Comparing three methods for quantifying change across a kindergarten program

Theresa Pham, Daniel Ansari, Marc Joanisse, Janis Cardy, & Lisa Archibald, The University of Western Ontario
tpham62@uwo.ca
@tpham62

Introduction

- The ability to assess change in skills is important for outcome measures used in programs and intervention
- However, there is no consensus about how to measure change
- In this study, we evaluated the following three methods for assessing change across a kindergarten program:
  - ***T-test***: Most commonly used in research to evaluate group differences. Change occurs if the p-value < .05
  - ***Reliable change index***[1,2]: Calculation of the smallest difference between pre- and post- scores that likely reflect a true change. Change occurs if a child has an RCI > 1.645
  - ***Normalization***: Most commonly used by speech-language pathologists (SLPs) to evaluate change in a struggling individual. Change occurs if a child moves from <16th percentile to > 25th percentile

Research Questions:

1) **How do the methods compare in capturing change?**
2) **Do the methods capture reliable change on different measures?**
3) **Are there individual differences in responsiveness?**

Methods

Participants: 157 children ages 4 to 5. Data from 25 different participants were used for test-retest reliability

Procedure: As part of a larger study, children completed nine tasks pre- and post-kindergarten: vocabulary test, sentence recall, alphabet knowledge, rapid colour naming, number line estimation, arithmetic skills, magnitude comparison, number name, and phonological awareness

Analysis: For each task, the following analyses were conducted

- T-test: Paired t-test (two-tailed) to evaluate group-level change. Significant p-value and effect size were indicative of change
- RCI: The following equation was used to evaluate individual-level change. An RCI > 1.645 was indicative of change[2]

$$RCI = \frac{D_i}{\sqrt{\left(S_{pre}\sqrt{1-R_{pre\text{-}post}}\right)^2 + \left(S_{post}\sqrt{1-R_{pre\text{-}post}}\right)^2}}.$$

- Normalization: Data was constrained to participants who scored below the 16th percentile on the given task. Participants who move from the 16th to 25th percentile are considered to have changed
- Odds ratios from logistic regressions and Cohen's kappa were further used to evaluate the methods

Results

Full sample	t-test *p < .001	RCI: Proportion responders (odds ratio)	Reliable change difference required
Vocabulary test	t(156) = -6.47*, d = 0.52	0.25 (0.18)	+1.51
Sentence recall	t(156) = -13.45*, d = 1.073	0.68 (0.91)	+3.59
Alphabet knowledge	t(155) = -12.01*, d = 0.96	0.54 (0.30)	+2.42
Rapid colour name	t(155) = 8.98*, d = 0.73	0.42 (1.064)	-17.39
Number line estimation	t(156) = -9.47,*, d = 0.76	0.48 (0.013)	+0.24
Arithmetic skills	t(153) = -10.25*, d = 0.83	0.40 (0.72)	+2.46
Magnitude comparison	t(153) = -5.84*, d = 0.47	0.27 (<.001)	+0.094
Number name	t(156) = -6.63*, d = 0.53	0.27 (0.040)	+1.14
Phonological awareness	t(156) = -17.52*, d = 1.40	0.34 (0.82)	+7.68

Low scorers	Proportion responders (odds ratio)		Cohen's Kappa
	Normalization	RCI	
Vocabulary test (n = 16)	0.38 (3.68)	0.94 (<.001)	0.077
Sentence recall (n = 39)	0.36 (0.97)	0.72 (1.028)	0.36
Alphabet knowledge (n = 28)	0.39 (1.29)	1 (1)	0
Rapid colour name (n = 19)	0.54 (0.97)	0.88 (0.97)	0.29
Number line estimation (n = 24)	0.52 (< .001)	0.52 (<.001)	1
Arithmetic skills (n = 0)			
Magnitude comparison (n = 26)	0.74 (4.10)	0.93 (<.001)	0.37
Number name (n = 20)	0.36 (1.20)	0.68 (1.01)	0.24
Phonological awareness (n = 28)	0.80 (1.28)	0.95 (>100)	0.35

Full sample:

- At the group-level, t-tests indicated significant changes on all measures with medium-large effect sizes
- At the individual-level, RCI identified 25-68% of students as making a reliable change. Odds ratios < 1 indicated that children with low (vs. high) pre-kindergarten scores were more likely to change

Low scorers:

- RCI identified more students as having changed than the normalization method. Agreement was relatively low

Conclusion

How do the methods compare in capturing change?
T-test results indicated improvements across measures, whereas the RCI and normalization methods were able to capture individuals who changed. Agreement was not high, however

Do the methods capture reliable change on different measures?
Depending on the measure, rates of change varied. For the full sample, the RCI method identified 25-68% of students as having changed. In contrast, for the low scorers, the RCI method identified more improvements (up to 100%) compared to the normalization method

Are there individual differences in responsiveness?
For the full sample, children with low pre-kindergarten scores were more likely to change using the RCI method, whereas for the low scorers, children with high pre-kindergarten scores (closer to 16th percentile) were more likely to change using the normalization method

Clinical Implications: Capturing change is important for SLPs but the methods available provide mixed results[1,3]. Nevertheless, these methods provide starting points for measuring change

References
1. Frijters, Lovett, M. W., Sevcik, R. A., & Morris, R. D. (2012). Four methods of identifying change in the context of a multiple component reading intervention for struggling middle school readers. *Reading & Writing, 26(4),* 539–563. https://doi.org/10.1007/s11145-012-9418-z. 2. Estrada, Ferrer, E., & Pardo, A. (2018). Statistics for Evaluating Pre-post Change: Relation Between Change in the Distribution Center and Change in the Individual Scores. *Frontiers in Psychology, 9,* 2696–2696. https://doi.org/10.3389/fpsyg.2018.02696. 3. Hendricks, & Fuchs, D. (2020). Are Individual Differences in Response to Intervention Influenced by the Methods and Measures Used to Define Response? Implications for Identifying Children With Learning Disabilities. *Journal of Learning Disabilities, 53(6),* 428–443. https://doi.org/10.1177/0022219420920379.

NSERC CRSNG

Disclosure Statement: All authors, no conflict of interest

Presented at SRCLD, June 2022