the lack of interaction between treatment and attention is more surprising considering
the general importance of sustained attention in learning (Shalev et al., 2016). However,
detecting such interactions often demands high statistical power. Further investigations of
differential responsiveness to interventions could be beneficial to gain knowledge regarding the
features characterizing responders and low responders and subsequently inform the development
of more differentiated interventions for children with Down syndrome.

**Implications for Practice and Future Research**

The combination of the positive intervention effect on trained vocabulary, high
implementer compliance and low attrition rate shows that the digital tool can be successfully
implemented with first graders with Down syndrome after implementers have participated in a
short educational course. The dosage is only 15 minutes daily. In comparison, the mean session
duration was 53 minutes in the review of language interventions for children with Down
syndrome by Smith et al. (2020). Thus, the proposed intervention is particularly feasible to
implement within an everyday school context. The short session duration time minimizes the
time spent outside the mainstream classroom and accounts for the short attention span of young
children in general (Reynolds & Romano, 2016) and children with Down syndrome in particular (Breckenridge et al., 2013). The 15-week length of the intervention period was also shorter than those in previous studies (e.g., 20 weeks in Burgoyne et al. (2012) and 9 months in Yoder et al. (2014)). Because the results from this short study were positive, a longer intervention period may further increase the effect and should be tested in future research.

Another method to increase the dosage of the DSL+ intervention is to apply collaborative implementation across the school and home contexts; parent involvement can increase the intervention dosage and provide the child with diverse opportunities to use the target words in new contexts and with other communication partners. A short, simplified version of the DSL+ intervention has been tested in the [country] for home implementation; the results show that the simplified DSL+ approach can be successfully implemented by parents (LeJeune et al., 2021).

Absence from school is found to predict vocabulary growth (Lervåg et al., 2019). However, the portability of mobile devices enables interventions to continue when children cannot attend school, e.g., due to pandemic restrictions and sick leave, and facilitates collaboration between schools and parents. Children with Down syndrome may be at greater risk than typically developing children of missing school because of comorbid physical conditions (e.g., congenital heart defects and anomalies in the digestive system and limbs; Lanzoni et al., 2019), which reduces access to teaching activities. The flexibility of online accessibility with mobile devices may eliminate the absence-from-school barrier to learning opportunities.

The DSL+ intervention is cost-effective, suitable for digital intervention in schools, easy to distribute digitally both throughout Norway and internationally (c.f. also LeJune et al., 2021), easy to administer, systematic, and predictable for both the child and the implementer. Although certain predefined content was tested in this project, the app is flexible and enables the
implementer to individualize the content with a specific child in mind by adding customized materials.

The outcomes of this study are important and timely in highlighting a method to help children with intensive needs, such as those with Down syndrome, increase their vocabulary via digital apps that are designed with their specific needs in mind. Because vocabulary development is vital for long-term academic and social outcomes, this intervention may increase the potential for lifelong participation in society for individuals with Down syndrome.

Limitations

The inclusion of children regardless of their verbal communication abilities may have increased the SDs. Future studies can potentially gain from distinguishing nonverbal and verbal children by blocking to maintain randomization when separately exploring the groups.

No assessment of long-term effects was included in the present study. It is important to assess maintenance and progress over time in future studies to help us understand how the intervention affects development over time and what support is needed in the long term to ensure progress. Previous intervention studies involving children with Down syndrome have found that targeted aspects of language showing no significant gains at the post-test significantly improved among the children in the trained groups at a later follow-up (see, e.g., Cleave et al., 2011).

Conclusions

DSL+ is a digital vocabulary intervention program especially developed for 6-year-old children with Down syndrome. Testing the effect in a nationwide Norwegian RCT of an age cohort of first graders with Down syndrome showed that working on the DSL+ intervention for approximately 15 minutes a day, five days a week for 15 weeks was effective in improving the trained expressive and receptive vocabulary breadth. There was no transfer effect to standardized
tests, but the intervention study showed that children with Down syndrome are able to learn both receptive and expressive vocabulary through systematic training. Based on these results, all children with Down syndrome should receive continuous vocabulary intervention, and the words included in their education should be carefully chosen.
References


template for intervention description and replication (TIDieR) checklist and guide. BMJ, 348(3), g1687. doi:10.1136/bmj.g1687


Figure 1. Showing Participant Recruitment, Allocation and Flow through the Study
Days 1-3

Picture Book Sharing with Scripts

The text-free picture books are developed for this intervention, and the story is based on a set of selected words. The picture book sharing is organized as a one-to-one activity and designed to encourage teacher-child dialogue.

To stimulate both literal and inferential talk, the teachers are provided with a script (separate booklet) containing questions, hints and model sentences for each book. The teachers are trained to include the script content as a minimum of input to the dialogue (with hints and model sentences when needed) and to encourage and follow up on any initiative from the child.

Each week, the same framework is followed:

Day 1: introduction of a new picture book; scripted dialogue at a concrete level.
Day 2: repetition of the book; scripted dialogue with an increased number of abstract questions.
Day 3: repetition of the book; scripted dialogue with more abstract questions.

Days 1-3

Systematic Training Tasks

The systematic training tasks are based on the same set of selected words as those in the picture books. The systematic training tasks are organized as one-to-one activities and include 8 categories of tasks that focus on language content, form and use:

1) variations (multiple referents to the target word); 2) relations (words associated with the target word); 3) category (group to which the target word belongs); 4) syntax (usage of the target word in context); 5) articulation (pronunciation of the target word); 6) phonology (first sound of the target word); 7) morphology (singular/plural tense of the target or related noun); 8) morphology 2 (present/past tense of the target word or related verb).

The requirement for the child’s responses is that they gradually increase in comprehension and expression.

Each week, the same framework is followed:

Day 1: listening comprehension tasks; viewing images and listening to words in different contexts.
Day 2: word understanding tasks; selecting images based on audio instruction.
Day 3: word production tasks; training articulation and pronunciation of words.

Days 4-5

Picture Book Sharing and Practical Activities

Picture book sharing and practical activities are based on the same books and set of selected words as those in days 1-3. The tasks focus on pragmatics and aim to provide the child with experience using the words in different contexts with their peers. The activities are delivered in small groups on day 4 and in a full class on day 5.

Each week, the same framework is followed:

Day 4: Retelling the picture book, listening comprehension tasks, and collaborative activities related to a target word and story (e.g., sorting tasks, play, and discussions). Words are used in new natural contexts to promote in-depth knowledge of the words.
Day 5: Reading aloud, listening comprehension tasks, and collaborative activities that focus on the narrative reconstruction: covering all aspects that have been learned on days 1-3 in new natural contexts and in more depth.

Figure 2. Overview of the Intervention Structure and Content
Figure 3. Latent Variable Model of Trained Vocabulary and Standardized Language
Table 1

*An Overview of the Measures and Scoring*

<table>
<thead>
<tr>
<th>Measure</th>
<th>Reference</th>
<th>Items</th>
<th>Max.</th>
<th>Description</th>
<th>Procedure</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Background measure</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Block Design</td>
<td>WPPSI-III</td>
<td>20</td>
<td>40</td>
<td>Measures nonverbal mental ability</td>
<td>The child is shown several building blocks in a pattern that is modeled by the assessor (items 1–13) or visualized in a picture (items 13–20). The child is asked to copy the block pattern.</td>
</tr>
<tr>
<td></td>
<td>Wechsler, 2002</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Primary outcome measures</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expressive Breadth¹</td>
<td>DSL+</td>
<td>35</td>
<td>35</td>
<td>Measures the ability to verbally express words</td>
<td>The child is asked to name single pictures.</td>
</tr>
<tr>
<td>Receptive Breadth¹</td>
<td>DSL+</td>
<td>31</td>
<td>31</td>
<td>Measures the understanding of words</td>
<td>The child is shown four pictures and is asked to point to the picture corresponding to the word spoken aloud by the assessor.</td>
</tr>
<tr>
<td>Test</td>
<td>Measurement</td>
<td>Details</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>------------------------------------------</td>
<td>-------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Receptive Depth</strong></td>
<td>DSL+</td>
<td>Measures the depth of understanding of words (e.g., cold) and four picture alternatives of a related word (e.g., sand, grass, snow and planks). The child is asked to choose the item related to the target word.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Secondary outcome measures</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Picture Naming²</td>
<td>WPPSI-III</td>
<td>Measures the ability to verbally express words.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Wechsler, 2002</td>
<td>The child is asked to name single pictures.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>British Picture Vocabulary</td>
<td>Dunn et al., 1997</td>
<td>Measures the understanding of words.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Norwegian ver. by</td>
<td>The child is shown four pictures and is asked to point to the picture corresponding to the word spoken aloud by the assessor.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scale-II (BPVS)²</td>
<td>Lyster et al., 2010</td>
<td>Measures receptive grammar.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for Reception of Grammar</td>
<td>Bishop, 2003</td>
<td>The child is shown 4 pictures and is asked to point to the picture corresponding to the sentence read by the assessor.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Norwegian ver. by</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lyster &amp; Horn, 2009</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The primary outcome measures assessed the direct effect of the intervention and thus were proximal to the content of the intervention. These measures were developed and piloted by the DSL+ team and were all created based on existing standardized measures, such as the BPVS and Picture Naming. The secondary outcome measures were more distal to the content of the intervention but were related to the desired transfer effects. The standardized measures used are internationally accepted and have satisfactory validity and reliability in research involving individuals with Down syndrome (e.g., Burgoyne et al., 2012; Mengoni et al., 2014; Næss et al., 2015).
Table 2

Alpha Reliability, Mean Raw Scores, SDs, and Effect Sizes of Group Differences

<table>
<thead>
<tr>
<th>Measure (max.)</th>
<th>α¹</th>
<th>Intervention group</th>
<th>Control group</th>
<th>Cohen’s d</th>
<th>[95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>N  M  SD</td>
<td>N  M  SD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Background measures</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age in months</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-test</td>
<td></td>
<td>50 76.84 5.20</td>
<td>53 77.45 5.60</td>
<td>.113</td>
<td>[-.274, .500]²</td>
</tr>
<tr>
<td>Nonverbal mental ability</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-test (40)</td>
<td></td>
<td>.84 50 10.56 6.14</td>
<td>52 10.71 6.46</td>
<td>.024</td>
<td>[-.364, .412]²</td>
</tr>
<tr>
<td>Attention</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-test (40)</td>
<td></td>
<td>.97 47 13.73 8.42</td>
<td>47 12.04 8.36</td>
<td>-.200</td>
<td>[-.605, .205]²</td>
</tr>
<tr>
<td>Primary outcomes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expressive vocabulary</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-test (35)</td>
<td></td>
<td>.90 49 4.37 3.50</td>
<td>50 4.04 3.51</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-test (35)</td>
<td></td>
<td>.93 47 9.68 8.06</td>
<td>53 6.28 5.74</td>
<td>.429</td>
<td>[.160, .699]³</td>
</tr>
<tr>
<td>Receptive vocabulary</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-test (31)</td>
<td></td>
<td>.92 48 11.58 7.67</td>
<td>52 11.08 7.73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-test (31)</td>
<td></td>
<td>.93 47 17.70 8.38</td>
<td>51 13.98 8.05</td>
<td>.447</td>
<td>[.193, .700]⁵</td>
</tr>
<tr>
<td>Vocabulary depth</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-test (22)</td>
<td></td>
<td>.76 50 2.84 3.10</td>
<td>51 2.12 2.08</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-test (22)</td>
<td></td>
<td>.75 46 3.57 3.05</td>
<td>50 2.72 2.58</td>
<td>.209</td>
<td>[-.108, .526]³</td>
</tr>
</tbody>
</table>
## Secondary outcomes

### Expressive vocabulary

<table>
<thead>
<tr>
<th>Pre-test (30)</th>
<th>.93</th>
<th>49</th>
<th>7.76</th>
<th>6.44</th>
<th>53</th>
<th>7.98</th>
<th>6.78</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post-test (30)</td>
<td>.94</td>
<td>48</td>
<td>8.54</td>
<td>6.92</td>
<td>53</td>
<td>8.83</td>
<td>7.05</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-.061</td>
<td>[-.246, .125]³</td>
</tr>
</tbody>
</table>

### Receptive vocabulary

<table>
<thead>
<tr>
<th>Pre-test (144)</th>
<th>.94</th>
<th>49</th>
<th>19.98</th>
<th>13.65</th>
<th>52</th>
<th>21.63</th>
<th>11.76</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post-test (144)</td>
<td>.94</td>
<td>48</td>
<td>26.10</td>
<td>14.29</td>
<td>52</td>
<td>25.19</td>
<td>12.49</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.154 [-.104, .411]³</td>
</tr>
</tbody>
</table>

### Grammar

<table>
<thead>
<tr>
<th>Pre-test (80)</th>
<th>.85</th>
<th>48</th>
<th>6.08</th>
<th>4.50</th>
<th>52</th>
<th>7.06</th>
<th>6.04</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post-test (80)</td>
<td>.87</td>
<td>48</td>
<td>7.69</td>
<td>5.78</td>
<td>51</td>
<td>9.04</td>
<td>6.30</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-.089 [-.368, .190]³</td>
</tr>
</tbody>
</table>

Note. Max. = maximum score. Nonverbal mental ability assessed by Block Design (WPPSI-III (Wechsler, 2002)). Attention was assessed by the Reynell Developmental Language Scales (Reynell & Gruber, 1997).¹ The average intercorrelation reliability among the items from Cronbach’s alpha.² Effect sizes of the between-group differences from the independent samples t-test.³ Effect sizes of the intervention based on the difference in progress between the groups from the ANCOVA model divided by the pooled SD of the measure at t1 (see Morris, 2008).