Teacher-assigned report card grades and standardized test scores in young children: Do results align?

Theresa Pham, Daniel Ansari, Marc Joanisse, Janis Cardy, & Lisa M.D. Archibald, The University of Western Ontario

Introduction

- Measures of achievement include:
  - **Teacher-assigned report card marks**: Criticized to be subjective and unreliable, but serve as the main communication of academic success to students and parents
  - **Standardized measures:** Said to be objective and reliable given uniform of testing protocols

- The relationship between marks and standardized tests is moderate[1, 2], correlation of .40-.60, for children in grades 3-6 and high school students. No research in younger children

- Girls are said to earn higher grades than boys, but boys score higher on standardized tests than girls[3]

Research Questions:

1) **How correlated are report card marks and standardized scores in grades 1 and 2?**
2) **Does the relationship depend on the domain, gender, and school?**

Methods

Participants: 218 participants in grade one (5-6-years-old), of which 124 were re-tested in grade two (6-7-years-old).

Procedure:

Report Card Marks: Letter grades were obtained and converted to a numeric 12-point scale, with 1 = D- (lowest) and 12 = A+ (highest)

- Confirmatory factor analysis identified three domains as the best fit:
  Grade 1 fit: CFI = .88; TLI = .86; RMSEA = .14; SRMR = .052
  Grade 2 fit: CFI = .98; TLI = .97; RMSEA = .090; SRMR = .030
  - Language (Oral Communication, Media Literacy)
  - Literacy (Reading, Writing)
  - Math (all strands of Mathematics)

Woodcock-Johnson III (WJ III): Standard scores were used for analysis

- Confirmatory factor analysis identified three domains as the best fit:
  Grade 1 fit: CFI = .99; TLI = .99; RMSEA = .044; SRMR = .030
  Grade 2: CFI = .96; TLI = .92; AIC = 6793.128; RMSEA = .11; SRMR = .051
  - Language (Story Recall, Understanding Directions)
  - Literacy (Reading Fluency, Letter-Word Identification, Passage Comprehension),
  - Math (Calculation, Math Fluency, Applied Problems)

Analysis: A z-score was created to compare grades and scores. Spearman's rho correlations reported, and analysis based on Bayesian statistics

Results

Grade 1	Grade 2
Correlations among Domains (Table 1)	
• Intradomain correlations were higher for Math and Literacy than Language • Interdomain correlations were also high. In fact, Language marks correlated higher with WJ III Literacy and Math than Language	
Grade vs. Test Scores across Domains (Table 2, Mean columns)	
• Language: WJ III (test score = 104) > marks (grade = 7.76 or B) by .24 points, BF_{10} > 100 • Literacy: Marks (grade = 7.68 or B) > WJ III (test score = 96) by .25 points, BF_{10} = 41	• Language : WJ III (test score = 104) > marks (grade = 8.11 or B) by .30 points, BF_{10} = 29
Gender Differences (Table 2, Gender columns)	
• No between gender differences • Girls: Literacy marks (7.78 or B) > WJ III (94), BF_{10} > 100 • Boys: Language WJ III (104) > marks (7.20 or B-), BF_{10} > 100	• Girls: Math marks (8.04 or B) > WJ III (95), BF_{10} = 9 • Boys: Language WJ III (103) > marks (7.93 or B), BF_{10} = 3
School Differences	
• Marks differed by school, but WJ III did not	• No school differences

Table 1. Correlations (Spearman's rho) among domains. Intradomain (similar domains) correlations are highlighted

	WJ III		
	Language	Literacy	Math
Grade 1			
Language	.44	.52	.50
Literacy	.37	.72	.53
Math	.48	.65	.65
Grade 2			
Language	.36	.52	.44
Literacy	.35	.67	.50
Math	.46	.64	.64

Table 2. Summary statistics reported as z-scores (SD) of report card marks and WJ III

	WJ III			Report card		
	Girls	Boys	Mean	Girls	Boys	Mean
Grade 1						
Language	0.22 (0.85)	0.28 (0.77)	0.26 (.81)	0.092 (0.79)	-0.062 (0.73)	0.015 (0.75)
Literacy	-0.41 (1.40)	-0.22 (1.44)	-0.28 (1.41)	0.039 (0.80)	0.40 (0.77)	-0.026 (0.78)
Math	-0.13 (0.95)	0.15 (0.90)	0.033 (.93)	0.0015 (0.76)	-0.018 (0.69)	-0.0046 (0.71)
Grade 2						
Language	0.29 (0.82)	0.19 (0.91)	0.25 (0.87)	0.052 (0.90)	-0.15 (0.78)	-0.046 (0.083)
Literacy	-0.15 (0.93)	-0.016 (0.85)	-0.060 (0.88)	0.059 (0.86)	-0.16 (0.87)	-0.050 (0.87)
Math	-0.36 (0.84)	-0.075 (0.83)	-0.20 (0.84)	-0.082 (0.85)	-0.084 (0.75)	-0.080 (1.98)

Conclusion

- Overlap between report cards and standardized tests, explained 15-52% of the variance, aligning with prior work[1, 2]
- Domain differences:
  - Intradomain correlations higher for Literacy and Math than Language
  - Discrepancies found in verbal subjects (Language and Literacy), with differences minimized for Literacy by grade 2
- *Within* gender differences: Female advantage in report card marks; male advantage on standardized tests, consistent with prior work[3]
- School differences: Differences were minimized by grade 2

Clinical Implications

- Teacher-assigned grades could be a more **valid measure** of student achievement for **literacy and math** than previously suggested, at least in young children. Evidence for language was less convincing

References

1. Bowers, A. J. (in press). Report card grades and educational outcomes. In S. M. Brookhart & T. R. Guskey (Eds.), What do we know about grading? Alexandria, VA: Association for Supervision and Curriculum Development (ASCD).
2. Ross, J. & Kostuch, L. (2011). Consistency of report card grades and external assessments in a Canadian province. Educational Assessment, Evaluation and Accountability, 23(2), 159–180. https://doi.org/10.1007/s11092-011-9117-3
3. Duckworth, A. & Seligman, M. E. P. (2006). Self-Discipline Gives Girls the Edge: Gender in Self-Discipline, Grades, and Achievement Test Scores. Journal of Educational Psychology, 98(1), 198–208. https://doi.org/10.1037/0022-0663.98.1.198

NSERC CRSNG

Disclosure Statement: All authors, no conflict of interest

Presented at SRCLD, June 2023