

Document Type: Statistical Analysis Plan

Official Study Title: Effectiveness and Predictors of Intervention Response for a
Technology-based Reading Intervention in the Home

NCT Number: NCT04568824

Unique Protocol ID: H20-0089

Document Date:08/24/2026

Overall Study Summary:

This project seeks to determine the effectiveness of in-home administered reading remediation (GraphoLearn), in a large, diverse, sample of children with reading disability (RD). To this end, we compare the performance of children randomized to GraphoLearn or to an active control (math game), for 12 weeks, and followed beyond treatment over another 12 weeks (retention period) (**Aim 1**). Beyond determining the overall effectiveness of GraphoLearn (for which we expect there to be substantial individual variability), the sample diversity and availability of extensive behavioral and neurobiological data on participants in this study allow the development of explanatory models of predictive factors for GraphoLearn intervention outcomes (**Aim 2**).

Only **Aim 1** is registered as a clinical trial; results from **Aim 2** will not be reported on clinicaltrials.gov.

Analytic plan for Aim 1: Evaluate the effectiveness of GraphoLearn in a large heterogeneous sample.

The design involves participants aged 6.0-10.0 years old entering the study across thirteen waves of enrollment (Oct-Nov, Feb-March, July-August, from October 2020 to November 2024). Note that the age range was extended to 11.0 years of age after study start to increase enrollment.

The three test points (i.e., baseline, post-intervention, retention) are examined in a multilevel (i.e., mixed-effects) regression model of time within participant. Analyses are divided in **primary and secondary analyses/outcomes**.

The primary analysis includes focuses on the Test of Word Reading Efficiency (TOWRE) Total Word Reading Efficiency (TWRE) measure, which combines the Phonemic Decoding Efficiency (PDE) and Sight Word Efficiency (SWE) raw scores (after they have been centered, scaled, and consequently averaged) for the first two timepoints only. Covariates include age, sex, and SES.

The secondary analyses focus on multiple reading-related outcomes (e.g., TOWRE, WJ, and CTOPP) for three timepoints (pre-test, post-test, and retention). These analyses also include an additional mid-point virtual timepoint for some measures, as well as some additional covariates to determine whether the time period of participation (e.g., summer vs. not; early vs. later in the pandemic) affected results.

Primary Analysis (pre- and post- test).

Our primary analysis utilizes a linear mixed effects approach has the general form:

$$Y_{ij} = \beta_0 + \beta_1 \text{Group}_i + \beta_2 \text{Time}_j + \beta_3 (\text{Group}_i \times \text{Time}_j) + \beta_4 \text{Age}_i + \beta_5 \text{Sex}_i + \beta_6 \text{SES}_i + u_i + e_{ij} \quad [1]$$

Where the outcome (Y_{ij}), i.e., the TOWRE reading score for participant i at time j (Pre- or Post- Test), is a function of treatment group (Group_i ; treatment vs control), time point (Time_j), and their interaction ($\text{Group}_i \times \text{Time}_j$). The group-by-time interaction captures differences in the rate of change in Y scores between groups and therefore represents the primary effect of interest. Age_i , Sex_i , and SES_i reflect participant-level background characteristics included as fixed-effect covariates, while u_i a participant-specific random intercept, accounting for the non-independence of repeated observations within participants, and e_{ij} is random error.

Secondary Analysis (pre-test, post-test, and retention).

Our secondary analysis utilizes a stepwise individual growth curve approach:

First. The unconditional individual growth curve approach has the general form:

$$Y_{ij} = \text{Intercept}_i + \text{TimePoint}_j * \text{Slope}_i + \text{Group}_{ij} + e_i \quad [2]$$

Where the outcome (Y_{ij}) for participant i at time j is a function of the participant's level at the first time point (pre-test) (Intercepts as a random effect), the TimePoint_{ij} of the measurement, and the participant's rate of change (Slopes as a random effect). Group_{ij} reflects a block of time-varying covariates (dummy variables) for the differing treatment effects for what the participant received in that time period: GraphoLearn or Math control.

Overall, this is a multilevel growth model reflecting the 3 time points during which participants received testing, along with treatment effects for the contrasts that occur – either at post-test (immediate post-test after the

clinical trial phase) or retention (after the retention phase). For outcome measures also collected at a virtual mid-point visit (i.e., 6 weeks into the clinical trial phase; which are, the two TOWRE subscales and CTOPP's Elision), the mid-point assessment is included as an additional measurement occasion in the growth models.

Second. The conditional growth curve approach has the general form:

$$Y_{ij} = \text{Intercept}_i + \text{TimePoint}_j * \text{Slope}_i + \text{Group}_{ij} + \text{Covariates}_i + e_i \quad [3]$$

This conditional model is the same as the model in Equation 2, but adds five variables to evaluate possible moderation effects of time and participant background: sex assigned at birth, age at pre-test, a mental health composite score, SES, and enrollment wave. In this equation, e_s is random error.

Third. The conditional growth curve with time-varying covariates approach has the general form:

$$Y_{ij} = \text{Intercept}_i + \text{TimePoint}_j * \text{Slope}_i + \text{Group}_{ij} + \text{Summer}_j + \text{Pandemic}_j + \text{Covariates}_i + e_i \quad [4]$$

This conditional model is the same as the model in Equation 3, but adds two time-varying covariate: Summer_j and Pandemic_j. Summer is a seasonal indicator variable that reflects the possible negative effects of a “summer slump” (i.e., decreased performance in the absence of regular instruction during the summer), which is a well-established phenomenon, and might affect test performance during the summer months. Similarly, Pandemic reflects the possible negative effects of a “covid slide” (i.e., decreased performance in the absence of regular instruction during the COVID19 pandemic), which is a well-described phenomenon and might affect test performance during and after the COVID19 pandemic.

Sensitivity analysis.

As a sensitivity analysis, the secondary analysis models (Equations 2-4) will be rerun using three groups (GraphoLearn and the two versions of the math control game) rather than combining both control games into a single group. This analysis accounts for the change in the math control game partway through the trial. Note that because the two math control games were implemented sequentially rather than offered as alternative conditions, with one version used earlier in the pandemic and the other introduced later, the pandemic covariate may be removed from these models to avoid adjusting for a variable that is intrinsically confounded with control-game version. Results will be compared with those from the primary two-group model to determine whether the estimated effect of GraphoLearn is sensitive to pooling the two control conditions.