

Supplemental Material S2. Reliability.

Entering Words Into the CSSS Dictionary

The CSSS includes a large dictionary of words that have already been parsed into elements and have legal spellings assigned for each element. The complete list of words from the WTW assessment was already included in the CSSS dictionary, but the CBM words were not. The author and a graduate student, also trained in the CSSS, together entered these words into the dictionary, coming to agreement for each word on the parsing into elements and the list of legal spellings for each element. In cases of uncertainty, new entries were compared to similar existing entries in the dictionary and agreement was reached between the two coders.

Coder Training and Intercoder Reliability

First, each coder underwent training. For CLS, this included reading the Aimsweb S-CBM training manual (Shinn & Shinn, 2002) and completing the five practice tests in the manual. For SSE and SSW, this included reading the CSSS manual (Masterson & Apel, 2015) and watching the online demonstration videos (<https://www.missouristate.edu/csd/LLL/CSSS.htm>).

Second, intercoder reliability was measured for each of the partial-credit scoring metrics. For CLS, an estimate of intercoder reliability was obtained using an intraclass correlation coefficient (ICC) for total scores. Acceptable criterion based on the literature is an ICC > .80 (Nunnally & Bernstein, 1994). Data for 10 participants (15%) were double-coded and an ICC of .998 was obtained. For SSW and SSE, multiple steps were taken to ensure reliability. First, children's spellings were examined for analyzability. Analyzable spellings are those that contain two elements that were either spelled correctly or represented by a legal grapheme. The author and a graduate student rated the spellings of a random 10 participants as analyzable or not. Initially, there was disagreement on two of the 370 items. The raters discussed the items and came to an agreement. Finally, children's spellings were imported into the CSSS software and compared to the dictionary. Each spelling was manually checked by the author to ensure that the parsings aligned as expected within the software prior to exporting the scores. Data for a random 10 participants (16%) was re-entered by a graduate student to check for reliability of the parsing alignment. An ICC of 1.00 was obtained for SSW and .996 for SSE.

References

- Masterson, J., & Apel, K. (2010). The spelling sensitivity score: Noting developmental changes in spelling knowledge. *Assessment for Effective Intervention*, 36(1), 35–45. <https://doi.org/10.1177/1534508410380039>
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill.
- Shinn, M. R., & Shinn, M. M. (2002). *AIMSweb training workbook: Administration and scoring of Spelling Curriculum-Based Measurement (S-CBM) for use in general outcome measurement*. Retrieved from <http://www.aimsweb.com>