Practical Assessment, Research, and Evaluation

2014

Exploring the Limitations of Measures of Students` Socioeconomic Status (SES)

Emily R. Dickinson

Jill L. Adelson

Follow this and additional works at: https://scholarworks.umass.edu/pare

Recommended Citation

Practical Assessment, Research & Evaluation

A peer-reviewed electronic journal.

Exploring the Limitations of Measures of Students' Socioeconomic Status (SES)

Emily R. Dickinson & Jill L. Adelson

University of Louisville

This study uses a nationally representative student dataset to explore the limitations of commonly used measures of socioeconomic status (SES). Among the identified limitations are patterns of missing data that conflate the traditional conceptualization of SES with differences in family structure that have emerged in recent years and a lack of theoretically-based guidance for how the components of SES should be combined. Using kindergarten achievement data, the study illustrates how both the observed relation between SES and achievement and the observed interaction between SES and kindergarten program would be impacted by the use of different measures of SES. This study also explores the measurement of SES within a structural equation modeling (SEM) framework, highlighting both the relevant conceptual and measurement issues.

Understanding the relationship between family social position and children's educational outcomes is one of the key areas where sociology informs educational research. Students' socioeconomic status (SES) is typically used as the variable that reflects inequality in access to family- and community-level resources that provide essential support for demonstrating academic achievement. Educational accountability systems recognize the importance of student SES by including it among the reporting categories for which states are required to demonstrate improvements in student achievement. In educational research, SES is a frequently used statistical control because empirical data support the notion that SES is a significant contributor, whether directly or indirectly, to both individual and group differences in educational outcomes (Alexander, Entwisle, & Thompson, 1987; Coleman, 1966; Mercy & Steelman, 1982; Roscigno, & Ainsworth-Darnell, 1999).

The purpose of this study is to explore the measurement of socioeconomic status (SES) and its relationship with student achievement using data that are available in a large, nationally representative dataset. Because of the prominent role student SES plays in both educational policy and research, its measurement should be subject to scrutiny, and there should be evidence that commonly used measures of SES reflect a similar construct, identify student SES in similar ways, and correlate with educational outcomes such as student achievement in similar ways. This study uses data from the Early Childhood Longitudinal Survey-Kindergarten (ECLS-K), a data source frequently used in sociological research (e.g., Covay & Carbonaro, 2010; Moller, et. al., 2013). It focuses on U.S. kindergarteners from the 1998-99 school year, addressing the following research questions:

1. Do available variables adequately reflect the SES construct?
2. Do commonly used SES measures identify student SES in similar ways?
3. Does the observed relationship between SES and achievement vary depending on the SES measure used?

This study employs a quantitative, variable-centered approach and uses regression analysis to model average relationships between variables within a population. SES in a regression model in which achievement scores are the outcome would thus give us some indication of

whether or not student SES is, on average, associated with achievement, or is associated with other variables that are also related to achievement. Such models are not suggesting a causal relationship between SES and educational achievement but are recognizing the larger patterns of inequality that exist and ensuring that those patterns are taken into account. SES has been found to be related to achievement from the early childhood years through post-secondary education (Baker, 2009; Chatterji, 2006; Howe, Lawlor, & Propper, 2013; Moore, 2003; Rumberger, 1995).

Theoretical Framework

Why SES Matters

Educational researchers typically cite Coleman's (1966) landmark study that highlighted how a student's background characteristics have implications for the quality of educational opportunities they are afforded. Coleman's work was crucial in establishing what has come to be a well-known social fact (Lee, 1994) by demonstrating that educational achievement is influenced by social rather than only individual factors. Since that time, a large amount of research has recognized the importance of social factors by routinely including SES as a control variable in statistical models. Similarly, in the current age of accountability, educational policy recognizes the importance of social factors by requiring evidence that student subgroups, including SES subgroups, demonstrate levels of performance on par with one another.

Social Stratification

The importance of social background characteristics for individual outcomes can be better understood through the lens of social stratification. The concept of social stratification is built on the assumption that individuals hold positions within a larger social structure and that these positions carry differential access to wealth, power, and prestige. One common view is that individuals are situated within social classes and these classes can be differentiated by the economic position or social prestige of their members or by their ability to exert their will through holding positions of authority or power (Weber, 1924). Early understandings of social stratification have been further expanded to describe the mechanisms through which individuals, depending on class membership, have access to varying levels of social and cultural capital, which in turn shapes their options for achieving desired outcomes (Bourdieu, 1986; Lin, 2000).

Historically, the SES of children has been regarded as a combination of parent income, parent educational attainment, and parent occupational prestige (Duncan, Featherman, & Duncan, 1972). On its face, this measure fits with Weber's view on stratification, with property classes, social status groups, and political parties (Tumin, 1967) roughly corresponding to variability in income, education, and occupation. However, issues with the operationalization of SES highlight some areas of weakness in understanding the link between SES and social class.

SES is often measured as a continuous variable, a single score derived from some combination of income, occupation, and education. This approach to the measurement of SES suggests an underlying continuum on which individuals may be located, thus treating SES as a "gradational concept" (Wright, 2009, p. 330) rather than as distinct classes. Wohlfarth (1997) argued that measuring SES on a gradient implies mobility and reflects that people assume their social positions based on individual merit rather than through class membership. As class-based models of stratification may allow for mobility between social classes (Weber, 1947), an SES continuum does not necessarily invalidate the concept of classes. It does, however, raise the question of where along the SES continuum particular classes should be located and at what SES level of SES one would be considered as moving into a different social class.

Another question is how education, income, and occupation are appropriately combined to create SES. According to Weber, "only persons who are completely unskilled, without property, and dependent on employment without regular occupation, are in a strictly identical class" (Weber, 1947, p. 425), suggesting variability among individuals within classes. It has been suggested that individuals experience advantages in some components of SES and disadvantages in others (Grusky & Weeden, 2009) and combining the elements of SES into a single indicator would fail to capture the interplay between its components.

At this point, it is worth considering the mechanisms through which each individual component of SES might relate to an individual outcome, using student achievement as an example. A family's income, for example, may influence the quality and safety of housing they can secure, which could have implications for children's health and subsequent school performance (Zhang, et. al., 2013). Income could also be

a key factor in the neighborhood in which a family can afford to live, which in turn can have implications for the quality of educational resources that are available and accessible (Klein, 2011). Occupational prestige, although correlated with income, would likely relate to student achievement in different ways. More prestigious jobs may help parents develop connections with others in prestigious positions within a community and draw on those connections for information and support for navigating the educational system (Horvat, Weininger, & Lareau, 2003). Higher levels of education might better equip parents to interact with teachers (Ciabattari, 2010) or lead them to hold higher education-related expectations for their children (Davis-Kean, 2005). Although this is not intended as an exhaustive list of the ways that the different SES components may relate to achievement, it serves an illustrative purpose. The components of SES are conceptually different. Although they may be highly correlated, it stands to reason that each plays a unique role in individual outcomes and strengths in an area could potentially offset deficits in other areas.

We have attempted to ground in theories of social stratification the combination of education, income, and occupation into a single SES indicator and to demonstrate gaps that exist between the conceptualization and operationalization of the SES construct. In the next section, we further discuss issues related to the measurement of SES in an educational context.

SES Measurement in the Context of Academic Achievement

Meta-analyses published two decades apart (Sirin, 2005; White, 1982) have documented the relation between SES and achievement. A common theme among these meta-analyses was that the measurement of SES mattered. Measures of SES that combined two or more indicators had higher correlations than any single indicator, and home atmosphere measures had higher correlations than did any single or combined group of traditional SES indicators (e.g., income and parent education; White, 1982). Effect sizes were larger when SES was measured as a continuous variable, when SES data were obtained from parent and secondary sources rather than from students, and when measured among older students (Sirin, 2005).

The quality of data on student SES that are available to researchers may vary considerably. Individuals may be unwilling to provide information about their household income (Turrell, 2000) or may be inclined to overestimate the true value of characteristics deemed socially desirable (Arnold & Feldman, 1981). A common practice in educational research is to rely on a student's free or reduced lunch status as a proxy for SES (see Ding & Lehrer, 2011; Ronfeldt, Loeb, & Wyckoff, 2013; Schwartz, Rhodes, Chan, & Herrera, 2011 for examples of studies using free/reduced lunch status variables), possibly because these data are maintained by schools and districts and may be more readily available than more detailed information on the components of SES. The practice of using lunch status as a proxy for SES has been called into question (Hauser, 1994). Free/reduced lunch status is determined by family income and thus only reflects one component of SES as it has been traditionally conceptualized. Moreover, it reflects participation in the program rather than eligibility, meaning that some families that would qualify for free/reduced lunch do not receive it and are categories with those who do not qualify. Additionally, it is a single indicator that has been dichotomized and so contains limited information about underlying differences in SES and may mask relationships that are not linear.

It is also common for studies to include only one or two of the SES components as a measure of SES (e.g., Balli, Demo, & Wedman, 1998; O'Connor & Spreen, 1988). Large-scale databases tend to be based on multiple surveys or other data sources, so they often contain both the component variables and an overall SES score, as well as a free/reduced lunch indicator. When multiple measures are available, it is possible to document the similarities and differences of the most widely available and commonly used variables and whether or not variables that are intended to measure a similar construct relate to outcomes in similar ways. This study uses a nationally representative dataset to explore the conceptual and empirical limitations of current approaches to the measurement of SES in the context of student achievement.

Sample

This study explored the measurement of SES using a nationally representative student sample of kindergarten students. Table 1 presents some descriptive statistics summarizing the gender and racial composition of the unweighted student sample.

This study focused on kindergarten students as they were the group in the dataset with the least amount of formal schooling, and so the study would be capturing the relation between family SES and achievement prior

to extensive exposure to school and peer SES factors. The measurement of academic achievement among young children is not without controversy, and assessment at the kindergarten level has been criticized for its focus on developmentally inappropriate content and for its potential negative consequences for children's educational experiences (Shepard, 1994). The direct cognitive assessment used for the ECLS-K study is a computer-based, adaptive assessment based on national and state educational standards and administered to children individually by a trained administrator. This approach was intended to ensure that individual children are assessed with the most appropriate test items (NCES, 2004). This study also included teacher ratings of student achievement to allow for a comparison of results using multiple achievement measures.

Table 1. Descriptive Statistics for the Unweighted Sample

	Percentage of unweighted sample (n=21,409)
Male	51.2
Female	48.8
White	55.2
Black	15.1
Hispanic	17.9
Asian	6.4
Other race/ethnicity	5.4

Measures

One benefit of the ECLS-K is that it provides analysts with several SES measures. These include an SES composite, a categorical measure of SES (quintiles), individual SES components (mother's and father's education level and occupational prestige and family income), as well as parent-reported student free/reduced lunch status. The following section provides descriptions of the variables, along with basic descriptive statistics for each.

Measures of SES

SES composite. The SES composite is computed for each student by averaging the values for mother's education, father's education, mother's occupational prestige, father's occupational prestige, and household income (NCES, 2004). The SES composite values among the unweighted sample ranged from -4.75 to 2.75 with a mean of 0.01 and a standard deviation of 0.80.

SES quintiles. SES quintiles were created by NCES by sorting the SES composite variable and dividing the sample into fifths. As one would expect, the SES quintile variable is highly correlated with the SES composite (Spearman's rho = .98). The resulting quintiles each contained between 18.7% and 21.6% of the sample.

Parents' education. Information about the education of mothers/female guardians and fathers/male guardians were collected via a parent interview conducted during either the fall or spring of the kindergarten school year. Response categories included: 8th grade or below, 9th-12th grade, High school diploma/equivalent, Vocational/Technical program, Some college, Bachelor's degree, Graduate/Professional school- no degree, Master's degree, and Doctorate or professional degree. The median value for mother's education was Vocational/Technical program, and the median value for father's education was some college. For both mother's and father's education, the modal value was High school diploma/equivalent.

Parents' occupational prestige. Information about the occupation of mothers/female guardians and fathers/male guardians were collected via a parent interview conducted during the fall of the Kindergarten school year. These occupations were then recoded based on the 1989 General Social Survey (NCES, 2004). Values for both mother's and father's occupational prestige ranged from 29.60 to 77.50. The average value for mother's occupational prestige was 43.43 (SD = 11.16), and the average value for father's occupational prestige was 43.17 (SD = 10.98).

Household income. Information about the occupation of mothers/female guardians and fathers/male guardians were collected via a parent interview conducted during the spring of the kindergarten school year. Income values in the unweighted sample ranged from $0 to $999,999.99 with a mean of $52,039.89, a standard deviation of $56,398.95, and a median of $40,000.00.

Free/reduced lunch status. Information about enrollment in the federal free or reduced lunch program was collected via a parent interview conducted during the spring of the kindergarten school year. This was a dichotomous variable indicating either that 'Yes,' the student received free or reduced lunch, or 'No,' the student did not receive free or reduced lunch. Of the students for which data on free/reduced lunch were

available, 44.4% did not receive free or reduced lunch, and 55.6% did.

Outcome Variables

IRT scores. The first of two academic achievement outcomes used in this study was Item Response Theory (IRT) scores. Students' IRT scores in mathematics and reading were derived from a direct cognitive assessment administered in the fall and again in the spring of the kindergarten year. The IRT score is calculated to reflect the relative difficulty of items on the assessment and is comparable over time. Mathematics IRT scores ranged between 11.57 and 113.80 points with a mean of 36.27 (SD = 12.00) for the spring administration. Reading IRT scores for the spring administration ranged from 22.23 to 156.85 points with a mean of 46.46 (SD =14.04).

Teacher ratings. The second achievement outcome used was teacher evaluations of students' achievement in the domains of language and literacy and mathematical thinking. These evaluations were on a 5-point Likert scale. The average mathematics teacher rating from the spring data collection was 3.54 (SD = 0.85), and the average literacy teacher rating was 3.37 (SD = 0.80).

Methods and Results

Research Question 1: Reflection of the construct of SES

Though the ECLS-K includes imputed values for the SES component variables (NCES, 2004), the data set still contains values that must be treated as missing (e.g., *Could Not Ascertain*). These additional values were applied, for example, when respondents did not complete an entire interview or survey or refused to answer a particular question. Our first step was to calculate the frequency in which these missing values were present. Table 2 presents the percentage of cases in the dataset that contained values that had to be treated as missing, as they did not contain meaningful data about education, occupational prestige, or income.

Table 2. Percentage of Sample Missing Data on Each SES Component (Unweighted n=21,409)

SES component	% of sample
Mother's education	7.5
Father's education	24.9
Mother's occupational prestige	38.2
Father's occupational prestige	31.0
Household income	5.9

Note: Percentages reflect the portion of the unweighted sample with missing values for each variable listed.

As Table 2 demonstrates, several cases are missing data for at least one SES component. Thus, an individual students' SES may be based on more or fewer components than other students.

Next, we counted the number of SES components for which each student had non-missing values. Table 3 presents the number of components from which each student's SES composite was computed. Table 3 shows that for over half of the students in the dataset, at least one of the SES components was not included in the SES calculation.

Table 3. Percentage of Sample Using Each Possible Number of SES Components (Unweighted n=21,409)

Number of SES components	Percent using number of components
Zero (SES composite missing)	5.9
One	0.0
Two	5.7
Three	16.9
Four	26.9
Five	44.5

Note: Percentages reflect the portion of the unweighted sample for which each number of SES components were included in the SES composite variable. For example, for 44.5% of the sample, all five SES components were included in the NCES creation of the SES composite variable.

One factor that might account for such a pattern of missing values would be the inclusion of single-parent households in which only one parent's education and occupational prestige would be available to contribute to the overall SES. Such an explanation is not satisfying, either empirically or conceptually. First, missing data are not consistently associated with the variable in the dataset that indicates the type of household (e.g., single-parent, two-parent). Secondly, this only further complicates our understanding of SES by introducing other factors related to family composition that have not been incorporated into traditional measures of SES. There has been discussion in the literature about expanding the conceptualization of SES to include family composition variables, as the traditional measurement of SES was based on a two-parent family (Mueller & Parcel, 1981). Though expanding the measurement of SES is beyond the scope of this article, this is an issue that should be addressed in the conceptualization and measurement of SES. Finally,

from a methodological viewpoint, these data are not all missing at completely at random, and this is rarely accounted for in analyses.

Research Question 2: Similarity and Differences in Identification of Student SES

Next, a crosstabulation between SES as measured by both free/reduced lunch status and SES quintiles was computed. Table 4 presents the results from the crosstabulation.

Table 4. Crosstabulation of SES Quintiles and Free/Reduced Lunch Status (Unweighted n=10,386)

SES quintile	Free/Reduced lunch	
	No	Yes
First	4.7	42.8
Second	15.9	26.6
Third	22.1	17.4
Fourth	27.6	9.4
Fifth	29.7	3.8

As Table 4 shows, although the majority of students identified as receiving free or reduced lunch also fall within the lowest two-fifths of the SES distribution, over 10% of free/reduced lunch students are identified as in the upper two-fifths of the SES distribution. Similarly, over 20% of students who are identified as not receiving free or reduced lunch fall in the two lowest SES quintiles. The correlation between the two variables is moderate (Spearman's rho = -.56), indicating that although similar, free/reduced lunch status and the SES composite seem to be measuring different constructs.

One alternative to using available composite SES measures is to model the construct of SES using the individual component variables. The treatment of SES as a construct is documented in the research literature, though there has been debate over whether it should be measured reflectively or formatively (Howell, Breivik, & Wilcox, 20007). Structural Equation Modeling, or SEM, is frequently used to model how observed measures are actually a reflection of an underlying construct that exists apart from the observed measures. Thus, our observations reflect imperfect manifestations of that construct, confounded by measurement error or other factors not related to the construct being measured (Kenny & Kashy, 1992). However, there may be instances in which the construct of interest is more appropriately interpreted as a combination of observed variables rather than a latent construct which is measured reflectively. In fact, SES is

one such variable that has been argued as being the result of a combination of education, income, and occupational prestige (Heise, 1972).

One benefit of modeling SES as a factor is that it takes into account the intercorrelations among the variables comprising SES, rather than simply averaging them. When modeling a formative factor, we are essentially regressing an unobserved variable on a number of observed variables. In order for such a model to be estimable, there must be some observed outcome included in the model (Howell, Breivik, & Wilcox, 2007). In estimating our SES factor, factor loadings are thus interpreted as regression coefficients that provide a sense of the magnitude at which each component contributes to the overall SES score that is predicting the observed outcome. Table 5 presents the factor loadings for SES as a combination of the ECLS-K component variables, predicting mathematics and reading achievement.

Table 5. Factor Loadings for SES Formative Factor Predicting Mathematics and Reading Achievement

	Mathematics		Reading	
	IRT score n=8,483	Teacher rating n=5,651	IRT score n=8,217	Teacher rating n=6,910
Mother's education	.412*	.490*	.424*	.484*
Father's education	.405*	.220*	.371*	.239*
Mother's occupational prestige	.106*	.146*	.144*	.198*
Father's occupational prestige	.113*	.141*	.157*	.078
Household income	.240*	.293*	.199*	.292*

Note. *N-counts are unweighted; * p < .05*

Table 5 illustrates first that the different components of SES do contribute differently to the overall SES factor and that this holds true for different achievement outcomes. For example, across the four content/measure combinations, occupational prestige for either parent tends to contribute less to SES than do education and income. However, Table 5 also illustrates how changing the outcome variable in the model can lead to differences in the relative contribution of particular components. For example, mother's and father's education appear to be more similarly weighted when predicting IRT score than when predicting teacher

rating. Interpretation and limitations of these models will be further discussed in the Discussion section of this article.

Research Question 3: The Observed Relationship between SES and Achievement with SES in the Model

To answer the final research question, we ran a series of regression models predicting end-of-year kindergarten achievement from the various SES measures. All analyses were run using Mplus version 7.11 (Muthén & Muthén, 1998-2012). Mplus software allows for both the application of sampling weights and the appropriate treatment of nested data (i.e., students within schools) by using the Complex analysis with the appropriate ECLS-K weight provided by NCES to appropriately account for disproportionate sampling, nonresponse, and differential coverage for national representativeness (BYCOMW0) and the school ID variable as a cluster ID to account for the design effect. Mplus was also used because it allowed for the modeling of the formative SES factor. Table 6 presents the regression results for IRT scores and teacher ratings.

Table 6. Standardized Regression Weights for Predicting IRT Scores from SES

	Mathe-matics IRT score	Reading IRT score	Mathe-matics teacher rating	Reading teacher rating
Free/ Reduced Lunch	-.358	-.303	-.265	-.254
SES Composite	.405	.354	.299	.312
SES quintiles	.409	.353	.311	.321
SES emergent factor	.373	.322	.257	.272

Notes. Each regression weight represents the relation between the SES variable listed in the row and the achievement variable listed in the column, with each variable added independently as a predictor of the outcome (i.e., one model included Free/reduced lunch while a separate model included the SES composite). For example, students who qualify for free or reduced-price lunch are expected to have on average a mathematics IRT score that is approximately .36 of a standard deviation lower than students who do not qualify for free or reduced-price lunch. When the SES composite is used as the predictor of mathematics IRT score instead, a one-unit increase in SES score is associated with an approximately .41 standard deviation increase in IRT score. All p-values are < .01.

Table 6 indicates that the magnitude of the effect of SES on achievement is similar across the various content/measure combinations. The magnitude of the effects range in absolute value from .25 to .41, though the interpretation of these effects differs as a function of the measurement characteristics of each variable. The negative sign of the free/reduced lunch coefficients reflects the difference in the coding of free/reduced lunch status such that a higher value reflects lower SES. The SES composite and SES emergent factor coefficients can be interpreted as the expected change in achievement score for every unit change in SES. Thus, although the magnitude of the coefficients is similar, these coefficients also capture the incremental increase in achievement as SES increases, rather than simply reflecting the average difference between free/reduced lunch groups. Similarly, the SES quintile coefficient reflects the expected achievement increase as students move up through the SES quintiles.

Another approach to looking at the differences in SES effects depending on the measure used is to document the effects of an educational program for students from different SES backgrounds, using different measures to demarcate student SES. Table 7 presents the standardized regression results from a regression analysis predicting end-of-year kindergarten achievement from whether students attended a full-day or half-day kindergarten program. Several regression equations were run predicting student scores as measured by both IRT-scaled assessments and teacher ratings. In addition to the kindergarten program type, the various SES measures were added into the equations to ascertain the program effect for students from different SES groups.

Table 7 demonstrates that although there are not large differences in the magnitude of the program effect when using different combinations of SES measure (as a control variable) and achievement outcomes, there are some notable patterns. Standardized regression coefficients reflecting the effect of full-day vs. half-day kindergarten tend to be larger when achievement is measured via IRT score and when SES is measured using either free/reduced lunch status or the SES factor. When the SES measure used is the available SES composite (measured both continuously and categorically) and the achievement outcome is the mathematics teacher rating, the effect of participating in full-day kindergarten is not statistically significant at *p* < .01.

Table 7. Standardized Regression Results Indicating the Effect of Program Type in a Model with each SES Measure

	Mathematics IRT score	Reading IRT score	Mathematics teacher rating	Reading teacher rating
Free/ Reduced Lunch	.109	.082	.065	.083
SES Composite	.064	.087	.041	.067
SES quintiles	.066	.089	.043	.069
SES emergent factor	.084	.106	.060	.086

Notes. Each regression weight represents the relationship between kindergarten program type with the achievement measure in each column, after controlling for the SES variable listed in each row. For example, the relationship between participation in full-day kindergarten with mathematics IRT score is approximately .11 of a standard deviation, after controlling for free or reduced-price lunch status. If SES composite is used as the control variable instead, the relationship between full-day kindergarten participation with mathematics IRT score is approximately .06 of a standard deviation. All p-values are < .05.

Findings

Research Question 1: Reflection of the construct of SES

One straightforward way to explore whether or not available variables adequately reflect the SES construct is to document the variables that comprise the SES composite. In the case of the dataset used here, missing values for the component variables resulted in the SES construct being measured differently for different students in the dataset. In some cases, SES contained information about both parents' education and occupational prestige, whereas in others such information was missing. To provide a simple example, two students in the dataset had an SES score of .62, but one reflected only household income and mother's education, whereas the other contained all five SES components.

Although these patterns of missing data may accurately reflect what factors are contributing to a particular students SES, it is difficult if not impossible to tease out whether the absence of particular data means that those variables are not contributing to the context in which a student's educational experience is happening. For example, a non-residential parent or other relative could be contributing income or social/cultural capital that is not being captured in the SES measure. The results of this study suggest that researchers should use caution in their use and interpretation of student SES even when large-scale data sets are available.

Whereas missing data problems highlight an empirical limitation to the measurement of SES, modeling SES within an SEM framework highlights the lack of conceptual clarity in what comprises the SES construct and how those components are to be appropriately combined. Although SES has historically included components of parent education, income, and occupation, there is no clear rationale for how these components should be combined to accurately reflect how they function together to create the context in which individual experiences occur. In the present study, father's occupational prestige was not a significant contributor to the SES construct when modeled as a predictor of teacher rating of reading achievement. Similarly, the standardized loading of father's education level was nearly twice as large when predicting IRT scores rather than teacher ratings. Taken together, these results suggest that combining SES into a single indicator that reflects an equal contribution of education, income, and occupation may not be appropriate. The SES composite does not allow for possible interactions among the components and fails to reflect that each component might contribute differently to the larger construct depending upon the outcome of interest.

Research Question 2: Similarity and Differences in Identification of Student SES

Although free/reduced lunch status may be a variable that is more readily available when using smaller-scale data sources such as school or district databases and that may be more easily interpreted, it is important to document the extent to which it can be used interchangeably with other measures of SES. This was done by comparing the free/reduced lunch status and relative SES standing of students with data available for both variables. One limitation of note was the large amount of missing data for the lunch status variable. The amount of missing lunch status data was slightly higher among the higher SES quintiles, though we cannot discern whether or not these students participated in or

would qualify for free or reduced lunch. Additionally, enrollment in the free/reduced lunch program is not necessarily equivalent to qualifying for it as parents must apply in order to enroll. The present analysis does support prior research that questions the validity of free-reduced lunch status as a proxy for SES.

Research Question 3: The Observed Relationship between SES and Achievement with SES in the Model

Regression results indicate that the magnitude of the effect of SES on achievement is similar across the various measures. Although the regression coefficients are similar in magnitude, they differ in their interpretation. The standardized regression coefficients for free/reduced lunch status reflect the differences in average achievement (in standard deviation units) between students receiving free/reduced lunch and students not, whereas the standardized regression coefficients for the SES composite and for the SES emergent factor reflect the differences in average achievement for each standard deviation increase in an SES score that combines information about some combination of parent education, income, and occupation. Because free/reduced lunch status is a dichotomized variable, the researcher cannot use it to investigate curvilinear relationships, whereas the researcher could with the SES composite or emergent factor. Regression coefficients for the SES quintiles reflect the differences in average achievement of students in adjacent fifths of the SES composite distribution. Thus, a student in the lowest SES quintile

would be expected to have a mathematics IRT score approximately 17.8 points lower than a student in the highest SES quintile, on average. It is important to note that the coefficient reflects the difference for all students in adjacent quintiles, so a student at the bottom of the fifth quintile and a student at the top of the fourth quintile have the same model-predicted difference in achievement as a student at the top of the fifth quintile and a student at the bottom of the fourth quintile.

As standardized regression coefficients can be interpreted as the expected change in the outcome in standard deviation units, predicted values can be calculated to illustrate differences in the relation between SES and achievement depending on the measures used. Table 8 presents predicted values for the four achievement measures using the regression results for the SES composite, SES emergent factor, and free/reduced lunch status variables.

Table 8 shows that if we define a low SES student as having an SES composite score 2 standard deviations below the mean we would expect low SES students on average to have a mathematics IRT score of approximately 27, which is roughly 18 points lower than a high SES student (defined as having an SES composite score 2 standard deviations above the mean). Teacher ratings of mathematics achievement for low SES students would be expected to be approximately one point lower (on a 1-5 scale) than those for high SES students. Modeling SES as an emergent factor yields consistently resulted in lower predicted scores for all three SES groups across the different achievement

Table 8. Predicted Values of Achievement Scores for SES Levels and Free/Reduced Lunch

	Composite			Emergent Factor				
	SES 2 SDs below mean	Mean SES	SES 2 SDs above mean	SES 2 SDs below mean	Mean SES	SES 2 SDs above mean	Free/Reduced lunch	No free/reduced lunch
Mathematics IRT score	27.23	36.05	44.87	14.77	22.93	31.09	35.47	39.30
Reading IRT score	36.90	45.89	54.88	23.31	31.76	40.21	45.55	49.26
Mathematics teacher rating	3.04	3.53	4.02	2.51	2.92	3.33	3.50	3.72
Reading teacher rating	2.89	3.37	3.85	2.31	2.71	3.11	3.34	3.54

Notes: SDs= standard deviations. The predicted mathematics IRT score for a student with an SES composite 2 standard deviations below average is 27.23, whereas the predicted mathematics IRT score for a student with an SES composite 2 standard deviations above average is 44.87.

measures (reflecting differences in the intercept values of the models), though the gaps between the groups are similar to those for the SES composite.

Using free/reduced lunch status as an indicator of low SES, we would expect a low SES student to have mathematics and reading scores IRT score roughly 4 points lower, on average, than a student not receiving free or reduced lunch, and teacher ratings of mathematics and reading achievement would be expected to be approximately 0.2 points lower for low SES students. Table 8 demonstrates that using an SES composite or emergent factor would identify a much larger gap between low SES and high SES students than would using free/reduced lunch as an indicator of low SES. Predicted values were also calculated based on the results of the analysis of full-day vs. half-day program effects when predicting kindergarten achievement and controlling for SES. Although regression results alone suggest that attending full-day kindergarten has small, positive, statistically significant effects on most of the measures of kindergarten achievement after controlling for SES measured in multiple ways, predicted values allow for a clearer demonstration of how different approaches to measuring SES could have implications for how the impacts of educational interventions are interpreted.

Figure 1 depicts the differences in predicted mean reading IRT scores between high SES and low SES groups for each of the three SES measures, as well as the differences between the groups when kindergarten program type is taken into consideration. High SES students are defined as those students with SES scores at two standard deviations above the mean for both the SES composite and SES emergent factor, or as students not receiving free or reduced lunch. Low SES students are defined as those students with SES scores two

standard deviations below the mean for both the SES composite and SES emergent factor, or as students receiving free or reduced lunch.

The first set of bars in Figure 1 shows the average difference between high and low SES groups regardless of program type (i.e., based on the model, the difference between high SES students and low SES students in full day kindergarten was the same as the difference between high SES students and low SES students in half day kindergarten). The second and third sets of bars show the interaction between SES and program type, with the second set displaying differences between high SES students in half day kindergarten and low SES students in full day kindergarten and the third set displaying differences between low SES students in half day kindergarten and high SES students in full day kindergarten. As shown by the height of the bars, across the three SES measures, the largest gap is between high SES students in full day kindergarten and low SES students in half day kindergarten. In other words, though gaps between high and low SES students exist across kindergarten program types, gaps are expected to be smaller when low SES students participate in full day kindergarten. Figure 1 also demonstrates that if student SES is identified using free/reduced lunch status rather than the SES composite or factor, substantially lower gaps in average predicted reading IRT scores would be expected, thereby reducing that measured benefit of participation in full-day kindergarten.

Discussion

Socioeconomic status is commonly included as a statistical control in models predicting educational and other social science variables due to its demonstrated ability for explaining large amounts of variance and because studies controlling for SES are deemed superior
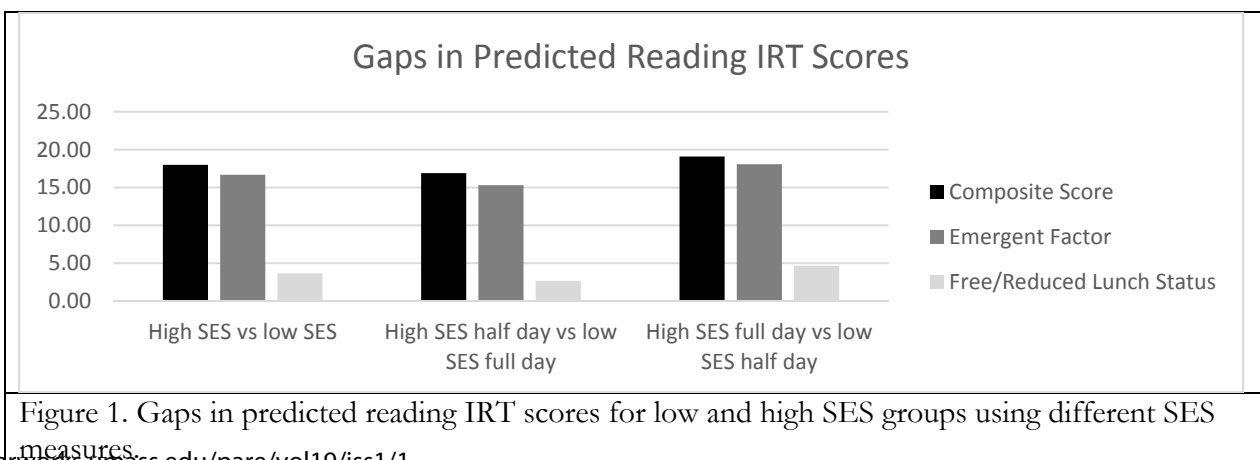

Figure 1. Gaps in predicted reading IRT scores for low and high SES groups using different SES measures.

to those that do not (Jeynes, 2004). One purpose of adding control variables into a statistical model is to attempt to account for the possible effects of other variables in order to get a more accurate estimate of the effects of the independent variable of substantive interest. For example, the positive effects of fathers' involvement in children's developmental outcomes may be observed in study populations, but controlling for SES provides evidence that observed correlations were not just a reflection of positive outcomes associated with the higher SES of fathers who tended to be more involved with their children (Sarkadi, Kristiansson, Oberklaid, & Bremberg, 2008). It is important to note that correlational analyses such as this example and those conducted for the current study do not provide evidence of causality. In other words, though controlling for SES may have helped to isolate the effects of fathers' involvement, it did not indicate that SES caused fathers to be more involved.

However, given that individual SES scores vary in the type and amount of information they contain, for what exactly are we controlling when we include SES? Regression and other correlation-based analyses presume that observed relations between variables can be appropriately applied to all members of what is assumed to be a homogenous population (Poncheri & Ward, 2008). These approaches are considered variable-centered because their focus is on variables that can be abstracted from the individuals or groups that embody them. When a variable is not measured consistently across the units of analysis, then conclusions drawn about its relations to other variables are necessarily flawed.

Though large-scale databases provide a robust source for estimating population-level relationships between variables, they are not without their limitations. A simple analysis of the pattern of missing data in our example database revealed that the measure SES provided in the database may not adequately capture SES as it has been traditionally conceptualized. Simultaneously, the missing data patterns also illustrate that the traditional conceptualization may be in need of an expansion to address the changing nature of family and household configurations. Increased variability in family configurations has been cited as a limitation of the use of the traditional socioeconomic status measure for children (Entwisle & Astone, 1994). Similarly, a recent report published by the National Center for Educational Statistics highlighted the need for an expanded measure of SES to be developed for the National Assessment of

Educational Progress (NAEP; NCES, 2013) and included household composition among the variables that could be incorporated into an expanded SES measure.

Also of concern is that different approaches to the measurement of SES may differently identify students as members of "at-risk" SES groups. Not only do different measures of SES capture different amounts of information about the underlying construct, but they also use different thresholds when defining a student as "at-risk" based on their SES. Regression results from this study demonstrated that achievement score gaps would be much smaller if students were identified as "at-risk" based on free/reduced lunch status and that the measurement of SES could have implications for the perceived benefit of educational programs and services for low SES students. Lunch status is a poor variable for measuring SES not only because it is based only on the income component but also because it is a dichotomous measure. The limitations of dichotomized variables are widely known to include less information about individual differences (MacCallum, Zhang, Preacher, & Rucker, 2002) and problems identifying complex associations between variables such as U-shaped relations (Ravichandran & Fitzmaurice, 2008).

Sociological theory provides tools for understanding how people can be organized into classes based on common levels of access to sources of wealth, power, and prestige. Every individual is located in several separate but overlapping realms of stratification, and the SES construct seeks to capture all of this information. Although a single SES score may be valued for its ability to explain variability, arriving at a consistent measure of student SES is clearly an area of concern. Until limitations in the conceptualization and operationalization of SES can be adequately addressed, quantitative researchers, in particular, need to give more thought to the mechanisms through which the components of SES relate to student achievement and model those instead.
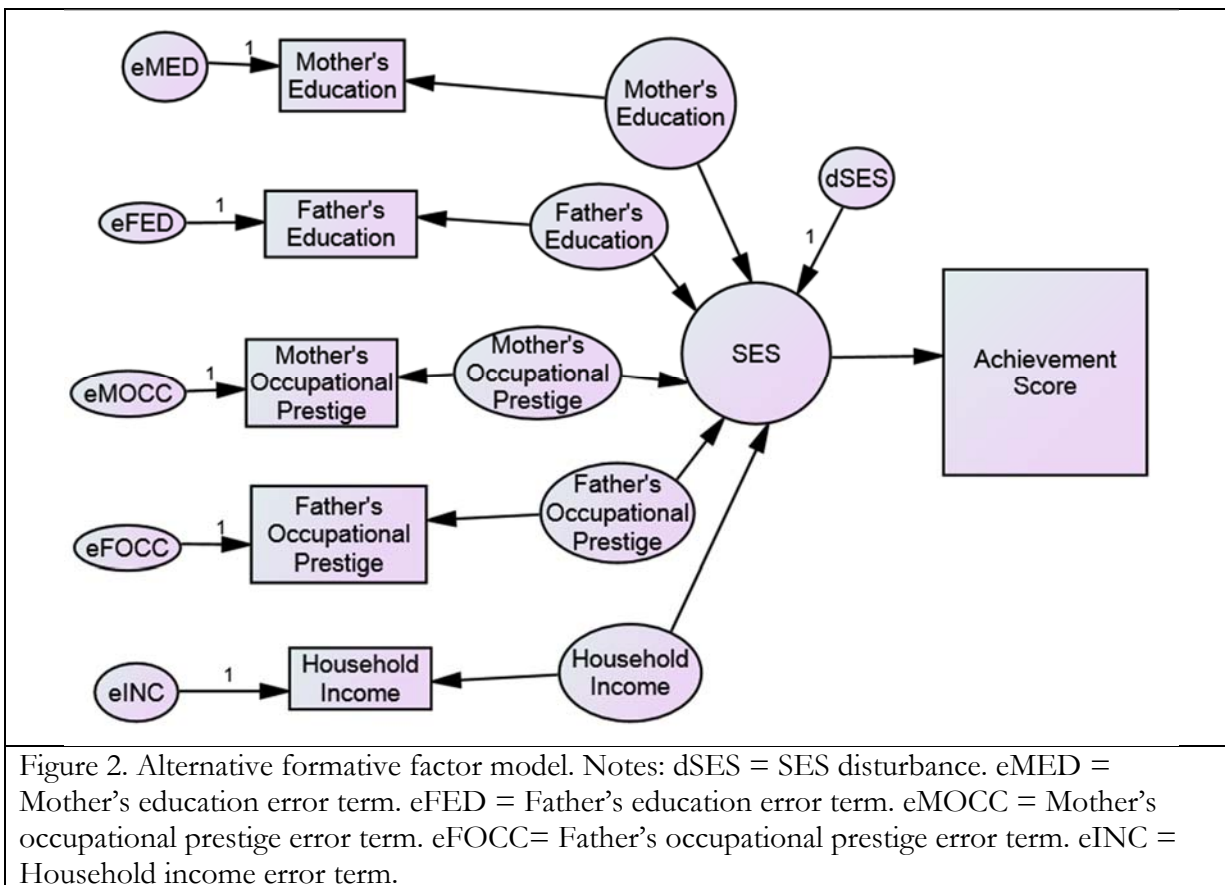
Limitations

Although this study highlights the limitations of existing data sources for adequately modeling the relationship between SES and student achievement, it is itself limited by issues related to missing data. Because so many students were missing information about their free/reduced lunch status, it is not clear the extent to which students would be differently labeled as low SES depending on the measurement used. There is some

Dickinson & Adelson, SES measures

evidence that students who would be labeled low SES using free/reduced lunch status as a proxy would not be so labeled using as SES measure.

Another characteristic of this study that may be considered a limitation is its focus on achievement at the kindergarten level. The validity of interpretations of scores derived from assessments of children at this developmental stage has been called into question (Shepard, 1994). This study attempted to account for

As Figure 2 demonstrates, SES may be modeled as a formative factor that emerges from latent factors representing each of the observed measures. The latent components are reflected in the measured variables, which can now be modeled with measurement error. Unfortunately, there were model identification issues that did not allow the model to be tested in the present study. Future research might explore ways to overcome such issues in modeling SES as a formative factor.

Figure 2. Alternative formative factor model. Notes: dSES = SES disturbance. eMED = Mother's education error term. eFED = Father's education error term. eMOCC = Mother's occupational prestige error term. eFOCC= Father's occupational prestige error term. eINC = Household income error term.

this potential limitation by including an alternative measure of achievement based on teacher ratings. One area for future research is to extend this analysis of SES to verify that similar patterns are observed at other grade levels.

Finally, there were limitations in the modeling of SES as a formative factor. In the model used, SES is treated as an endogenous variable and so fails to capture the measurement error in the observed variables that comprise the SES formative factor. Edwards and Bagozzi (2000) presented an alternative model that would allow for this measurement error to be taken into account. Figure 2 illustrates this model.

References

Alexander, K. L., Entwisle, D. R., & Thompson, M. S. (1987). School performance, status relations, and the structure of sentiment: Bringing the teacher back in. *American Sociological Review, 52*, 664-682.

Arnold, H. J., & Feldman, D. C. (1981). Social desirability response bias in self-report choice situations. *Academy of Management Journal, 24*, 377-385.

Balli, S. J., Demo, D. H., & Wedman, J. F. (1998). Family involvement with children's homework: An intervention in the middle grades. *Family Relations, 47*, 149-157.

Dickinson & Adelson, SES measures

Bourdieu, P. (1986). The forms of capital. In J. Richardson (Ed.) *Handbook of Theory and Research for the Sociology of Education* (pp. 241-258). New York: Greenwood.

Chatterji, M. (2006). Reading achievement gaps, correlates, and moderators of early reading achievement: Evidence from the early childhood longitudinal study (ECLS) kindergarten to first grade sample. *Journal of Educational Psychology, 98*, 489-507.

Ciabattari, T. (2010). Cultural capital, social capital, and educational inequality. *Childhood Education, 87*, 119-121.

Coleman, J. S.   (1966). Equality of Educational Opportunity (COLEMAN) Study (EEOS), Ann Arbor, MI: Inter-university Consortium for Political and Social Research.

Covay, E., & Carbonaro, W. (2010). After the bell: Participation in extracurricular activities, classroom behavior, and academic achievement. *Sociology of Education, 83*, 20-45.

Davis-Kean, P. E. (2005). The influence of parent education and family income on child achievement: The indirect role of parental expectations and the home environment. *Journal of Family Psychology, 19*, 294-304.

Ding, W. & Lehrer, S. F. (2011). Experimental estimates of the impacts of class size on test scores: Robustness and heterogeneity. *Education Economics, 19*, 229-252.

Dumais, S. A. (2006). Early childhood cultural capital, parental habitus, and teachers' perceptions. *Poetics, 34*, 83-107.

Duncan, O. D., Featherman, D. L., & Duncan, B. (1972). *Socioeconomic background and achievement*. New York: Seminar Press.

Edwards, J. R. & Bagozzi, R. P. (2000). On the nature and direction of relationships between constructs and measures. *Psychological Methods, 5*, 155-174.

Entwisle, D. R. & Astone, N. M. (1994). Some practical guidelines for measuring youth's race/ethnicity and socioeconomic status. *Child Development, 65*, 1521-1540.

Grusky, D. B. & Weeden, K. A. (2008). Are there social classes? A framework for testing Sociology's favorite concept. In A. Lareau & D. Conley (Eds.), *Social Class* (pp. 65-89). New York: Russell Sage Foundation.

Hauser, R. M. (1994). Measuring socioeconomic status in studies of child development. *Child Development, 65*, 1541–1545.

Heise, D. R. (1972). Employing nominal variables, induced variables and block variables in path analysis. *Sociological Methods and Research, 1,* 147–173.

Horvat, E. M., Weininger, E. B., & Lareau, .A. (2003). Social ties to social capital: Class differences in the relations between schools and parent networks. *American Educational Research Journal, 40*, 319-351.

Howe, L. D., Lawlor, D. A., & Propper, C. (2013). Trajectories of socioeconomic inequalities in health, behaviours, and academic achievement across childhood and adolescence. *Journal of Epidemiology and Community Health, 67*, 358-364.

Howell, R. D., Breivik, E., & Wilcox, J. B. (2007). Reconsidering formative measurement. *Psychological Methods, 12*, 205-218.

Jeynes, W. H. (2002). The challenge of controlling for SES in social science and education research. *Educational Psychology Review, 14*, 205-221.

Kenny, D. A., & Kashy, D. A. (1992). Analysis of the multitrait-multimethod matrix by confirmatory factor analysis. *Psychological Bulletin, 112*, 165-172.

Klein, S. (2011). The availability of neighborhood care and education resources and the maltreatment of young children. *Child Maltreatment, 16*, 300-311.

Lee, D. J. (1994). Class as a social fact. *Sociology, 28*, 397-415.

Lin, N. (2000). Inequality in social capital. *Contemporary Sociology, 29*, 785-795.

MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods, 7*, 19-40.

Mercy, J. A., & Steelman, L. C. (1982). Familial influence on the intellectual attainment of children. *American Sociological Review, 47*, 532-542.

Moore, B. L. D. (2003). *An Exploration of the influence of demographic factors on individual and aggregate student achievement measurements in the Kentucky accountability system.* (Unpublished doctoral dissertation). University of Louisville: Louisville, KY.

Moller, S., Mickelson, R. A., Stearns, E., Banerjee, N. & Bottia, M. C. (2013). Collective pedagogical teacher culture and mathematics achievement: Differences by race, ethnicity, and socioeconomic status. *Sociology of Education, 86*, 174-194.

Mueller, C. W. & Parcel, T. L. (1981). Measures of socioeconomic status: Alternatives and recommendations. *Child Development, 52*, 13-30.

Muthén, L. K., &Muthén, B. O. (1998-2012). Mplus Users Guide (7th ed.). Los Angeles, CA: Muthén & Muthén.

O'Connor, S. C. & Spreen, O. (1988). The relationship between parents' socioeconomic status and education level, and adult occupational and educational achievement of children with learning disabilities. *Journal of Learning Disabilities, 21*, 148-153.

Poncheri, R. M., & Ward, S. J. (2008, April). Personality remix: Integrating variable-centered and person-centered approaches to personality. Paper presented at the 23rd annual meeting of the Society for Industrial and Organizational Psychology, San Francisco, CA

Ravichandran, C. & Fitzmaurice, G. M. (2008). To dichotomize or not to dichotomize? *Nutrition, 24*, 610-611.

Ronfeldt, M., Loeb, S., & Wyckoff, J. (2013). How teacher turnover harms student achievement. *American Educational Research Journal, 50*, 4-36.

Roscigno, V. J., & Ainsworth-Darnell, J. W. (1999). Race, cultural capital, and educational resources: Persistent inequalities and achievement returns. *Sociology of Education, 72*, 158-178.

Rumberger, R. W. (1995). Dropping out of middle school: A multilevel analysis of students and schools. *American Educational Research Journal, 32*, 583-625.

Sarkadi, A., Kristiansson, R., Oberklaid, F., & Bremberg, S. (2008). Fathers' involvement and children's developmental outcomes: a systematic review of longitudinal studies. *Acta Paediatrica, 97*, 153-158.

Schwartz, S., Rhodes, J., Chan, C. & Herrera, C. (2011). The impact of school-based mentoring on youth with different relational profiles. *Developmental Psychology*, 47, 450-462.

Shepard, L. A. (1994). The challenges of assessing young children appropriately. Retrieved from: http://www.cse.ucla.edu/products/parents/cresst_challengesshepard.pdf.

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research, 75*, 417-453.

Tumin, M. M. (1967). *Social Stratification*. New Jersey: Englewood Cliffs.

Turrell, G. (2000). Income non-reporting: Implications for health inequalities research. *Journal of Epidemiology and Community Health, 54*, 207-214.

U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics. (2004). ECLS-K base year public-use data files and electronic codebook. Retrieved from: http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2001029rev.

U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics. (2013). Improving the measurement of socioeconomic status for the National Assessment of Educational Progress: A theoretical foundation. Retrieved from: http://nces.ed.gov/nationsreportcard/researchcenter/socioeconomic_factors.aspx.

Weber, M. [1924] 1978. *Economy and Society*, edited by Guenther Roth and Claus Wittich. Berkeley: University of California Press.

Weber, M. (1947). *Theory of Social and Economic Organization*. Glencoe, IL: The Free Press.

White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin, 91*, 461-481.

Wohlfarth, T. (1997). Socioeconomic inequality and psychopathology: Are socioeconomic status and social class interchangeable? *Social Science & Medicine, 45*, 399-410.

Wright, E. O. (2009). Logics of class analysis. In A. Lareau & D. Conley (Eds.), *Social Class* (pp. 329-349). New York: Russell Sage Foundation.

Zhang, N., Baker, H. W., Tufts, M., Raymond, R. E., Salihu, H., & Elliott, M. R. (2013). Early Childhood Lead Exposure and Academic Achievement: Evidence from Detroit Public Schools, 2008-2010. *American Journal of Public Health, 103*, 72-77.

Citation:

Authors:

Emily R. Dickinson
Educational & Counseling Psychology, Counseling, and College Student Personnel Department
College of Education and Human Development
University of Louisville
Louisville, KY 40292
emily.dickinson [at] louisville.edu

Jill L. Adelson
Educational & Counseling Psychology, Counseling, and College Student Personnel Department
College of Education and Human Development
University of Louisville
Louisville, KY 40292
jill.adelson [at] louisville.edu