

# Analyzing Relationships Between School Libraries and Academic Achievement

Keith Curry Lance

Director

Library Research Service

Colorado State Library &

University of Denver

# Outline

- Background
- Research questions
- Data types & sources
- Statistical concepts & techniques
- “Success stories”

# Background

- A half century of previous school library research
- The political climate of education & libraries in the late '80's
- The School Match Incident
- The first Colorado study
- The political climate of education & libraries in the late '90's
- The second Colorado study & successor studies by Lance, Rodney & Hamilton-Pennell
- Successor studies by others

# Research Questions

- Are students more likely to “pass” tests if they have a school library than if they don't?
- Are students likely to score higher on tests if they have a school library than if they don't?
- As the school library improves, do test scores rise?
- How are different qualities of school libraries, schools, and communities related to each other?
- Do school libraries & test scores improve together, even when other school & community conditions are taken into account?

# Types of Data

- Nominal
  - Categories
  - No necessary quantitative dimension
  - ***Pass/fail, library/no library***
- Ordinal

---

  - Degrees of difference
  - No equal intervals
  - Zero is just a code
  - Usually limited number of values
- Interval/Ratio
  - Equal intervals
  - True zero (have none of something)
  - Usually large number of values
  - ***Weekly hours of librarian staffing, test scores***

# Types of Variables

- Dependent variable
  - “The effect” in a cause-and-effect relationship
  - **Reading test scores** used to “operationalize” concept of academic achievement
- Independent variables
  - “The causes” in a cause-and-effect relationship
  - Characteristics of **school libraries**, schools & communities
    - “Treatment” or predictor variables
    - “Control” variables

# State Test Scores

- Standards-based tests v. "standardized" tests
- Test scores, % proficient & above v. % "passed" v. percentile rankings
- Reading scores are key
- Difference between existing & available data (actually acquiring data file in a usable format & on a timely schedule)

# Other Data Sources

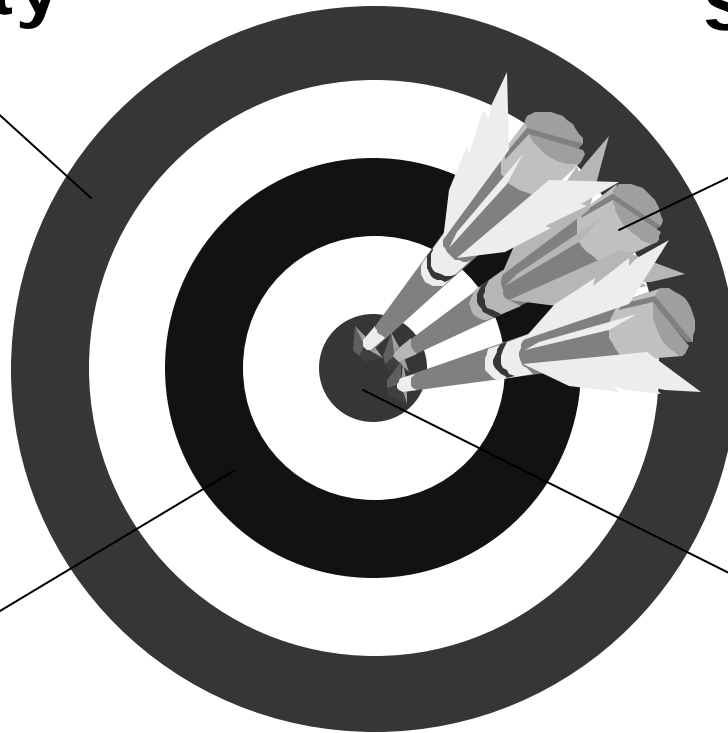
Data items	Source
<b>Library</b> <ul style="list-style-type: none"><li>▪ School library hours</li><li>▪ Staffing &amp; staff activities</li><li>▪ Collections, technology &amp; usage</li><li>▪ Expenditures</li></ul>	<b>Survey</b>
<b>School</b> <ul style="list-style-type: none"><li>• District expenditures per pupil</li><li>• Teacher-pupil ratio</li><li>• Teacher education, experience &amp; salaries</li></ul>	<b>State ED dept.</b>
<b>Community</b> <ul style="list-style-type: none"><li>▪ Students by NSLP status (poverty), race/ethnicity</li><li>▪ Adult educational attainment</li></ul>	<b>State ED dept., census</b>



# The Data Model

**Community**

**School library**



**School**

**Test scores**

# Experiment v. Statistical Analysis

- Experiment
  - Older studies
  - Smaller samples
  - More precise units of analysis (student)
  - More control over independent variables
  - Matching issues
  - Easier to explain, communicate
- Statistical analysis
  - Newer studies
  - Larger samples
  - Less precise units of analysis (school)
  - Less control over independent variables
  - Data availability issues
  - More precise measurement of effects

# Statistical Significance

- Likelihood the sample results are representative of the universe under study
- Most common notation:
  - $p < .05$ ,  $< .01$ ,  $< .001$
- Difference between statistical significance & confidence interval (i.e., margin of error)
- No statistical test of SUBSTANTIVE significance (i.e., how important is this?)

# Statistical Analysis Software

## ■ Market leaders:

- SPSS: Statistical Package for the Social Sciences
- SAS: Statistical Analysis Software

## ■ Software Issues:

- Available statistical techniques: correlation, comparison of means, factor analysis, regression
- Data management features: sort, sample, compute, recode, if
- Case limits (maximum number of cases allowed)
- Cost (education discount)

# Cross-tabulation

- ***Are students more likely to pass tests if they have a school library than if they don't?***
- Two nominal variables or one nominal and one ordinal (small range)
- Pass/fail on tests, librarian/no librarian
- Turning interval or ratio variables into nominal or ordinal ones
- Chi-square ( $X^2$ ) indicates statistical significance

# Test Scores by Time Spent Teaching Information Literacy: Alaska, 1998

Time on information literacy	Average & above scores	Below average scores	Total
<b>Median &amp; above</b>	56 82%	12 18%	68 100%
<b>Below median</b>	33 53%	29 47%	62 100%
<b>Total</b>	89 69%	41 31%	130 100%

Chi-square = 12.743,  $p < .001$

# Comparison of Means

- ***Are students likely to score higher on tests if they have a school library than if they don't?***
- One nominal (2 dimensions), one interval or ratio variable
- Pass/fail on test, hours of librarian staffing
- Generates means (averages) for 2 groups
- Levene's test indicates equality (or inequality) of variances between groups
- t test indicates statistical significance of difference between groups

# Student Visits for Information Literacy Instruction for Higher & Lower Scoring Elementary Schools: Alaska, 1998

<b>Schools by reading scores</b>	<b>Student visits for IL instruction per 100 students</b>
<b>High-achieving schools</b>	81
<b>Low-achieving schools</b>	43

$t = 3.963, p < .001$



# Correlation (r)

- ***As the school library improves, do test scores rise?***
- Two interval or ratio variables
- LM expenditures per student, volumes per student
- Pearson's product-moment correlation (r)
- Expressed in decimal form
  - Perfect correlation = 1.00
  - + & - indicate positive & negative relationships  
(+ = both rise or fall, - = one rises, other falls)
  - $r = .60-.80$  v.  $.80+$  & factor analysis
  - $r^2$  = percent of variation explained

# Bivariate Correlation Coefficients for LM Program Development Variables: Colorado Middle Schools, 1999

LM Development variables	1	2	3	4	5	6
<b>1. LMS hours/100</b>	1.00					
<b>2. Total hours/100</b>	.696	1.00				
<b>3. Volumes/student</b>	.695	.703	1.00			
<b>4. E-reference/100</b>	.668	.779	.668	1.00		
<b>5. Subscriptions/100</b>	.701	.646	.680	.640	1.00	
<b>6. LM exp. per student</b>	.788	.790	.837	.755	.802	1.00

$p < .001$

# Factor Analysis

- ***How are different qualities of school libraries (schools, communities) related to each other?***
- Analyzes relationships between and among variables
- Key statistics:
  - Percent of variance explained
  - Factor loadings
  - Factor scores
    - Allow mixing items on different scales
    - Data reduction technique

# Factor Analysis of LM Program Development Variables: Colorado Middle Schools, 1999

---

<b>LM Program Development Variable</b>	<b>Factor Loading</b>
<b>LMS hours/100 students</b>	<b>.863</b>
<b>Total hours/100 students</b>	<b>.877</b>
<b>Volumes per student</b>	<b>.874</b>
<b>E-reference/100</b>	<b>.863</b>
<b>Subscriptions/100</b>	<b>.847</b>
<b>LM exp. per student</b>	<b>.949</b>

Initial eigenvalue = 4.638, 77% variance explained

# Regression (R, R<sup>2</sup>)

- ***Do school libraries & test scores improve together, even when other conditions are taken into account?***
- Need to conduct correlation—and often factor—analyses first
- Linear regression
- Stepwise regression
- Multiple R, R square & R square change
- Standardized beta coefficients (indicate positive or negative direction)
- Included v. excluded variables

# Regression Analysis of 4<sup>th</sup> Grade Scores with LM, School, & Community Predictors: Colorado, 1999

---

<b>Predictor added</b>	<b>R</b>	<b>R Square</b>	<b>R Square Change</b>	<b>Beta</b>
% Poor	.638	.407	.407	-.471
<b>LM Factor</b>	<b>.694</b>	<b>.482</b>	<b>.075</b>	<b>.238</b>
% Minority	.709	.502	.021	-.225

$p < .01$

Excluded variables: teacher-pupil ratio, per pupil expenditures, teacher characteristics

# “Success Stories”

- Even the strongest statistical evidence can be made more persuasive by compelling “success stories”

# Characteristics of Good “Success Stories”

- One clear point: value of librarian as teacher (technology coordinator, in-service provider)
- Variety of voices: librarians, students, teachers, principals, parents
- “Short & sweet”
- A quotable quote