

# Impact of COVID-19 on WJ IV ACH Administration and Scoring Errors

Benson, N.<sup>1</sup>, Lockwood, A.<sup>2</sup>, Klatka, K.<sup>2</sup>, Farmer, R.<sup>3</sup>, Parker, B.<sup>1</sup>, Kubick, R., & Lilly, K.

Baylor University<sup>1</sup>, Kent State University<sup>2</sup>, Oklahoma State University<sup>3</sup>



## Background & Purpose

- COVID-19 is hypothesized to hinder the accuracy with which school psychologists administer and score the WJ IV Tests of Achievement (WJ IV ACH)
- We recruited 35 practitioners to submit test records from before and during the pandemic

## Project Design

- Wilcoxon signed-rank test was used to identify differences in accuracy before and during the pandemic

## Implications for Research & Practice

- The impact of COVID-19 on administration and scoring of the WJ IV ACH appears to be minimal
- Future research is needed that specifically examines the impact of remote administrations

## References

- Lockwood, A. B., Farmer, R., Bohan, K., Winans, S., & Sealander, K. (2021). Academic achievement test use and assessment practices: A national survey of special education administrators. *Journal of Psychoeducational Assessment*, 39(4), 436-451. <https://doi.org/10.1177/0734282920984290>
- Lockwood, A. B., Sealander, K., Gross, T. J., & Lanterman, C. (2019). Teacher trainees' administration and scoring errors on the Kaufman Test of Educational Achievement. *Journal of Psychoeducational Assessment*, 38(5), 551-563. <https://doi.org/10.1177/0734282919871144>
- McGrew, K. S., LaForte, E. M., & Shrank, F. A. (2014). *Technical manual: Woodcock-Johnson IV*. Riverside Publishing.
- Schrank, F. A., Mather, N., & McGrew, K. S. (2014). *Woodcock-Johnson IV Tests of Achievement*. Riverside.
- Styck, K. M., & Walsh, S. M. (2016). Evaluating the prevalence and impact of examiner errors on the Wechsler scales of intelligence: A meta-analysis. *Psychological Assessment*, 28(1), 3-17. <https://doi.org/10.1037/pas0000157>
- Viner, R. M., Russell, S. J., Croker, H., Packer, J., Ward, J., Stansfield, C., Mytton, O., Bonell, C., & Booy, R. (2020). School closure and management practices during coronavirus outbreaks including COVID-19: A rapid systematic review. *The Lancet Child & Adolescent Health*, 4(5), 397-404. [https://doi.org/10.1016/S2352-4642\(20\)30095-X](https://doi.org/10.1016/S2352-4642(20)30095-X)

## Findings

**Table 1**

*Critical Errors by Subtest*

Subtests	N	% Missing	% Error Free, Pre	% Error Free, COVID	Median, Pre & COVID	M, Pre	M, COVID	SD, Pre	SD, COVID
Letter-Word ID	34	2.9	77.1	64.7	0	.40	.50	.81	.79
Applied Prob.	34	2.9	77.1	76.5	0	.31	.29	.63	.63
Word Attack	30	14.3	63.7	61.3	0	.47	.45	.82	.62
Passage Comp.	35	0	62.9	65.7	0	.57	.40	.95	.60
Spelling	29	17.1	37.9	41.9	1	1	.97	1	1.02
Writing Samp.	34	2.9	41.2	38.2	1	1.29	1.41	1.19	1.34
Calculation	33	5.7	23.5	36.4	1	1.26	1.18	.99	1.19
Oral Reading	27	22.9	77.8	85.2	0	.53	.36	.53	.36

Note. N is the number of pairs (pre- and post-COVID) of subtests. Wilcoxon signed-rank tests did not identify any statistically significant differences in critical errors committed prior to and during the pandemic.

**Table 2**

*Types of Critical Errors*

Critical Error	N	% Error Free, Pre	% Error Free, COVID	Median Errors, Pre & COVID	M, Pre	M, Covid	SD, Pre	SD, COVID
Items Scored Incorrectly	28	35.7	27.6	1	1.71	1.86	1.82	1.85
Ceiling	28	32.1	31	1	1.14	1.41	1.04	1.3
Start Point	28	53.6	44.8	1	.71	.72	1.12	.84
Reversal	28	35.7	37.9	1	1.25	1.07	1.32	1.1
Calculation	28	75	79.3	0	.36	.24	.73	.51
Total	28	7.1	0	5	5.18	5.31	3	2.87

Note. N is the number of pairs (pre- and post-COVID) of test records. Errors on fluency subtests are not included due to high rates of missing data as many practitioners did not administer them. Wilcoxon signed-rank tests did not identify any statistically significant differences in critical errors committed prior to and during the pandemic.

**Table 3**

*Types of Non-Critical Errors*

Non-Critical Errors	N	% Error Free, Pre	% Error Free, COVID	Median, Pre	Median, COVID	M, Pre	M, COVID	SD, Pre	SD, COVID
Failure to Record Verbatim	29	6.9	3.3	21	17	23	21.33	18.11	15.69
Failure to Record Qualitative Observation	28	57.1	48.3	0	1	2.46	2.86	3.25	3.34
Failure to Test by Page	30	76.7	66.7	0	0	.37	.6	.77	1.04
Failure to Record DOB	35	91.4	91.4	0	0	.09	.09	.28	.28
Failure to Record Date of Testing	35	94.3	100	0	0	.06	0	.24	0
Total	27	0	0	25	19	25.85	24.83	17.9	15.44

Note. N is the number of pairs (pre- and post-COVID) of completed test records. Paired comparisons were tested using Wilcoxon signed-rank tests did not identify any statistically significant differences in non-critical errors committed prior to and during the pandemic.

