D*mned Lies and Statistics?: Critical Consumerism of Large-Scale Research

Amanda L. Sullivan, PhD
Mollie Weeks, MA
Tara Kulkarni, MA
University of Minnesota

Scott Graves, PhD
The Ohio State University
Jamilia Blake, PhD
Texas A&M University

Disclaimer

The contents of this report were developed under a grant from the U.S. Department of Education, # H325D160016. However, those contents do not necessarily represent the policy of the U.S. Department of Education, and you should not assume endorsement by the Federal Government. Project Officer, Sarah J. Allen, Ph.D.
WELCOME AND INTRODUCTIONS

Learner Objectives

Understand common methodological features, including strengths and limitations, of secondary analysis of large-scale data

Identify key conceptual, ethical, and methodological considerations for critical consumerism of published research

Apply guiding considerations to critically evaluate applicability of published research to practice
BACKGROUND AND CONTEXT

The Visibility of Secondary Analysis

Aggressive Policing and the Educational Performance of Minority Youth

Can Maternal Fluoride Consumption During Pregnancy Lower Children's Intelligence?

Minorities Are Disproportionately Underrepresented in Special Education: Longitudinal Evidence Across Five Disability Conditions

(Barnum, 2019; Green et al., 2019; Harris, 2019; Legewie & Fagan, 2019; Morgan & Farkas, 2015; Morgan et al., 2015)
Our Responsibilities as School Psychologists

- School Psychologists have an ethical responsibility to contribute to the school psychology knowledgebase (Principle IV.5).
  - Conducting and disseminating research;
  - Grounding research methods in sound practice;
  - Not fabricating or falsifying data;
  - Making data and other information available to other researchers; and,
  - Correcting errors when made aware.

(NASP, 2010)

PART 1: INTRODUCTION TO LARGE-SCALE RESEARCH METHODS
What is Secondary Analysis?

(Boslaugh, 2007; Smith, 2008)

Key Words and Definitions

- **Sample Size**: the number of participants in a study drawn from a population of interest. The more participants, the greater the power.
- **Nationally-Representative**: a sample designed to approximate a target population on a national level.
- **Weighting**: corrects for disproportionate representation in a sample (i.e., makes a sample more like the population of interest).
- **p-value**: the likelihood we would obtain the size of our effect if a null hypothesis was true.
- **Effect Size**: the magnitude of an effect (i.e., its practical significance). Large samples may have significant p-values and negligible effect sizes.

(Howell, 2013; Smith et al., 2011)
Advantages and Disadvantages of Secondary Analysis

<table>
<thead>
<tr>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
</table>
| - Efficient  
- Macro-ecological questions  
- Large samples  
- Longitudinal data  
- Under-represented populations  
- Access for scholars  
- Replicability and reproducibility  
- Cross-cultural analysis | - Labor-intensive and complex  
- Consideration of method and variable conceptualization  
- Match between questions and data  
- Age of data  
- No control over participants  
- No control over measures  
- Sources of error by informant |

(Andersen et al., 2011; Greenhoot & Dowsett, 2012; Silberzhan et al., 2018; Smith, 2008)

WHAT MAKES A SECONDARY DATA ANALYSIS HIGH QUALITY?

**Avoids Causal Language**

- Influences
- Effect
- Bolster
- Enhance
- Improve
- Leads to Decreases/Increases

*Associated
*Related
*Higher probability
*Less/more likely
*Higher/lower risk

https://cpb-us.w2.wpmucdn.com/u.osu.edu/dist/c/2883/files/2015/02/Causal-language-infographic-17ae0x0.png

**Explicit Sampling and Sample Characteristics**

2.1. Data and analytical sample

We analyzed data from the Early Childhood Longitudinal Study-Kindergarten Class of 1998–1999 (ECLS-K), a data set collected and administered by the National Center for Education Statistics (NCES). The ECLS-K is a nationally representative, longitudinal cohort of children who were followed from kindergarten entry through the end of eighth grade. A multistage, probability sampling design was employed in which approximately equal public and private schools were sampled from 100 geographic regions. About 16,000 kindergarten students were recruited from each school. Data collection continued throughout elementary and middle school. Surveys were administered at the beginning of third, fifth, and eighth grade. The NCES provided sampling weights to account for attrition across sample waves. The study’s sample contained information on 67,460 students who participated in the ECLS-K from kindergarten through eighth grade. As described below, we used multiple imputation to account for missing data for this sample.
Measures are Best Available Given Research Question

Parent recall of number of suspensions

OCR Data
SWIS Data

Analytic Sample is Explicit and Justified

<table>
<thead>
<tr>
<th>Model 1: Baseline</th>
<th>Model 2: Inclusion of Prior Problem Behavior (PPB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N = 4,101</td>
<td>N = 2,737</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model 1: Within Whites</th>
<th>Model 2: Within Blacks</th>
</tr>
</thead>
<tbody>
<tr>
<td>N = 2,448</td>
<td>N = 289</td>
</tr>
</tbody>
</table>

Sample
Data for the present study came from the Early Childhood Longitudinal Study, Kindergarten Class of 1998-1999 dataset (ECLS-K). The ECLS-K includes over 21,000 children and is the largest nationally representative sample of kindergartners, parents, teachers, and both public and private schools in the United States. The ECLS-K is sponsored by the U.S. Department of Education and
Limitations of Generalizability Explicit

We find little evidence that U.S. elementary and middle schools are suspending SWD more frequently than otherwise similar students without disabilities. Our findings do not support policies or reports (Kim et al., 2010; Losonc & Gillespie, 2012; OCR, 2014) that having a disability is itself associated with an increased risk of entering the school-to-prison pipeline via school suspension (Behrken et al., 2014; Mowen & Brent, 2010) conditional on this study's other explanatory factors. This includes SWD who are racial or ethnic minorities, who were suspended no more frequently than SWD who are White. That SWD were not suspended more frequently...
data did not allow us to distinguish between students who were only in one category or in multiple categories. Second, the data for this study provide one glimpse of exclusionary data. Data reporting for years prior to 2013–14 were not required, and data collection for 2015–16 was only the second year of mandatory collection. Therefore, data collection procedures may have been unfamiliar to data collectors. Relatedly, we were unable to confirm the reliability of the data reported by each school. Future years of data may be more accurate. Third, the nature of the model is in generalization of individual school results at the state-level. There may have been a subset of exemplary schools within each state that demonstrated proportional representation for students with and without disabilities who received disciplinary exclusions. Further research is necessary to identify these schools and conduct targeted studies to identify what makes these schools exemplary and how their results can be replicated across schools.

### Effect Sizes and Confidence Intervals Provided

<table>
<thead>
<tr>
<th>Variables</th>
<th>Model 1</th>
<th>95% confidence interval</th>
<th>Model 2</th>
<th>95% confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-5.64</td>
<td>0.28</td>
<td>-5.64</td>
<td>0.28</td>
</tr>
<tr>
<td>Boy</td>
<td>1.44***</td>
<td>0.27 2.50</td>
<td>1.44***</td>
<td>0.30 2.07</td>
</tr>
<tr>
<td>Black</td>
<td>0.75*</td>
<td>0.34 1.10</td>
<td>0.75*</td>
<td>0.52 1.91</td>
</tr>
<tr>
<td>Hispanic</td>
<td>-0.75</td>
<td>0.41 0.22</td>
<td>-0.75</td>
<td>0.41 0.28</td>
</tr>
<tr>
<td>Other race</td>
<td>0.07</td>
<td>0.38 1.07</td>
<td>0.07</td>
<td>0.29 1.10</td>
</tr>
<tr>
<td>SES</td>
<td>-0.21</td>
<td>0.18 0.17</td>
<td>-0.21</td>
<td>0.18 0.68</td>
</tr>
</tbody>
</table>

**Note:** * indicates p < .05, ** indicates p < .01, *** indicates p < .001.
Popular Data Repositories

- National Center for Educational Statistics
  - https://nces.ed.gov/pubsearch/licenses.asp
- Administration for Children and Families
  - https://www.acf.hhs.gov/opre/research
- Health Resources and Services Administration
  - https://www.childhealthdata.org/
- Inter-University Consortium for Political and Social Research
  - https://www.icpsr.umich.edu/icpsrweb/ICPSR/

Common Large-Scale Surveys: NSCH

2016 National Survey of Children’s Health (n = 52,129)

- Purpose and Scope: to survey health and wellness of children ages 0-17
- Strengths: sample size, representativeness, over-sampling
- Limitations: complex sampling, missing data, not longitudinal, information from caregivers
- Considerations: powerful survey, use of effect sizes

**Common Large-Scale Surveys: ECLS**

<table>
<thead>
<tr>
<th>2010-2011 Early Childhood Longitudinal Study (n = 20,250)</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Purpose and Scope: four longitudinal studies exploring child development and education</td>
</tr>
<tr>
<td>• Strengths: educational improvement, large sample, special populations, longitudinal collection</td>
</tr>
<tr>
<td>• Limitations: generalizability, age of cohorts, missing data, time, restricted use</td>
</tr>
<tr>
<td>• Considerations: theory-driven questions, context and educational policy</td>
</tr>
</tbody>
</table>

(NCES, n.d.)

**Common Large-Scale Surveys: NLTS**

<table>
<thead>
<tr>
<th>2000-2009 National Longitudinal Transition Studies (n &gt; 11,000)</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Purpose and Scope: experiences of students with special healthcare needs during transition from high school</td>
</tr>
<tr>
<td>• Strengths: representativeness, longitudinal collection, training modules (podcast), perspective of students, link to SEELS</td>
</tr>
<tr>
<td>• Limitations: restricted use database, some response bias, retrospective data</td>
</tr>
<tr>
<td>• Considerations: no comparison to children without disabilities</td>
</tr>
</tbody>
</table>

(Javitz & Wagner, 2003; NCSER, n.d.; NLTS2, n.d.)
Common Large-Scale Surveys: SECCYD

1991 NICHD Study of Early Child Care and Youth Development (n > 1,300)

- Purpose and Scope: examines relationship between child development and childcare.
- Strengths: access to training workshop materials from SECCYD, flexible examination of child and family factors (including parent beliefs) in relation to childcare, longitudinal, some medical information
- Limitations: age of cohort, not representative
- Considerations: sample comes from communities around universities which limits generalization.

Resources

https://nces.ed.gov/training/datauser

STRUCTURED PANELIST Q&A

Question

Briefly, what are your experiences with conducting and disseminating large-scale secondary analyses?
Question

Given that you engage in secondary analysis, how do you ensure the quality of your work?

Question

How do you approach determining the quality of secondary analyses as well as claims made by study authors?
Question

In your opinion, what are the most pressing concerns in current research regarding the use and interpretation of secondary analysis?

Question

In what ways do you see researchers misusing secondary analysis within educational research?
Question

In what ways can our field help ensure that scholars are disseminating and using secondary analysis responsibly?

Question

What are your recommendations for novice and senior scholars to consider when conducting secondary analyses?
Question

How can practitioners and other make defensible decisions about the quality of scientific findings without access to the studies?

Question

How can researchers facilitate practitioners’ access to the research?
AUDIENCE Q&A AND DISCUSSION

Secondary Analysis Resources and References

- [How Statistics Can Lie](#) - YouTube video for beginners.
- [Effect Size](#) - YouTube Video explaining effect sizes
- [Thresholds for interpreting effect sizes](#)
- [Why the P value is not enough (Sullivan & Feinn, 2012)](#) - This is a great public access article with examples.
- [Choose The Correct Statistics to Answer Research Questions](#) - Very useful resource for researchers and practitioners.
- [Answering Your Research Questions with Descriptive Statistics Diana Suhr, University of Northern Colorado](#) - This appendix has a handy guide to research questions and corresponding statistical analysis.
References


