Multi-level multi-instrument surveys and assessments in the NCES Secondary Longitudinal Studies

Steven Ingels and Dan Pratt

www.rti.org

Outline

- Overview of High School Longitudinal Study of 2009 (HSLS:09)

- Data collection methodologies and technologies
  - Student data collection: survey and math assessment
  - Parent and staff surveys

- Developing multiple- and multi-level instruments
  - Maximizing coverage
  - Minimizing burden
  - Enhancing quality

HSLS:09 Overview

- Fifth NCES-sponsored high school longitudinal study (NLS-72; HS&B; NELS:88; ELS:2002; HSLS:09)
- Conducted for ED/NCES with additional funding from NSF
- Longitudinal study of more than 21,000 fall 2009 9th graders in more than 900 U.S. high schools
- Nationally representative; state-representative for subset of states
- Base year data collection began in fall 2009

HSLS:09 Overview

- First follow-up with same cohort planned for spring 2012
- Field test conducted a year prior to main study (fall 2008; spring 2011)
- Additional rounds of data collection planned
    - Follow the same cohort beyond high school into/through postsecondary education, workforce, and family formation
    - Administrative records (High school transcripts, postsecondary transcripts, other linked data)

Data Collection Methodologies

- **Students**
  - In-school group administration of survey and math assessment
  - Computerized, web-based
  - Entire session limited to 90 minutes
  - Out-of-school CATI follow-up for those not reached in school
- **Parents, math/science teachers, counselors, principals**
  - Self-administered web
  - CATI follow-up

RTI
INTERNATIONAL

In-School Data Collection Challenges

- Single 90 minute session for approximately 30 students
- Limit time for project personnel in school
- Utilize school computers and Internet connectivity but control environment
  - Ensure that no software left on school equipment
  - Ensure that no software could monitor student sessions
  - Allow access to survey site but no other web-sites during session
  - Ensure consistent testing environment for each student
- Determine whether school computers/network capable
- Provide alternative if school equipment use infeasible or not allowed

In-School Data Collection Approach

- Sojourn: Customized distribution of Linux operating system
    - Ensured system compatibility
    - Standardized user experience
    - Secured data being collected
    - Cost-effective solution
- Project laptops as alternative or supplement
- Sojourn used in 84% of participating schools

In-School Data Collection Approach

- 90 minute session

- 15 minutes for instructions, consent, set-up, wrap-up

- 35 minutes for student survey

- 40 minutes for 2-part 40-question adaptive math assessment

  - 15 minutes for 15 question router

  - 25 minutes for 2nd stage test based on router performance

- Automated resumption of survey after math assessment if time remaining and survey not complete

- Rotating order of sections in student survey to minimize item non-response for items at end of survey

RTI
INTERNATIONAL

Efficient Use of Multiple Sources

- 1. Rounded picture via multiple perspectives
- 2. Higher data quality by pursuing best possible source
- 3. Reduction of student (and staff) burden

1. Rounded Picture

- Lack of convergence across sources, even on relatively factual matters, is not necessarily error, but may be part of a larger multidimensional picture

- A holistic picture of the student depends on contextual data from non-student sources

- Example:
  - Do you think your child has a learning disability?
  - Do you (teacher) think this student has a learning disability?
  - Do you (student) have a learning disability?
  - Do school records say this student has an IEP/disability?

Note: overlap = <50% but construct operates for all groups

2. Best source: validity criterion, quality data

- For example, parents are good reporters on family income, students are not

- School administrators are good reporters on school policies; students are not

- Students are the only viable source of information about student beliefs and attitudes

3. Reducing burden

- Over recent years, the Secondary Longitudinal Studies have had to reduce demand on students (though not on teachers) owing to schools' burden concerns

- Some reductions:
  - Assessment time 1972, 1980-82, 1988-1992 = 90 minutes, 2009 = 40 minutes
  - Questionnaire time – older studies = 45-60 minutes, 2009 = 35
  - Total survey session – older studies = half a school day, 2009 = 90 minutes

  General strategy:

  Ask students only what they can uniquely supply

  Otherwise, seek administrative records, contextual sources

HSLS:09 student – test and questionnaire

- Tested achievement in algebraic reasoning
- Basic demographics
- Recent school experiences
- Math and science self-efficacy
- Attitudes about school, math, and science
- High school, college and career plans
- Plans for the year after HS
- Perceptions of college costs

HSLS:09 Parent questionnaire

- Family origin and language use
- Parental education, occupation, income
- Parent on student's earlier educational history
- Parents' school involvement
- Plans and actions in connection with child's future

HSLS:09 School administrator questionnaire

- School characteristics
- Student population
- Teaching staff
- Course offerings
- Principal's background and opinions

Note on minimizing administrator burden

- HSLS:09 counselor questionnaire permits additional school-level information collection
- Administrator questionnaire may be hived off
    - Factual questions may be delegated
    - Principal may respond to background items only
    - Different log-in credentials for designee, admin.
    - Limited access for designee

HSLS:09 Counselor

- Counseling staffing and practices
- Programs, especially regarding transitions (into HS, out of HS and into college)
- Math and science placement/tracking policies

Administrative records (later rounds)

- High school transcripts

- Merged or linked data from external sources (e.g., Census, College Board and ACT, National Student Loan Data System, FAFSA [Central Processing System])

- Postsecondary transcripts

More teacher levels (NELS:88, ELS:2002)

- Teacher reports on school level (ethos or climate, etc.)
- Teacher reports at teacher level (e.g. own background)
- Teacher reports at classroom level for given student(s)
- Teacher rates/evaluates specific students
- Teacher provides an indirect student assessment (e.g., writing skills)
- Teacher provides instructional artifacts (lesson plans, exams/quizzes etc.)
- Teacher reports on textbooks used and extent of use

RTI
INTERNATIONAL

Multiple sources and data quality, A

- Fetters, Stowe and Owings on transcript versus coursework and grades self-reporting
  - Student-reported grades correlate 0.77 with transcript report
  - Coursework correlations vary by subject (math 0.70, science 0.66, social studies 0.39, English 0.28)
  - Students overstated math coursetaking by 1.1 semesters
  - Students overstated specific math courses completed (e.g., 55% completed geometry per student report, 49% per transcript)

Multiple sources and data quality, B

- ## Kaufman and Rasinski on NELS 8th graders vs. parents

  - Family background, parent report as validity criterion shows race/ethnicity correlated 0.77, father's education 0.82, and mother's education 0.76.
  - Validity of reports varied with student characteristics (e.g., females, high SES, good readers = better reporters/closer to parent reports on parental background)

Multiple sources and data quality, C

- McLaughlin and Cohen, NELS students vs. parents
- Query: How often does the choice of source for a common item make a difference for conclusions in analysis?
  - Many tests of statistical significance of predictive correlates of outcomes depended on the source—whether student, or parent, was used—of the predictor information.
  - For NELS almost a fifth of the overlap items had different outcomes
  - Had only one source been asked these questions, analytic conclusions may have been affected

Coda

- Multiple sources for key classification variables may make for multiple sources of incomplete and contradictory information.   Composite variables, which rank sources in terms of likely quality and fill in where the superior source is missing, may minimize the need for imputation.

More Information

Steven Ingels

202.974.7834

sji@rti.org

Dan Pratt

919.541.6615

djp@rti.org