Journal of Modern Applied Statistical Methods

Volume 14 | Issue 1 Article 20

5-1-2015

Special Education Distributions and Analysis

Valerie Felder
Department of Education, State of Michigan, felderv@michigan.gov

Shlomo S. Sawilowsky
Wayne State University, shlomo@wayne.edu

Journal of Modern Applied Statistical Methods
May 2015, Vol. 14, No. 1, 238-252.

ISSN 1538 − 9472

Special Education Distributions and Analysis

Valerie Felder
Department of Education
State of Michigan

Shlomo S. Sawilowsky
Wayne State University
Detroit, MI

Micceri (1989) examined the distributional characteristics of 440 large sample general education achievement and psychometric measures. All the distributions were found to be statistically significantly different from the normal distribution. In this study, 395 special education datasets were examined. Although there were some normally distributed datasets, most were not, and some were markedly different in shape from those found by Micceri (1989). Implications for statistical testing and making special education policy decisions were given.

Keywords: Nonnormal data sets, statistical testing, special education

Special education distributions

Micceri (1989) conducted an investigation of the distributional characteristics of 440 large sample educational achievement and psychometric measures. The data sets were obtained from general education and the behavioral and social sciences, including ability tests, achievement tests, criterion or mastery level tests, psychometric measures, and pre- and post-intervention scores. All were found to be non-normal based on the Kolmogorov-Smirnov test with nominal $\alpha = 0.01$. Factors that contributed to a non-Gaussian error distribution in the population include (a) subpopulations within a target population, (b) ceiling/floor effects, and (c) variability in the items within a measure. This has implications in terms of statistical testing, because classical parametric tests require normality in order to maintain acceptable robustness and comparative power properties (Sawilowsky & Blair, 1992). If ignored, costly errors may occur in making policy decisions.

The prevalence of non-normally distributed data permeates many fields. Previous studies that demonstrated this include Bradley (1977, 1982), Hill and

Dr. Valerie Felder is Research Consultant with the Michigan Department of Education. Email her at: felderv@michigan.gov. Dr. Sawilowsky is a Professor in the Department of Evaluation and Research, College of Education. Email him at shlomo@wayne.edu.

Dixon (1982), Ito (1980), Pearson and Please (1975) and Tan (1982). However, they, as well as Micceri (1989), did not have special education and disability assessments as a focus.

Assessment of students in special education is frequently different than for students in general education, because often the focus is on process or progress as opposed to specific learning outcomes. This may include adaptive behavior, development, and screening. Adaptive behavior skills are those skills that are useful in daily functioning. Developmental skills pertain to fine- and gross-motor, communication and language, social, cognitive, and self-help skills. Screening helps find children who might be below the norm in different areas (Rosenberg, Westling, & McLeskey, 2010).

Purpose of the study

Given the paucity of representation of special education data sets in the studies mentioned above, the purpose of this study is to canvass that literature to determine the distributional shape commonly encountered. This will help inform the appropriate statistical method (i.e., parametric or nonparametric) to be used in measuring the progress of students in special education.

Methodology

The distribution patterns of special education data sets were obtained from published, peer-reviewed journal articles from the years of 2007-2011. In addition, research studies that focused on special education assessment were considered for inclusion. A Google Scholar search with the key terms "special education" and "data" returned 396,397 related publications.

To construct a confidence level of 95% and margin of error of ±5%, a sample size of 384 data sets was needed from that population. It was estimated a return response rate of 25% was needed to accommodate lack of responses, and therefore 1,540 survey requests were made from selected authors of those published studies. Assessment data sets were also solicited from various state departments of education. Requests were made via email and telephone. The request included instructions to de-identify student information. Initial contact via email and phone was made from October - December, 2012. Follow-up phone calls and email messages were made in January, 2013.

Criteria for inclusion

Potential studies were reviewed to determine if the instrument used to collect data was supported by adequate reliability and validity information. However, there was no preset type or minimum reliability index or validity methodology required for inclusion.

Reliability is "the consistency that a test measures whatever it measures" (Sawilowsky, 2007, p. 516). As noted by Sawilowsky (2000), reliability is a psychometric property of a test. If the test produces similar results under consistent conditions then it is considered reliable. There were different types of reliability information obtained:

- Internal consistency, which is the extent items on an instrument relate to each other.
- Test-retest, which is the consistency over time (i.e., stability) of an instrument.
- Inter-rater reliability, which is the degree of agreement among raters.

Validity is "the degree that a test measures what it purports to measure (Sawilowsky, 2007, p. 166). There are different types of validity, including content-related validity, construct validity, and predictive validity (Cicchetti, 1994):

- Content-related validity, which is how well the content of the test relates to what is being assessed.
- Construct validity. "A construct is a fiction that is used to explain reality" (Cuzzocrea & Sawilowsky, 2009, p. 215), such as aptitude, intelligence, or self-determination. Hence, construct validity is the degree that a test measures that fiction used to explain reality.
- Predictive validity, which is the extent a test predicts some criterion measure.

Results

There were 744 authors contacted via email. Note that many authors had obtained multiple data sets in their study, exceeding the 1,540 data set requirement. Follow-up phone calls and emails were conducted where necessary after 3 months. There were n = 333 data sets collected from journal article authors, as compiled in

Table 1. In addition, academic achievement special education assessment test scores were requested from state education departments. Twenty four state departments of education, randomly selected, were contacted from which an additional $n = 62$ data sets were obtained from Alaska, Florida, Michigan, Minnesota, Missouri, and South Carolina, as compiled in Table 2. Thus, there were a total $N = 395$ data sets. Based on an estimated accessible population, the obtained sample size yielded a confidence level of 95% with a ±4.25% margin of error.

Table 1. Summary of Canvassed Authors (744) and Data Sets (4,362)

	Total	Total % of Articles
Acceptable Reliability	1760	40.30%
Acceptable Validity	1600	36.70%
Acceptable Articles*	1002	23.00%
Acceptable Data Sets	333	7.60%

*Note: An acceptable article required acceptable reliability and validity evidence.

Table 2. Data Sets from State Departments of Education

Florida	16	Minnesota	19
South Carolina	8	Alaska	15
Missouri	3	Michigan	1
		Total	**62**

Cronbach alpha coefficients for the instruments used to obtain these data sets ranged from .70 to .93. Test-retest reliability coefficients ranged from .65 to .97, and alternate-forms reliability ranged from .91 to .92. Concurrent validity indices ranged from .70 to .89, and predictive validity indices ranged from .65 to .86. (The author of one study used Item response theory (IRT) in a measurement model (i.e., Rasch one-parameter logistic (1PL) partial credit model for polytomous scoring).

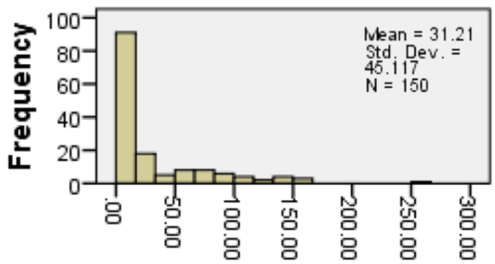
Distribution shapes

The histograms was analyzed and categorized. Histograms that resembled Micceri's (1986) distributions were named accordingly. Histograms that did not resemble Micceri's distributions were given a name based on the shape of each distribution. Figure 1 contains typical shapes obtained from the data sets. The types of distributions and the percentage of each distribution that were collected are indicated in Table 3. There were 258 (65.31%) special education data sets that were different and 137 (34.67%) similar to Micceri's (1989) shapes.

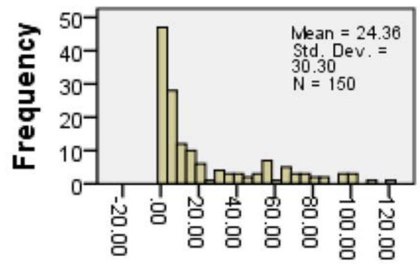
The data sets were also analyzed for normality and compared with Micceri's data sets. Based on the Kolmogorov-Smirnov and Shapiro-Wilks tests, 318 (81%) data sets were non-normally distributed and 77 (19%) data sets were normally distributed. Recall that Micceri (1986, 1989) found 100% of the distributions to be significantly non-normally distributed at the $\alpha = .01$ level. There were 19 out of 440 distributions, or 4.3%, that were considered reasonable approximations to the Gaussian distribution only in the sense that they were smooth symmetric with light tails. As compared with Micceri's (1986, 1989) results, this study shows special education assessment data sets were somewhat more likely to be normally distributed, but the number of different data sets shapes was higher than those found by Micceri (1986, 1989).

Table 3. Type, Number, and Percentage and Distribution Shapes

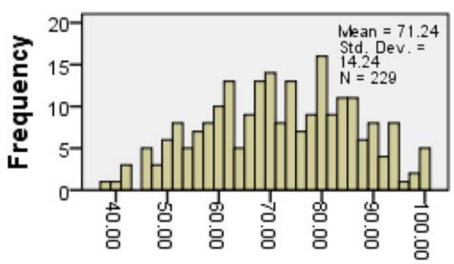
Type of Distribution	Number	Percentage
Extreme Bimodality	106	26.84%
Equimodal	96	24.30%
Unimodal and Smooth	79	20.00%
Bimodal and Smooth	31	7.85%
Slight Asymmetry	25	6.33%
Multimodal and Lumpy	19	4.81%
Unimodal and Slightly Smooth	10	2.53%
Extreme Asymmetry	6	1.52%
Slightly Asymmetric and Digit Preference	6	1.52%
Digit Preference	4	1.01%
Unimodal and Slightly Lumpy	4	1.01%
Equimodal and Symmetric	3	0.76%
Extreme Mass at Zero	2	0.51%
Mass at Zero	1	0.25%
Smooth Symmetric	1	0.25%
Equimodal and Slight Asymmetry	1	0.25%
Slightly Smooth and Symmetric	1	0.25%



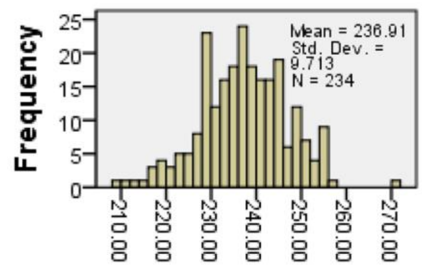
Dataset 1. Skew = 2.090, PATM Pre-test



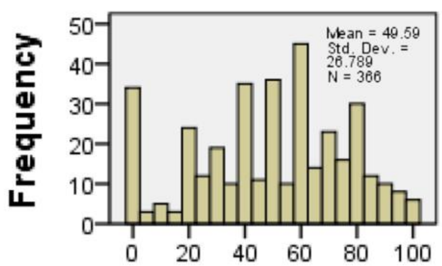
Dataset 2. Skew = 1.340, PATM Post-test



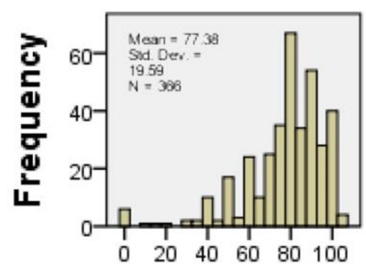
Dataset 3. Skew = -.111,
CAAVES Reading Assessment



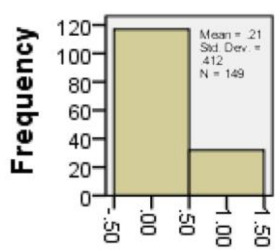
Dataset 4. Skew = -.080,
CAAVES Math Assessment



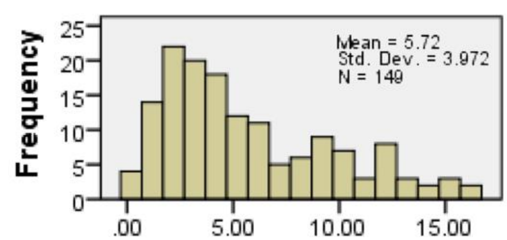
Dataset 5. Skew = -.246, Pre-test
Tomlinson's differentiated instruction
strategies adapted assessment



Dataset 6. Skew = -1.543, Post-test
Tomlinson's differentiated instruction
strategies adapted assessment



Dataset 7. Skew = 1.291
Grade 2, Dyslexiacriteria, Spring



Dataset 8. Skew = .896
Grade 1, Fluency Word Recognition, Fall

Figure 1. Special Education Data Sets

Discussion

There were more classifications of special education data sets as extreme bimodality (n = 106, uni-modal, and smooth and equimodal than found in other disciplines. There were 106 extreme bimodality distributions and 57%, or 60 data sets, were non-normal. There were 46 distributions that were normal. There were 79 unimodal and smooth distributions and 29%, or 23 data sets, were non-normal. The remaining category, which had a large amount of distributions, is the equimodal category. There were 96 distributions and 70%, or 67, were non-normal. Thirty percent of the equimodal distributions were normally distributed based on the Kolmogorov-Smirnov and/or Shapiro-Wilks normality tests.

These data sets that were non-normally shaped pertained to curriculum-based assessments in writing, alternative assessments, applied problem solving, calculations, mathematics operations, reading, letter-word identification, segmenting words, and letter naming. Assessments of achievement, and fine- and gross-motor skills tended to be shaped normally.

In terms of policy, it is important to consider statistical robustness and comparative power when analyzing special education assessments. The results of this survey confirm the importance of considering nonparametric alternatives to parametric methods. As has been conducted throughout the Monte Carlo literature of the past century for data in many disciplines (e.g., general education, psychology, medicine, nursing), a study is warrant to determine the extent to which robustness and power of parametric tests may be compromised when analyzing special education data.

The new special education data shapes in this study may overlap with Micceri’s (1989) data shapes. Due to the small sample size of the special education data sets, some of the shapes were different than Micceri’s data shapes, but a larger sample sizes may show the data converges to one of Micceri’s shapes.

For example, consider the data sets from the Florida Alternate Assessment. They were separated by grade level and a distribution was created for each data set, because the achievement of students in special education is measured based on a set of academic standards for each grade level. However, if the sample size is increased by combining a single grade with all grade levels, the resulting shape, identified by Micceri (1989) as a discrete mass at zero with gape, will result, as noted in Figure 2.

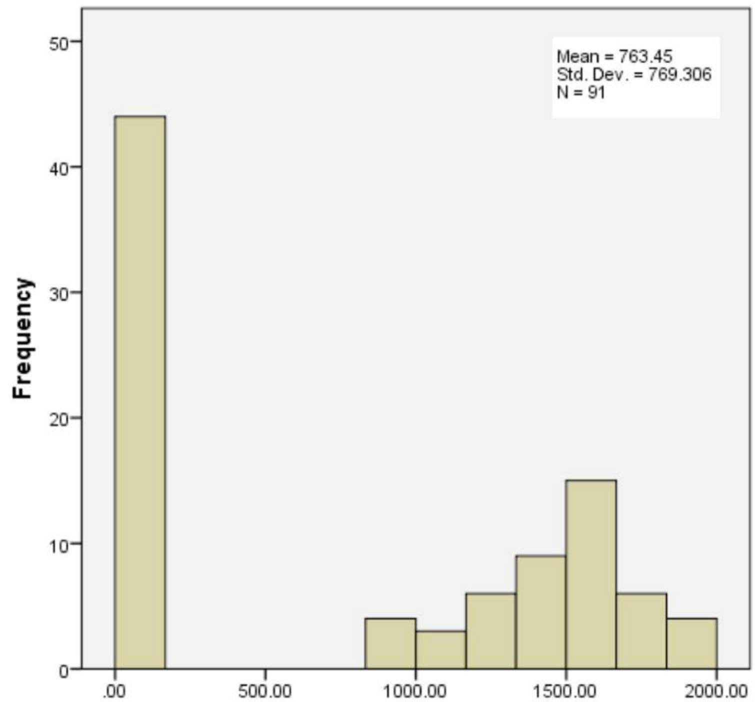


Figure 2. Concatenated Florida Alternate Assessment Special Education Data Set for All Grade Levels

---

In summary, Micceri's (1989) seminal article on 440 real data sets from general education achievement and psychometric constructs, shockingly, found them all to be non-normally distributed. This led to a major overhaul in techniques for analyzing quantitative data, as is known in the statistical literature, in those fields. Unfortunately, progress in revising and updating statistical strategies into other fields has been slow. Workers have the tendency to hold fast to techniques learned many years prior in graduate school, and furthermore, with the uptick in qualitative research, the lessons learned from Micceri (1989) obtain little voice until such surveys are replicated in their fields. On the basis of 395 special education data sets obtained in this study, differences from Micceri's (1989) rubric were noted, particularly the emergence of new non-normal distribution shapes. We believe this survey will help motivate quantitative workers in the special education field update their data analytic choices.

References

Aldridge, J. (2008). Narrowing the gaps for special-needs students. *Childhood Education, 84*(3), 182.

Barkley, R.A. (1997). *ADHD and the nature of self-control.* New York: Guilford Press.

Biancarosa, G. & Snow, C.E. (2004). *Reading next: A vision for action and research in middle and high school literacy. A report to the Carnegie Corporation of New York*. Washington, DC: Alliance for Excellent Education.

Bluman, A. (2007). *Elementary statistics. A step by step approach*. New York, NY: McGraw-Hill Higher Education.

Bradley, J. W., (1977). A common situation conducive to bizarre distribution shapes. *The American Statistician*, *31*(4), 147-150. doi:10.2307/2683535

Bradley, J. W. (1982). The-insidious L-shaped distribution. *Bulletin of the Psychonomic Society*, *20*, 85-88. doi:10.3758/BF03330089

Browder, D., Wakeman, S., & Flowers, C. (2006). Assessment of progress in the general curriculum for students with disabilities. *Theory Into Practice, 45*(3), 249-259.

Caffrey, E., Fuchs, D., & Fuchs, L. S. (2008). The predictive validity of dynamic assessment: A review. *The Journal of Special Education*, *41*(4), 254-270. doi:10.1177/0022466907310366

Calhoon, M.B., Sandow, A., & Hunter, C. (2010). Reorganizing the instructional reading components: could there be a better way to design remedial reading programs to maximize middle school students with reading disabilities' response to treatment? *Annals of Dyslexia*, *60*, 57-85. doi:10.1007/s11881-009-0033-x

Cicchetti, D. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6*(4), 284-290. doi:10.1037/1040-3590.6.4.284

Clarke, B., Baker, S. & Smolkowski, K. (2008). An analysis of early numeracy curriculum-based measurement: Examining the role of growth in student outcomes. *Remedial and Special Education*, *29*(1), 46-57. doi:10.1177/0741932507309694

Cuzzocrea, J., & Sawilowsky, S. S. (2009). Robustness to non-independence and power of the I test for trend in construct validity. *Journal of Modern Applied*

Statistical Methods, *8*(1), 215-225. Available at: http://digitalcommons.wayne.edu/jmasm/vol8/iss1/19

Eckes, S., Swando, J. (2009). Special education subgroups under NCLB: Issues to consider. *Teachers College Record, 111*(11), 2479-2504.

Elbaum, B. (2007). Effects of an oral testing accommodation on the mathematics performance of secondary students with and without learning disabilities. *The Journal of Special Education, 40*(4), 218-219. doi:10.1177/00224669070400040301

Fore, C., Boon, R. T., Burke, M. D. & Martin, C. (2009). Validating curriculum-based measurement for students with emotional and behavioral disorders in middle school. *Assessment for Effective Intervention, 34*(2), 67-73. doi:10.1177/1534508407313234

Fuchs, D., Mock, D., Morgan, P. L., & Young, C. L. (2003). Responsiveness-to-intervention: Definitions, evidence, and implications for the learning disabilities construct. *Learning Disabilities Research & Practice, 18*(3), 157-171. doi:10.1111/1540-5826.00072

Graham, S., & Harris, K. R. (1989). Components analysis of cognitive strategy instruction: Effects on learning disabled students' compositions and self-efficacy. *Journal of Educational Psychology, 81*, 353-361. doi:10.1037/0022-0663.81.3.353

Hardman, M. L., Drew, C. J., & Egan, M. W. (2002). *Human exceptionality: Society, school, and family*. Boston: Allyn & Bacon.

Heckaman, K., Conroy, M., Fox, J., & Chait, A. (2000). Functional assessment-based intervention research on students with or at risk for emotional and behavioral disorders in school settings. *Behavioral Disorders, 25*(3), 196-210.

Helwig, R. & Tindal, G. (2003). An experimental analysis of accommodation decisions on large-scale mathematics tests. *Council for Exceptional Children, 69*(2), 211-225. doi:10.1177/001440290306900206

Hill, M., & Dixon, W. J. (1982). Robustness in real life: A study of clinical laboratory data. *Biometrics*, *38*(2), 377-396. doi:10.2307/2530452

Hosp, J. L., Howell, K. W. & Hosp, M. K. (2003). Characteristics of behavior rating scales: Implications for practice in assessment and behavioral support. *Journal of Positive Behavior Interventions*, *5*(4), 201. doi:10.1177/10983007030050040301

Hughes, C.A., Schumaker, J.B., & Deshler, D.D. (2005). *The essay test-taking strategy*. Lawrence, KS: Edge Enterprises, Inc.

Ito, P. K. (1980). Robustness of ANOVA and MANOVA test procedures. In P. R. Krishnaiah (Ed.), *Handbook of Statistics* (Vol. 6, p. 199-236). Amsterdam: North-Holland.

Jacobson, L. & Reid, R., (2010). Improving the persuasive essay writing of high school students with ADHD. *Exceptional Children, 76*(2), 157.

Katz, L. A., Stone, C. A., Carlisle, J. F., Corey, D. L. & Zeng, J. (2008). Initial progress of children identified with disabilities in Michigan's Reading First schools. *Exceptional Children, 74*(2), 235-256. doi:10.1177/001440290807400206

Kohl, F. L., McLaughlin, M. J., & Nagle, K. (2006). Alternate achievement standards and assessments: A description investigation of 16 states. *Exceptional Children, 73*(1), 107-123. doi:10.1177/001440290607300106

Kover, S. T. & Atwood, A. K. (2013). Establishing equivalence: Methodological progress in group-matching design and analysis. *American Journal on Intellectual and Developmental Disabilities, 118*(1), 3-15. doi:10.1352/1944-7558-118.1.3

Lane, K. L., Carter, E. W., Pierson, M. R. & Glaeser, B. C. (2006). Academic, social, and behavioral characteristics of high school students with emotional disturbances or learning disabilities. *Journal of Emotional and Behavioral Disorders, 14*(2), 108-117. doi:10.1177/10634266060140020101

Mayes, S. D., Calhoun, S. L., & Crowell, E. W. (2000). Learning disabilities and ADHD: Overlapping spectrum disorders. *Journal of Learning Disabilities, 33*(5), 417-424. doi:10.1177/002221940003300502

McConaughy, S., & Ritter, D. (2002). Best practices in multidimensional assessment of emotional or behavioral disorders. *Best practices in school psychology IV* (pp. 1303-1336). Bethesda, MD: National Association of School Psychologists.

McGinnis, E., Kiraly, J., & Smith, C. R. (1984). The types of data used in identifying public school students as behaviorally disordered. *Behavioral Disorders, 9*(4), 239-246.

Mertens, D. M. & McLaughlin, J. A. (2004). *Research and evaluation methods in special education*. (pp. 170-178). Thousand Oaks, CA: Sage Publication Ltd.

Mervis, C. B. & Klein-Tasman, B. P. (2004). Methodological issues in group-matching designs: Alpha levels for control variable comparisons and measurement characteristics of control and target variables. *Journal of Autism and*

Developmental Disorders, 34(1), 7-17. doi:10.1023/B:JADD.0000018069.69562.b8

Micceri, T. (1986, November). A futile search for that statistical chimera of normality. Paper presented at the annual meeting of the Florida Educational Research Association, Tampa, FL.

Micceri, T. (1989). The unicorn, the normal curve, and other improbable creatures. *Psychological Bulletin, 105*(1), 156-166. doi:10.1037/0033-2909.105.1.156

Mosteller, F., & Tukey, J. (1977). *Data Analysis and Regression*, (pp. 55). Reading, MA: Addison-Wesley Publishing Company.

No Child Left Behind (NCLB) Act of 2001, Pub. L. No. 107-110, § 115, Stat. 1425 (2002).

Olson, L. (2000). Worries of a standards “backlash” grow. *Education Week*, *30*, 1-13.

Pearson, E. S., & Please, N. W. (1975). Relation between the shape of population distribution and the robustness of four simple test statistics. *Biometrika*, *62*(2), 223-241. doi:10.1093/biomet/62.2.223

Rosenberg, M., Westling, D. & McLeskey, J. (2010). *Special education for today’s teachers: An introduction*. Upper Saddle River, NJ: Pearson Education.

Runyon, R., Coleman, K & Pittenger, D. (2000). *Fundamentals of behavioral statistics*. New York, NY: McGraw-Hill Higher Education.

Salahu-Din, D., Persky, H., & Miller, J. (2008). *The nation’s report card: Writing 2007* (NCES 2008-468). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education, Washington, DC.

Sawilowsky, S. S. (2000). Psychometric versus datametrics: Comment on Vacha-Haase’s ‘Reliability Generalization’ method and some EPM editorial policies. *Educational and Psychological Measurement*, *60*(2), 157-173. doi:10.1177/00131640021970439

Sawilowsky, S. S. (2007). KR-20 and KR-21. In (N. J. Salkind, Ed.), *Encyclopedia of Measurement and Statistics*. Thousand Oaks, CA: Sage, p. 516-519.

Sawilowsky, S. S. & Blair, R. C. (1992). A more realistic look at the robustness and type II error properties of the t test departures from population normality. *Psychological Bulletin. 111*(2), 352-360. doi:10.1037/0033-2909.111.2.352

Sawilowsky, S. S., Blair, R.C., & Micceri, T. (1990). REALPOPS.LIB: a PC Fortran library of eight real distributions in psychology and education. *Psychometrika, 55*(4), 729.

Sawilowsky, S. S. & Fahoome, G.F. (2003). *Statistics through Monte Carlo simulation with Fortran.* Michigan: JMASM, Inc.

Silberglitt, B. & Hintze, J. M., (2007). How Much Growth Can We Expect? A Conditional Analysis of R-CBM Rates by Level of Performance. *Exceptional Children, 74*(1), 71. doi:10.1177/001440290707400104

Tan, W. Y. (1982). Sampling distributions and robustness of t, F and variance-ratio in two samples and ANOVA models with respect to departure from normality. *Communications in Statistics*. *A11*, 2485-2511.

Therrien, W. J., Hughes, C., Kapelski, C. & Mokhtari, K. (2009). Effectiveness of a test-taking strategy on achievement in essay tests for students with learning disabilities. *Journal of Learning Disabilities 42*(1), 14-23. doi:10.1177/0022219408326218

Thorndike, R., Hagen, E. (1986). *The Stanford-Binet intelligence scale, fourth edition: Guide for administering and scoring*. Chicago, IL: Riverside Publishing Co.

Tindal, G. & Fuchs, L.S. (1999). *A summary of research on test change: An empirical basis for defining accommodation.* Lexington: University of Kentucky, Mid-South Regional Resource Center.

Tomlinson, C. A. (1995). Deciding to differentiate instruction in the middle school: One school's journey. *Gifted Child Quarterly, 39*(2), 77-114. doi:10.1177/001698629503900204

Tukey, J.W. (1977). *Exploratory data analysis*. (pp. 63) Reading, MA: Addison-Wesley Publishing Company.

Wechsler, D. (1991). *The Wechsler intelligence scale for children*. San Antonio, TX: Psychological Corporation.

Woodcock, R., Mather, N., McGrew, K. (2001). *Woodcock-Johnson III Tests of Cognitive Abilities Examiner's Manual*. Itasca: Riverside.

Ysseldyke, J., Thurlow, M., Langenfield, K., Nelson, J.R., Teelucksing, E., & Seyfarth, A. (1998). *Educational results for students with disabilities: What do the data tell us?* (Tech. Rep. No. 23). Minneapolis: University of Minnesota, National Center of Educational Outcomes.

Appendix: Journals used in the survey

Journals marked with an "*" were used in the survey. The data is available from the first author of this study.

**American Annals of Deaf*
**American Educational Research Journal*
**American Journal on Intellectual and Developmental Disabilities*
**Annals of Dyslexia*
**Applied Measurement in Education*
Australasian Journal of Special Education
Behavioral Disorders
British Journal of Special Education
Career Development for Exceptional Individuals
Child Development Perspectives
Developmental Psychology
Early Childhood Research Quarterly
Education and Training in Mental Retardation and Developmental Disabilities
**Education and Treatment of Children*
Educational Assessment
**Educational and Psychological Measurement*
**Elementary School Journal*
**Exceptional Children*
**Exceptionality: A Research Journal*
International Journal of Disability
**Journal of Adolescent and Adult Literacy*
**Journal of Applied Behavior Analysis*
Journal of Applied Developmental Psychology
Journal of the Association for Persons with Severe Handicaps
Journal of Attention Disorders
**Journal of Autism and Developmental Disorders*
Journal of Deaf Studies and Deaf Education
**Journal of Disability Policy Studies*
**Journal of Early Intervention*
Journal of Educational Psychology
Journal of Educational and Behavioral Statistics
Journal of Educational Measurement
**Journal of Emotional and Behavioral Disorders*

Journal of Intellectual Disability Research
**Journal of the International Association of Special Education*
**Journal of Learning Disabilities*
Journal of Policy and Practice in Intellectual Disabilities
**Journal of Positive Behavior Interventions*
**Journal of Psychoeducational Assessment*
Journal of Research and Development in Education
**Journal of School Psychology*
**Journal of Special Education*
Journal of Speech and Hearing Research
**Journal of Visual Impairment and Blindness*
**Learning and Individual Differences*
**Learning Disability Quarterly*
**Learning Disabilities Research and Practice*
Mental Retardation
Peabody Journal of Education
**Preventing School Failure*
**Psychology in the Schools*
**Reading and Writing*
Reading Psychology
Reading Research Quarterly
**Remedial and Special Education*
Research in Developmental Disabilities
**Review of Educational Research*
**School Psychology Quarterly*
**School Psychology Review*
Teachers College Record
Teaching Exceptional Children
**Volta Review*