

# Identifying SES Factors in the National Household Education Surveys

Amanda Delore<sup>a</sup>, Ashley Jeanty<sup>b</sup>, & Ron Brown<sup>\*b</sup>

<sup>a</sup>Graduate School of Arts and Sciences, Lesley University, Cambridge, MA

<sup>b</sup>School of Liberal Arts, Dean College, Franklin, MA

<https://doi.org/10.33697/ajur.2024.125>

Students: adelore333@gmail.com; ashleyjeanty01@gmail.com

Mentor: rbrown@dean.edu\*

## ABSTRACT

The purpose of this study was to determine if conceptually and psychometrically rigorous measures of socioeconomic status could be developed for the National Household Educational Surveys (NHES). Exploratory factor analyses were performed for the following NHES: (1) Parent and Family Involvement in Education; (2) School Readiness; (3) Early Childhood Program Participation; (4) After School Programs and Activities; and (5) Adult Education/Adult Education and Lifelong Learning over several years of this administration, including two of the NHES surveys conducted in 2019. The results, while preliminary, suggest that rigorous measures of socioeconomic status (SES) can be constructed for most of the surveys based on Hauser and Warren's model which used income, education, and occupation. This should enhance inferences based on these data and simultaneously increase the appeal of this database to the educational research community. If the measurement of SES is enhanced, it will fundamentally raise the level of quality of research findings and make them more attractive to educational researchers who use the NHES survey database.

## KEYWORDS

Educational Attainment; Exploratory Factor Analysis; National Surveys; Parent Involvement; School Readiness; Socioeconomic Status; Data Analysis Methods; Educational Research Databases

## INTRODUCTION

The NHES surveys constitute a comprehensive database that provides data on the educational activities of the US population from early childhood through adulthood and, as such, provide fertile ground for the exploration of SES, which has had a model proposed by Hauser and Warren.<sup>1</sup> The NHES comprises five different biannual surveys carried out with households at different periods.<sup>2</sup> A key NHES survey dataset feature is the need to consider SES to avoid compromising inferences that may occur due to the widely documented impact of SES on a range of educational outcomes.<sup>3,4</sup> This requires effectively engaging conceptually and psychometrically rigorous measures of SES, which was the ultimate purpose of this study. The present study aims to improve SES measures and, thereby, enhance the strength and validity of inferences based on the NHES surveys while making the NHES survey database more useful and appealing to a much larger audience within the general educational research community.

The availability of such measures should enhance inferences based on these data and increase the appeal of this database to researchers more broadly. Hauser and Warren provided a model arguing that SES is a multidimensional construct composed of householder occupation, education, and income.<sup>Error! Bookmark not defined.</sup> This conceptualization produces a hierarchy implying that an NHES survey respondent whose SES is "high" has a higher level of education, occupational status, and/or income than a respondent whose SES is "low." We adopt this conceptualization here.

The NHES surveys are one of several databases managed by the National Center for Education Statistics (NCES); others include the National Assessment of Educational Progress (NAEP), High School and Beyond (HSB), the Early Longitudinal Childhood Study (ELCS-K), Current Population Survey (CPS), and the Educational Longitudinal Study of 2002 (ELS). Among the surveys contained in the NCES, raw data are typically available for public download, as they were for the present study. NHES surveys were initially administered in the early 1990s and have since been conducted biannually. Five surveys comprise the NHES: (1) Parent and Family Involvement in Education (PFI); (2) School Readiness (SR); (3) Early Childhood Program Participation (ECP); (4) After School Programs and Activities (ASPA) for Middle School and High School Students; and (5) Adult Education/Adult Education and Lifelong Learning (AE/AELL). Data are collected via phone interviews using a list-assisted sampling method. Not every survey is administered in each NHES survey data collection and only data for 2005 and 2007 are available online. Most data collected are dichotomous yes/no responses; however, responses also include continuous, fill-in-the-blank, and Likert items. A key feature of NHES surveys is that data are collected at the household level.

Although precise usage statistics for NCES databases are not readily available, evidence suggests that NHES surveys are among of the least frequently used by educational researchers. For example, no study published between 2000-2010 in three prominent educational research journals (*American Educational Research Journal*; *Educational Evaluation & Policy Analysis*; *Sociology of Education*) reported using the NHES survey database, whereas several studies used other NCES databases (e.g., ELCS-K, ELS). An aim of this study is to make the NHES survey datasets more appealing to educational researchers through the present analyses and practical explorations of the benefits of using these datasets.

While NHES surveys do not appear to be widely used in the field of educational research, some educational researchers have reported results based on these surveys. For example, Silva *et al.*<sup>4</sup> studied barriers to adult participation in education, and Zill *et al.*<sup>5</sup> examined the adjustment of children who enter kindergarten late or repeat kindergarten. One characteristic shared by many analyses of widely used NCES datasets, including NHES, is the need to consider SES in the analyses because a failure to do so can seriously compromise inferences. This occurs because of the widely documented impact of SES on a range of educational outcomes.<sup>3,4,6</sup>

To successfully control for the effects of SES, there must be one or more conceptually and psychometrically rigorous measures of the concept. These measures can then serve as a control variable in data analyses; for instance, they may act as covariates, matching variables and/or they may figure in propensity models. However, constructing a rigorous measure of SES is challenging due to a potential mismatch between the conceptualization of SES and available variables in a database, significant measurement error associated with variables sometimes used as measures of SES (e.g., eligibility for a free/reduced price lunch), and the presence of missing data.<sup>7</sup> These problems increase the likelihood that inferences pertaining to outcomes of interest (e.g., householder satisfaction with their child's school) will be biased by a failure to adequately take SES into account.

#### *Purpose of the Study*

The NHES survey database offers much to educational researchers because of its detailed focus on pre-primary education, childcare utilization, school safety and discipline, and adult education, especially if one or more rigorous measures of SES are available. However, no previous research has constructed measures of SES for the NHES survey database. The purpose of this study was to determine if conceptually and psychometrically rigorous measures of SES for the NHES survey database can be constructed. Ultimately, it could be a great contribution to future studies and to the utility that the NHES survey database can make in educational research if rigorous measures of SES were available at this level. Though this study may be regarded as an "exploratory study", the availability of such measures should enhance inferences from these data and simultaneously increase the appeal of this database to the educational research community.

#### LITERATURE REVIEW

##### *Socioeconomic Status*

Defining and operationalizing SES is an ongoing process in the scholarly activity of several disciplines including education, psychology and public health, among others. Research shows that living in a high-SES area influences people's daily health choices.<sup>8</sup> In fact, greater SES has been demonstrated to be protective against increasing HbA1c (Glycated hemoglobin), with physical activity partially mediating this relationship.<sup>9</sup> SES has also been linked to mental health, with Reiss *et al.*<sup>10</sup> finding that children are at a higher risk for mental health problems when their parents have a lower level of education and subsequently more stressful lives.

However, exploring SES has been complicated by differences in how it is defined and measured across fields. For example, in sociology the terms "social class," "social stratification," or "social mobility" are generally used interchangeably and are associated with a particular theory<sup>11</sup> that treats this construct as a latent variable. On the other hand, in education and psychology, the term most used is SES, which is defined by the variables used to measure it in a study rather than a theory of SES, and thus is atheoretical.<sup>12</sup> Educational literature utilizes SES through different definitions such as Hauser and Warren's.<sup>13</sup> Some literature focusing on education while associating it with SES defined the term in relation to the number of books a parent or guardian has in their home or in relation to their geographical location.<sup>13</sup> Tan *et al.*<sup>14</sup> provided information on the examination of parent involvement in their children's education regardless of the parent's SES and defines it as the educational level the parent achieved. Bollen *et al.*<sup>15</sup> convincingly argues the feasibility and need for grounding a definition of SES [and class] within research. In order to provide the grounding Bollen is referring to, research, including measures of SES, require consistency in how SES is conceptualized. The NHES surveys provide multiple data that research could consider of interest when examining SES from different fields of study. A few of the datasets within the NHES surveys include those which investigate parental involvement in a child's life and educational success in areas such as cognition, language, and social development (Oswald, 2018).<sup>16</sup> For the purpose of this study, an education-based definition of was utilized.

An examination of SES measures used in educational research reveals few that demonstrate a clear conceptual or psychometric foundation. Moreover, these measures are generally dated (e.g., Hollinghead's index, Siegel Prestige Scale;<sup>17</sup> Duncan Socioeconomic Index) and do not conceptualize SES as a latent variable. Educational research typically views SES as unidimensional, in contrast to Hauser and Warren who view SES as multidimensional<sup>Error! Bookmark not defined.</sup>, Hauser and Warren suggest SES is a latent variable composed of the three facets of householder occupation, education, and income.<sup>1</sup> The resulting multidimensional conceptualization of SES represents shorthand expression for variables that enable the placement of persons, families, households and aggregates such as statistical local areas, communities and cities in some hierarchical order, reflecting their ability to produce and consume the scarce and valued resources of society.<sup>Error! Bookmark not defined.</sup> Hauser and Warren's<sup>Error! Bookmark not defined.</sup> conceptualization implies that a householder (NHES survey respondent) whose SES is "high" has more education, higher occupational status, and/or income than a respondent whose SES is "low", thus suggesting a "hierarchical order" with high SES ranked at the top.

SES has been discussed extensively in terms of its importance to educational outcomes. For example, Tan *et al.*<sup>14</sup> examined the patterns associated with parental involvement variables and the achievement level of students from K-12. Another study by Yerdelen-Damar and Peşman<sup>18</sup> explores SES in relation to students' physics achievement where high school students were observed in the hope of finding a correlation between the different genders and their SES to their achievements in learning and understanding physics. Students benefit from both metacognitive strategies and self-efficacy, but self-efficacy plays a key role in translating metacognitive skills into actual achievement. The study revealed that a parent's SES is connected to a range of different learning strategies used by students, and the results convey a link between the learning methods, meta-cognition and physics self-efficacy, and differences in SES and the different physics achievements. SES had a small but positive relation to students who use different strategies to achieve their goals, in this case learning and succeeding at physics.

### *Exploratory Factor Analysis*

#### *Origins and Purpose*

The central goal of factor analysis is to represent a set of observed variables with a set of factors (potentially latent variables) through inter-correlations to generate factors.<sup>19</sup> In the social sciences, factor analysis is generally used for one of three purposes. One is as a purely empirical data reduction technique, usually involving factor analyzing data for the purpose of identifying  $P$  factors underlying  $T$  observed variables. An example of this use (typically labeled exploratory factor analysis or EFA) can be seen in the work of Bilder *et al.*,<sup>20</sup> who analyzed items reflecting positive and negative symptoms of schizophrenia to produce three factors.

A second, and related, purpose is captured by Gorsuch,<sup>19</sup> who defined factor analysis as a form of summary of connected relationships between variables in a manner that is both accurate and concise and that is helpful in developing a conceptualization. Here, interest centers on factors explaining information in the observed variables in ways that are linked to a study's purpose (e.g., identifying subsets of items on a questionnaire reflecting conceptually meaningful subscales). Typically, this is accomplished through EFA. Lyden *et al.*<sup>21</sup> provide an example of this with their focus on the structure of a modified version of The National Institutes of Health Stroke Scale to compare this scale to an original version.

Some researchers have raised concerns over these two purposes and the use of EFA, most of which center on the potential subjectivity in the process of identifying factors. However, Henson and Roberts<sup>22</sup> argued that this process inherently requires a researcher to make informed decisions, and, when done carefully, represents an important and appropriate use of EFA. The research question and methodology which includes defining and operationalizing SES is presented below. The development of an SES measure involves aggregating variables (items) in the NHES survey dataset, which implies the use of factor analytic methods. Specifically, exploratory factor analytic methods were carried out on the surveys analyzed for this research using principal factor analysis (PFA) to examine factors that emerged and whether they can be used as proxy measures for SES, as described below. In the present study, given the treatment of SES as a latent variable (construct/factor), principal factor analysis (PFA) was used, as it extracts common variance, an important element when examining underlying extracted factors.

#### *Extraction Method*

Multiple options exist for extracting factors including unweighted least squares, generalized least squares, maximum likelihood, PAF, alpha factoring, and image factoring.<sup>23</sup> According to Fabrigar *et al.*,<sup>24</sup> maximum likelihood is optimal because "it allows for the computation of a wide range of indexes of the goodness of fit of the model [and] permits statistical significance testing of factor loadings and correlations among factors and the computation of confidence intervals." Maximum likelihood provides loadings that are most likely and permits hypothesis testing for the loadings. Additionally, the use of maximum likelihood was advantageous given that there are many fit indices available when making use of this extraction method, making it more broadly used. However, Fabrigar *et al.*<sup>24</sup> caution that if the assumption of normality is violated, then other methods are recommended. This study used maximum likelihood as the extraction method for each NHES survey (see Mueller<sup>6</sup> for details of this method).

### *Number of Factors to Extract*

A critical step in EFA is determining the number of factors to extract. This is typically done through methods such as identifying eigenvalues greater than one, the scree test, and Velicer and Jackson's MAP criteria and parallel analysis.<sup>25</sup> Costello and Osborne<sup>23</sup> suggest that the scree test is generally the most effective method and involves examining a scree plot and identifying the break where the curve flattens out; data points above the break are considered an adequate number of factors to retain.<sup>19</sup> Examining the scree plot produced for each of the surveys in preliminary analyses helped to determine how many factors to extract and led us to set the number of extracted factors per survey at three (see below). Given this work is exploratory in nature and that a threshold of .30 for factor loadings is commonly used to determine the practical significance of an item in defining a factor,<sup>26,27,28</sup> it was specified in this study that a variable loading on a factor (i.e., correlation between an item and a factor) had to exceed .30 to treat that variable as helping to define the factor.<sup>29</sup>

### *Rotation Method*

Extracted factors are sometimes rotated to improve factor structure with respect to interpretation and simplifying and clarifying the data structure. Rotation can be oblique, in which the resulting factors are correlated, and orthogonal, in which the resulting factor loadings are interpreted as simple correlations of items with factors.<sup>30</sup> Varimax rotation is considered the most popular method for orthogonal rotation and was used in the factor analyses carried out in this study.

### *Sample Size*

Clear guidance on ideal sample size for factor analysis is arguably not available. Costello and Osborne<sup>23</sup> found that a respondent-to-variable ratio of 10:1 is commonly utilized largely without formal justification. Fortunately, sample size is not an issue in the current study as each of the five surveys used for this study involved at least 2,633 respondents and a ratio of respondents to items exceeding 50:1. Secondary data analyses using factor analyses were carried out on five surveys of the NHES. IRB approval was granted for this study at each participating institution and the data are available to the public and not individually identifiable. Conceptually and empirically distinct SES measures are needed for each survey. The surveys were analyzed to determine if SES measures could be constructed for each NHES survey using Hauser and Warren's<sup>Error! Bookmark not defined.</sup> definition. Using the NHES survey database, we sought to determine what factors emerge that can act as a proxy measure for SES.

### *Research Question*

The present study was guided by the following research question:

RQ: Can conceptually and psychometrically rigorous measures of SES be constructed for the NHES survey database using variables and data available within this database? If so, what is the nature of the evidence supporting their construction?

## **METHOD AND PROCEDURES**

### *Design of the Study*

As noted, the NHES surveys consist of five surveys (PFI, SR, ECPP, ASPA, AE/AELL) with a focus on different groups (e.g., elementary age children, high school students). These surveys are conceptually and empirically distinct, and separate SES measures are needed for each. A four-step procedure was used to determine if measures of SES could be constructed for each NHES survey. In all cases, Hauser and Warren's<sup>Error! Bookmark not defined.</sup> definition guided the construction of the measures using the variables available in each survey. First, the purpose of each survey was determined using NCES materials to ascertain what types of household information was gathered. Second, pertinent variables in each survey were identified. Third, EFA was conducted to identify potential SES factors. Fourth, construct validity evidence was generated by examining Hauser and Warren's<sup>Error! Bookmark not defined.</sup> definition and the resulting factors, and correlations between the factors and variables were identified and explored. This procedure was followed for each of the five surveys in the NHES survey database.

### *Respondents*

The PFI survey of 2007 was completed with the parents of 10,681 children in kindergarten through 12th grade, including 10,370 students enrolled in public or private schools and 311 home-schooled children. The PFI for 2019 was carried out with 14,075 participants. The SR survey data for 2007 were provided by 2,633 children aged 3-6 and kindergarten through second grade. The ECPP survey of 2007 used interviews with 7,209 parents of children aged three through third grade, children from birth through age six, and those not yet in kindergarten. The ECPP survey for 2019 utilized survey data from 7,092 participants. The ASPA of 2005 sample of 11,684 is nationally representative of non-institutionalized students enrolled in kindergarten through eighth grade with a maximum age of 15 years. The AE/AELL survey produced a national sample that provides information on civilian and non-institutionalized persons aged 16 or older and not enrolled in grade 12 or below. Sample sizes for the AE/AELL surveys ranged from 6,697-19,722 non-institutionalized adults aged 16 and older, not enrolled in 12th grade or below, and not on active duty in the US armed forces. The 2005 AE survey was based on 8,904 adults.

*Procedures*

The four-step procedure for carrying out analyses is detailed below for the 2005 AE survey, with the same steps repeated for the remaining surveys in the current study. The 2005 AE survey collected information on householder participation in college and university programs taken for work-related reasons, postsecondary vocational/technical diploma reasons, apprenticeships, work-related courses, and work-related informal learning (e.g., courses taken to obtain a GED or obtain ESL skills). Only data for the most recently administered surveys that are available online were analyzed.

In step one, the purpose of the AE as outlined in NCES documentation was examined. In general, the AE attempts to characterize adult involvement and attainment in both formal and informal learning activities. Formal activities include those courses where an instructor is present while informal activities are those with no instructor present. Questions appearing in the NCES materials include “What is the highest grade or year of school that you completed?,” “Do you have/Did you later receive a high school diploma or its equivalent, such as a GED?,” and “Did you work at a job for pay or income at any time in the past 12 months, including self-employment?,” among others. Certain response patterns suggest different levels of adult involvement in learning activities, different reasons for taking courses, and an employer’s financial contribution to the adult’s learning activity. Next, several variables in the AE data consistent with Hauser and Warren’s definition based on householder occupation, education, and income were identified. Examples of candidate AE variables for an SES measure for this survey included: 1) Educational Level (Highest grade/YR completed and HS diploma or GED); 2) Occupational Status (Year last worked for pay or income / Hours worked per week); and 3) Income (Total household income range / Currently works for employer / Amount of earnings). The availability of multiple variables reflecting, for example, educational level, strengthens the analyses because these serve as “items” and having multiple items assessing the same facet should ultimately increase the reliability and ideally the validity of the SES measure.

EFA of AE variables deemed to be consistent with Hauser and Warren’s facets of SES (including those above) to preliminarily identify SES factors using SPSS version 20<sup>31</sup> were performed. These analyses provided evidence of an identifiable SES factor structure corresponding to household educational level and income.

The last step involved developing conceptual and empirical arguments about the extent to which a factor supports valid inferences about household SES for the AE. This focused on assessing construct validity using a combination of two complementary lines of evidence. The first is the agreement between the factor analytic findings and the conceptualization of SES attributed to Hauser and Warren. Specifically, the emergence of identifiable SES factors corresponding to householder occupational status, education, and/or income provides evidence that NHES survey respondents can be placed in a hierarchical order and thus indicates construct validity. The second line of evidence comes from correlations between the derived factors and other variables in the AE survey.

**RESULTS**

*Adult Education/Adult Education and Lifelong Learning (AE/AELL, 2005)*

The AE/AELL survey (see Table 1) was based on 8,904 adults, with respondents ranging from 6,697-19,722 non-institutionalized adults aged 16 and older, not enrolled in 12th grade or below, and not on active duty in the US armed forces.

AE/AELL items	Factor loading		
	1	2	3
Factor 1: Education Related to Employment			
Earned college credit	.98	-.05	.07
Meet requirements for public assistance	.98	-.05	.07
Attend college or vocational school	.98	-.05	.07
Get a raise or promotion	.98	-.05	.07
Get a new job	.97	-.05	.07
ESL for work or personal reasons	.96	-.04	.07
Currently taking ESL classes	.95	-.04	.07

AE/AELL items	Factor loading		
	1	2	3
Instruction provider was employer	<b>.70</b>	-.02	.04
Work while taking ESL classes	<b>.69</b>	-.02	.04
Employer required to take ESL classes	<b>.68</b>	-.01	.04
Being paid while taking ESL classes	<b>.68</b>	-.02	.04
Employer suggested to take ESL classes	<b>.65</b>	-.01	.03
Total hours attending ESL classes	<b>.58</b>	-.03	.02
Factor 2: Higher Educational Attainment			
Type of program-professional degree	.03	<b>.97</b>	-.06
Get/keep certificate or license	.03	<b>.97</b>	-.06
Type of degree-another degree	.03	<b>.97</b>	-.05
Type of program-doctorate	.03	<b>.96</b>	-.06
Type of program-master's degree	.03	<b>.93</b>	-.06
How long enrolled in program	.03	<b>.92</b>	-.06
Type of program-bachelor's degree	.03	<b>.92</b>	-.05
Type or program-associate's degree	.03	<b>.92</b>	-.05
Total credit hours enrolled	.02	<b>.75</b>	-.05
Factor 3: Lower Educational Attainment			
Basic skills classes	.06	-.08	<b>.99</b>
Other high school equivalency program	.07	-.08	<b>.99</b>
GED preparation classes	.07	-.08	<b>.99</b>
Ever taken ABE/GED classes	.03	-.08	<b>.95</b>
Has high school diploma or GED	.02	.01	<b>.66</b>
Actual grade 9-11 completed	.02	-.05	<b>.54</b>
Actual grade 0-8 completed	.04	-.05	<b>.52</b>
Took ESL classes	.12	-.01	<b>.33</b>
Ever taken ESL classes	-.06	-.00	<b>.31</b>

Note: N = 8,904. Items with factor loadings < .3 are omitted. Bold indicates a factor loading above .3 and to indicate which factor (1, 2 or 3). ESL = English as second language; ABE = Advanced basic English classes.

Table 1. Adult Education/ Adult Education and Lifelong Learning (AE/AELL; 2005) survey

The three factors to emerge were “Education Related to Employment” (factor 1), “Higher Educational Attainment” (factor 2), and “Lower Educational Attainment” (factor 3), with results representing a pattern that is partly consistent with the conceptualization of SES as having an educational facet.

*Before and After School Programs and Activities (ASPA; 2005)*

The ASPA (see Table 2) sample of 11,684 is nationally representative of non-institutionalized students enrolled in kindergarten through eighth grade with a maximum age of 15 years.

ASPA items	Factor loading		
	1	2	3
Factor 1: Father’s Educational Attainment			
Dad enrolled in school	<b>.93</b>	-.29	-.02
Education attainment of child’s father/guardian	<b>.93</b>	.13	-.03
Months dad worked in past year	<b>.88</b>	-.15	-.02
Highest grade/yr of school dad completed	<b>.87</b>	.23	-.03
Dad worked for pay last week	<b>.84</b>	-.34	-.01
Hours per week dad works for pay	<b>.79</b>	-.08	-.02
Work status of dad/stepdad/foster dad/guardian	<b>.72</b>	-.36	-.01
Dad’s ease of leaving work	<b>.69</b>	-.07	-.00
Household income below or above \$50K	<b>.55</b>	.38	-.02
Total household income range	<b>.50</b>	.32	-.01
Family received Medicaid in past 12 months	<b>.40</b>	.28	-.03
Family received food stamps in past 12 months	<b>.38</b>	.20	-.03
Factor 2: Mother’s Educational Attainment			
Highest grade/yr mom completed	.30	<b>.79</b>	-.03
Highest level of parent/guardian education	.50	<b>.75</b>	-.04
Educational attainment of child’s mother/guard	.28	<b>.77</b>	-.03
Actual grade 0-8 mom completed	-.12	<b>-.38</b>	.00
Actual grade 9-11 mom completed	-.19	<b>-.30</b>	.01
Months mom worked in past year	-.00	<b>.30</b>	.05
Hours per week mom works for pay	-.04	<b>.29</b>	.08
Mom has high school diploma or GED	-.23	<b>-.37</b>	.00
Factor 3: Type of Financial Childcare Support Household Received			
TANF helps pay for relative care	-.06	-.00	<b>.99</b>
Other helps pay for relative care	-.06	-.00	<b>.99</b>
Social services/welfare helps pay for care	-.06	.00	<b>.99</b>

Note: N = 11,684. Items with factor loadings < .3 are omitted. Bold indicates a factor loading above .3 and to indicate which factor (1, 2 or 3); TANF = Temporary Assistance for Needy Families

Table 2. Results from the Before- and After-School Programs and Activities (ASPA; 2005) Survey

The first factor was defined by “Father’s Educational Attainment”. The second factor contained variables loading highly on a component of mother or mother-figure’s education, resulting in “Mother’s Educational Attainment.” The third factor was labeled “Types of Financial Childcare Support a Household Received.”

*Early Childhood Program Participation (ECPP, 2005)*

The ECPP survey (see Table 3) used interviews with 7,209 parents of children aged three through third grade, children from birth through age six, and those not yet in kindergarten.

ECPP items	Factor loading		
	1	2	3
Factor 1: Father’s Educational Attainment			
Educational Attainment of child’s father/guardian	<b>.99</b>	.08	-.04
Highest grade/yr of school dad completed	<b>.97</b>	.05	-.03
Dad enrolled in school	<b>.75</b>	.12	-.05
Months dad worked in past year	<b>.74</b>	.12	-.03
Dad worked for pay last week	<b>.70</b>	.10	-.05
Highest level of parent/guardian education	<b>.69</b>	-.08	-.02
Hours per week dad works for pay	<b>.65</b>	.14	-.03
Work status of dad/stepfather/fosterdad/guard.	<b>.59</b>	.06	-.05
Dad’s ease of leaving work	<b>.57</b>	.06	-.02
Household income below or above \$50K	<b>.54</b>	-.15	-.01
Total household income range	<b>.47</b>	-.14	-.00
Family received Medicaid in past 12 months	<b>.47</b>	-.14	-.01
Family received WIC in past 12 months	<b>.41</b>	-.13	-.02
Family received food stamps in past 12 months	<b>.40</b>	-.08	-.02
Own or rent home or other arrangement	<b>-.39</b>	.08	.03
Percent under 18 below poverty line	<b>-.35</b>	.04	.05
Factor 2: Mother’s Work Status			
Work status of mom/stepmom/fostermom/guard.	-.04	<b>.97</b>	-.03
Mom looking for work in past four weeks	-.05	<b>.95</b>	-.03
Mom worked for pay last week	-.09	<b>.85</b>	-.02
Hours per week mom worked for pay	-.02	<b>.84</b>	.05
Mother/guardian works full time	-.02	<b>-.84</b>	.05
Mom’s ease of leaving work	.05	<b>-.80</b>	.05
Months mom worked in past year	.06	<b>-.77</b>	.05
Factor 3: Type of Financial Childcare Support Household Received			
Other helps pay for relative care	-.06	-.11	<b>.98</b>
TANF helps pay for care	-.06	-.11	<b>.98</b>
Social services/welfare helps pay for care	-.05	.11	<b>.98</b>

Note: N = 7,209. Items with factor loadings < .3 are omitted. Bold indicates a factor loading above .3 and to indicate which factor (1, 2 or 3)

**Table 3.** Results from a factor analysis of the Early Childhood Program Participation (ECPP; 2005)

The first factor was indicated by items which loaded highly on a latent variable again labeled “Father’s Educational Attainment.” The second was labeled “Mother’s Work Status” as it related to the employment status and hours of the mother in a household. The third factor was indicated by high loadings on items related to employer, state, or other sources of financial aid and was labeled “Type of Financial Childcare Support a Household Received.”

*Parent and Family Involvement in Education (PFI; 2007)*

The PFI survey (see Table 4) was completed with the parents of 10,681 children in kindergarten through 12<sup>th</sup> grade.



PFI items	Factor loading		
	1	2	3
Factor 1: Father's Educational Attainment			
Dad enrolled in school	<b>.98</b>	-.05	-.02
Dad worked for pay last week	<b>.87</b>	-.10	-.00
Months dad worked in past year	<b>.86</b>	.01	-.04
Educational attainment of child's resident dad/guardian	<b>.86</b>	.20	-.03
Highest grade/yr dad completed	<b>.78</b>	.26	-.03
Work status of dad/stepdad/foster dad/guard.	<b>.76</b>	-.13	.00
Hours per week dad works for pay	<b>.74</b>	.05	-.03
Factor 2: Mother's Educational Attainment			
Educational attainment of child's resident mom/guardian	.11	<b>.97</b>	.01
Highest grade/yr mom completed	.13	<b>.97</b>	.01
Mom enrolled in school	-.03	<b>.43</b>	-.00
Months mom worked in past year	-.06	<b>.34</b>	-.02
Mom has high school diploma or GED	-.13	<b>-.30</b>	.01
Actual grade 0-8 mom completed	-.06	<b>-.30</b>	-.02
Factor 3: Type of Financial Support Household Received			
Receives services from other source	-.03	-.01	<b>.99</b>
Receives services from state/local/social services	-.03	.00	<b>.98</b>

Note. N = 10,681. Items with factor loadings < .3 are omitted. Bold indicates a factor loading above .3 and to indicate which factor (1, 2 or 3)

**Table 4.** Results from a factor analysis of the Parent and Family Involvement (PFI; 2007) Survey

The factor loadings for the PFI survey suggested an initial factor was once again labeled “Father’s Educational Attainment,” with a few income-related variables and the pattern of loadings for the second factor led to the label “Mother’s Educational Attainment.” Lastly, factor three was labeled “Type of Financial Support a Household Received.”

*School Readiness (SR; 2007)*

The SR survey data (see Table 5) were provided by 2,633 children aged 3-6 and kindergarten through second grade.

School readiness items	Factor loading		
	1	2	3
Factor 1: Father's Educational Attainment			
Educational attainment of child's resident dad/guardian	<b>.99</b>	-.11	-.03
Highest grade/yr dad completed	<b>.97</b>	.09	-.02
Dad enrolled in school	<b>.74</b>	.08	-.02
Months dad worked in past year	<b>.73</b>	.12	-.04
Dad worked for pay last week	<b>.70</b>	.08	-.01

Hours per week dad works for pay	<b>.64</b>	.14	-.04
Work status of dad/stepdad/foster dad/guard	<b>.57</b>	.04	.00
Factor 2: Mother’s Work Status			
Work status of mom/stepmom/foster mom/guardian	-.04	<b>.99</b>	.02
Mom looking for work in the past four weeks	-.05	<b>.92</b>	.01
Mom worked for pay last week	-.10	<b>.87</b>	.01
Mom/guardian works full time	-.01	<b>.86</b>	.02
Hours per week mom works for pay	.01	<b>-.83</b>	-.02
Months mom works in past year	.03	<b>-.72</b>	-.02
Factor 3: Type of Financial Support Household Received			
Receives state/local/social services	-.04	.03	<b>.99</b>
Receives services from other source	-.05	.03	<b>.98</b>

Note. N = 2,633. Items with factor loadings < .3 are omitted. Bold indicates a factor loading above .3 and to indicate which factor (1, 2 or 3)

Table 5. Results from a factor analysis of the School Readiness (SR; 2007) Survey

A pattern of loadings similar to other surveys led to labeling the first factor for SR “Father’s Educational Attainment,” although a few items were inconsistent. The other factors were labeled “Mother’s Work Status” and “Type of Financial Support a Household Received.”

Early Childhood Program Participation (ECPP; 2019)

The ECPP (see Table 6) contained a total of 7,092 participants who were parents of children aged three through third grade, children from birth through age six, and those not yet in kindergarten.

ECCP items	Factor loading		
	1	2	3
Factor 1: Socioeconomic status as it relates parent/guardian employment status			
First parent/guardian employment status	<b>-.53</b>	.10	.07
Factor 2: Socioeconomic status as it relates to household total income			
Total income	-.08	<b>.90</b>	.25
Factor 3: Socioeconomic status as it relates to household receiving public assistance			
Received food stamps in past 12 months	-.03	-.12	<b>.55</b>

Note. N = 7,092. Items with factor loadings < .3 are omitted. Bold indicates a factor loading above .3 and to indicate which factor (1, 2 or 3)

Table 6. Results from a factor analysis of the Early Childhood Program Participation (ECPP; 2019) Survey

Three factors emerged as salient after factor analysis. Factor 1 related to employment status and months and hours worked. Factor 2 is concerned with socioeconomic status as it relates to household total income. Factor 3 related to receiving public benefits and correlates with total income.

Parent and Family Involvement (PFI, 2019)

The PFI (see Table 7) had a total of 14,075 participants and surveyed children in kindergarten through 12<sup>th</sup> grade.

PFI items	Factor loading	
	1	2
Factor 1: Socioeconomic status as it relates to age respondent became a parent		
First parent/guardian age when became parent	.86	-.03
Factor 2: Socioeconomic status as it relates to number of hours worked by first parent/guardian		
First parent/guardian hours worked per week	.03	.53

Note. N = 14,075. Items with factor loadings < .3 are omitted.

Table 7. Results from a Factor Analysis of the Parent and Family Involvement (PFI; 2019) Survey

Two factors emerged from this analysis. Factor 1 was the most salient and related to the age the participant became a parent, which related directly to SES and grade level completed. Findings also showed a relationship between hours a parent worked and highest grade level completed. Factor 2 related to hours worked. Thus, the age when the participant became a parent may have impacted their educational level.

Validity Evidence

This research operationalizes SES according to the multidimensional approach suggested by Hauser and Warren<sup>1</sup>, in which SES is considered a latent variable encompassing householder occupation, education, and income. This is consistent with identified factors such as “Father’s Educational Attainment” and “Type of Financial Support a Household Received,” which correspond directly to Hauser and Warren’s conceptualization of SES. This helped us to ensure a comprehensive approach to SES, better catching its complex nature than the one-dimensional measures discussed earlier.

Further empirical evidence of validity is provided by a more detailed examination of the factor loadings and items, or, in other words, the bivariate correlations of the actual items and factors obtained. Several items from each survey were compared with their associated factors. To begin, the AE/AELL 2005 survey displayed a first factor that was labeled “Education Related to Employment.” Its bivariate Pearson correlations with similar items were high. For example, a .98 correlation was found between this factor and the items “Earned college credit” and “Get a raise or promotion” as reasons for the outcomes of pursuing after-school education.

The second factor in this survey was labeled “Higher Educational Attainment.” As such, it had a strong correlation with items such as “Earned college credit” at .97. However, this factor correlated at -.04 with items such as “Took basic skills classes” and “Took GED prep classes,” items it was not expected to correlate with. Lastly, factor three was labeled, “Lower Educational Attainment.” In this case, the opposite effect was observed. Higher correlations were observed for items related to lower educational attainment, whereas higher educational attainment correlated negatively. For example, the item labeled, “Grade 9-11 completed” correlated with this factor at .54. Bivariate correlations provide evidence of an educational component within the NHES survey database. This educational component can be ranked hierarchically, and thus can be used in mapping it onto the educational component of Hauser and Warren’s<sup>Error! Bookmark not defined.</sup> definition.

Next, the first factor in the ASPA 2005 survey was labeled, “Father’s Educational Attainment.” Evidence for this was the fact that correlations between this latent factor and the items “Dad enrolled in school” and “Work status of dad/step/fosterdad/guardian” were high at .93 and .72 respectively. However, low correlations appeared for items related to mother’s educational attainment. High correlations appeared for this factor among items related to father’s enrollment in school, educational attainment, months work, highest grade completed, worked last week, hours worked and work status, income, and financial support household received. These items will be discussed later.

The second factor in this survey was labeled, “Mother’s Higher Educational Status.” Items correlating strongly with this factor included “Highest grade/yr mom completed”, “Highest level of parent/guardian education” and “Educational attainment of child’s mother/guardian” at .79 to .77 respectively. However, items that showed a weak correlation involved mother’s lower educational attainment as well as father’s educational status and items related to mother’s work status. These correlations provided further validity evidence for the specific label of this factor. Negative loadings were found for the items "Actual grade 0-8 completed," "Actual grade 9-11 completed," and "Mom has high school diploma or GED". These items likely have negative loadings because they oppose the second factor. These items contain lower values for the factor "mother's educational attainment," while positive loadings would correspond to higher levels of educational attainment (e.g., some college, college degree).

The third factor was labeled “Types of Financial Childcare Support Household Received.” Correlations in the .99 range were found for items related directly to childcare householder support. However, other types of financial support such as the household receiving support from WIC in the past 12 months and whether the family receives support from state, local, or social services was unrelated to this third factor. Like the AE survey, an educational element emerged that can be ranked vertically (high/low). A financial component also emerged that related to financial support a household receives for childcare. This factor is positioned as a strong link to Hauser and Warren’s Error! Bookmark not defined. component of SES involving income; for example, a household requiring financial support for childcare may rank lower in income.

The ECPP 2005 survey displayed similar corroboration between factors and items. Factor one was labeled “Father’s Education Attainment” and correlated strongly at .97 with the item “Highest grade/yr of school dad completed,” but at .50 with the respective item related to the mother’s highest grade/yr of school completed. This is an indication that the item was related to the father’s educational attainment and not the mother’s. However, items related to father’s work status also correlated strongly with this factor. The negative factor loadings for “Own or rent home or other arrangement” and “Percent under 18 below poverty line” suggest that these items are inversely related to the underlying factor of “Father’s Educational Attainment.” Homeownership is associated with a lower number, thus it is positively skewed and generating a negative factor. Percentage of children living below the poverty line represents a socioeconomic disadvantage that contrasts with the positive socioeconomic indicators linked to higher education.

Next, the second factor was labeled “Mother’s Work Status” and correlated strongly on items related specifically to this factor such as “Mom worked for pay last week” at .85, but showed a weaker correlation on items related to the number of hours the father works for pay at .14, demonstrating that this factor pertained specifically to a mother’s work status. This item correlated weakly or negatively with items related to income or household financial support. A factor such as “Mother’s Work Status” permits a connection between this factor and Hauser and Warren’s Error! Bookmark not defined. definition as having an “occupation” component. Future research might examine the precise ranking of this factor. Some negative loadings were found for the second factor labeled, “Mother’s work status”; specifically, the items: “Mother/guardian works full time”, “Mom’s ease of leaving work” and “Months mom worked in past year.” These are negative because they have an inverse relationship to the theme of this factor. These items are ranked with multiple options for the employment values, potentially reflecting the opposite endorsement pattern. Lastly, factor three, called “Types of Financial Childcare Support Household Received,” correlated positively with the three types of financial support for childcare that this survey examined. However, it, correlated in the low or negative range for items related to state support, educational attainment and parent’s work status, and questions related to income.

Next, the PFI 2007 survey demonstrated similar evidence of validity. For example, the first factor was labeled “Father’s Educational Attainment.” An item that correlated highly with this factor was, for example, “Educational attainment of child’s resident dad/guardian” at .86, but another item called, “Highest grade/yr mom completed” correlated at only .13, indicating that this factor pertained most significantly to the householder father’s educational attainment. It should be noted that high correlations appeared for this factor among items related to a father’s work status, indicating that this factor may not only relate to father’s educational attainment, but may also have an educational component. Again, these discrepancies will be unpacked in the discussion section. The second factor of “Mother’s Educational Attainment” loaded high on “Highest grade/yr mom completed” at .97, but low on an item related to the father’s school status, “Dad enrolled in school,” at -.05. This evidence suggests that the factor pertained to the householder mother. Lastly, the third factor, called “Type of Financial Support Family Received,” correlated very strongly with the items related to a family’s sources of financial support, but in the negative range on items related to parental education and work status. Like the ASPA 2005 survey, the PFI 2007 survey consisted of 2 items that loaded negatively, “Mom has HS diploma or GED” and “Actual grade 0-8 mom completed”. Again, these items may have negative loadings because these two variables are associated with lower educational attainment.

As discussed among the results of the previous surveys, the father and mother’s educational attainment provides a direct connection to the educational component of SES. Reflecting similar connections as the financial childcare support a family receives as found in the ASPA and ECPP surveys, “Type of Financial Support a Household Received” supplies a factor that again maps onto the “income” component of SES as used in this paper.

The first factor in the SR 2007 survey was labeled, “Father’s Educational Attainment.” Validity evidence for this factor came from high correlations among items relating to it such as “Educational attainment of child’s resident dad/guardian” at a .99 correlation and “Dad enrolled in school” at .74. However, items that correlated lower with this factor are those related to the educational status of the mother. Items related to family financial support correlated in the negative range for this factor. This provided further empirical evidence that this factor, in fact, relates to a father’s educational attainment and was thus named appropriately.

However, items related to the father's number of paid work hours and the father's work status also demonstrated high correlations.

A second factor was labeled "Mother's Work Status." Items correlating strongly were directly related to the householder mother's work status, indicating whether she is employed but not work duration. A negative correlation appeared for items related to duration of work such as "Hours per week mom work for pay" at -.83. Items related to householder father's work status and duration all demonstrate low or negative correlations, adding further evidence that this factor was related to the householder mother. Two negative loadings emerged in this pattern of factor analysis, "Hours per week mom works for pay" and "Months mom works in past year". This may indicate different employment patterns or job stability compared to the other items. To conclude, the last factor in this survey was called, "Type of Financial Support Household Received." Items directly related to this type of support correlated above .90, such as "Receives state/local/social services" which correlated at .99. "Father's Educational Attainment," "Mother's Work Status" and "Type of Financial Support a Household Received" grant direct links to each element of Hauser and Warren's<sup>Error! Bookmark not defined.</sup> definition of SES.

The ECPP 2019 survey had one item load as negative, "First parent/guardian employment status" at -.53. This negative loading might represent that this item's particular employment status (such as unemployment or irregular employment) is associated with lower socioeconomic status. This negative loading could indicate that less stable or non-standard employment of the first parent/guardian is inversely related to higher socioeconomic status, reflecting economic instability or disadvantage. Because only one item per factor loaded above .30, the ECPP 2019 survey may be a weak proxy for SES.

The PFI 2007 survey does not contain any negative loadings; however, again, because only one item per factor loaded above .30, the PFI 2019 survey may be a weak proxy for SES.

## DISCUSSION

Results suggest that conceptually and psychometrically rigorous measures of SES should be constructed for the NHES survey database. This would likely enhance inferences based on these data and increase the appeal of the database to the wider educational research community. For example, Breit-Smith *et al.*<sup>32</sup> utilized the 2005 administration of the ECPP survey in the NHES survey database to investigate children with a developmental disability, children with a single disability in comparison to children with multiple disabilities, and speech language disability. The results of the study indicate that it is crucial to take emerging literacy and home literacy experiences into account when evaluating the talents of young patients with disabilities and making clinical recommendations. In another study, Park *et al.*<sup>33</sup> used the 2005 NHES survey ASPA survey to investigate Latino immigrant children who are disadvantaged and vulnerable due to their limited English skills and a lack of educational resources. The findings prompt consideration of whether standard after-school programs are suitable for Latino children from immigrant homes. Further, these findings emphasize the necessity of including culturally appropriate elements in the curriculum for regions with a high concentration of Latino immigrant households.

Analyses of the NHES surveys produced results consistent with the educational facet and, to a lesser extent, the occupational facet. Generally, items correlated highly with factors they would be expected to correlate with, and vice versa. The survey data have further applications if used as proxy measures of SES by educational researchers such as in the exploration of such questions as whether parents with a higher income are more likely to be involved in their child's homework.<sup>34</sup>

Using Hauser and Warren's<sup>Error! Bookmark not defined.</sup> multidimensional conceptualization of SES involving the facets of householder education, occupation, and income, analyses of the NHES surveys produced results generally consistent with the educational facet and the occupational facet. Factors labeled "Father's Educational Attainment," "Mother's Educational Attainment," "Mother's Work Status," "Type of Financial Support a Household Received," and "Type of Financial Childcare Support a Household Received" appeared consistently across surveys. This consistency of socioeconomic factors provides preliminary evidence of construct validity.

Further evidence of construct validity was provided through an examination of the bivariate correlations between the factors and the items in each survey. Primarily, items correlated highly with factors they would be expected to correlate with, and the same can be said for low correlations. However, some items related to the factor "Father's Educational Attainment" presented a special challenge. While this factor correlated strongly (above .95) with items directly related to this construct such as father's work status, number of hours worked, a father's income and household financial support, its correlation with other items such as "Mom has high school diploma or GED," "Mother's Work Status," and "Type of Financial Support a Household Received" were problematic. In this preliminary exploratory analysis, the factor structure found was coherent overall, but the discrepancies found require further study.

An interesting feature of the results, with the exception of the AE survey, was the appearance of a factor characterizing the type of financial support received. This variable was composed of employer, state, and other support sources and may turn out to be a

suitable measure of the income component of SES as research continues to show that it is generally inappropriate to use direct measures of income, such as earnings.<sup>17</sup>

Our findings confirm the utility of Hauser and Warren's multidimensional operationalization of SES for capturing these nuances. There are other operationalizations that provide useful light, focusing on single indicators such as parental education or geographic location, for example, but which could not capture the breadth and complexity of socioeconomic factors addressed by Hauser and Warren's model. The use of a multidimensional approach in this study underlies the robustness and theoretical grounding for the analysis of SES.

#### *Number of Factors Extracted*

Upon examining the surveys that were able to produce two, three or four factors, in part based on the interpretation of the scree plot, utilizing three factors permitted the clearest solution in terms of its application to an SES construct. To examine this empirically, EFA was conducted specifying the extraction of two, three or four factors. For example, in relation to the ECPP 2005 survey, EFA specified the extraction of two factors. The result was that the first factor grouped items related to father and mother's educational attainment, father's work status, and types of financial support a household receives. The second factor included items related to financial childcare support such as Temporary Assistance for Needy Families. Items related to mother's work status did not load high enough (above the .30 threshold) to be included.

Performing EFA on the ECPP 2005 survey for four factors grouped variables related to father's educational attainment and father's work status as factor one while factor two clearly delineates items related to mother's work status. Factor three appeared to group items related to father and mother's educational attainment, income range, and household financial support. This factor may be combined to a factor that can be called "earning potential" or "income." Factor four was a mixture of other items and difficult to define. In short, a three-factor solution provided the clearest factor structure and factor interpretation for the 2005 and 2007 surveys, while a two-factor solution appeared most suitable for the 2019 NHES surveys.

Although a four-factor deduction of the items on the ECPP 2005 survey lends itself to an expanded conceptualization of factors that can be mapped on to the components of Hauser and Warren's<sup>Error! Bookmark not defined.</sup> definition and may be ranked hierarchically, it is cumbersome, especially given the nature of factor three. Although the scree plot for this ECPP survey suggests four factors, examining the survey with two or four factors does not allow the factors to be examined in a more detailed manner, which aids in applying the factors theoretically to Hauser and Warren's<sup>Error! Bookmark not defined.</sup> definition. However, the importance of a solution consisting of four factors should not be understated, and such a solution may prove useful in future research such as when exploring alternative definitions and conceptualizations of SES, such as in the field of sociology when considering facets of social stratification or social mobility. However, in this research, the fourth factor was set aside temporarily in favor of the three-factor solution with the specific focus on Hauser and Warren's<sup>Error! Bookmark not defined.</sup> definition of SES.

#### *LIMITATIONS*

One limitation was that the 2005 and 2007 surveys contained more items that factor loaded and correlated with one another and, thus, exhibited stronger themes related to SES. However, with the two 2019 surveys (ECPP and PFI), factor loadings were lower, and it was more difficult to determine which factors emerged as prominent. Although the SES factors were similar across all surveys, the 2019 surveys had two factors while the 2005 and 2007 surveys had three. In terms of participants, the surveys had an adequate sample size and a diverse sample of the US population. Future educational researchers may build on this work to further explore these datasets and adjust what items to include, as well as explore alternative rotation and extraction methods.

#### *CONCLUSIONS*

Overall, the NHES survey database has much to offer educational researchers because of its detailed focus on pre-primary education, childcare utilization, school safety and discipline, and adult education. For example, the ECPP survey from the 2005 NHES survey database was investigated to explore whether there was a difference in home literacy practices, and emergent literacy skills of children with and without developmental disabilities.<sup>32</sup> Furthermore, Constanzo and Magnuson<sup>35</sup> use the NHES survey ECPP data to examine differences in childcare arrangements between families with children who have a disability and families with children who do not. Using the NHES surveys as a proxy measure of SES may lend itself to using SES-related variables as indicator variables in statistical models. For example, using a similar database to examine if SES is a hidden factor behind four indicators (poverty ratio, personal earnings, educational attainment, and employment status), Martinez *et al.*<sup>36</sup> analyzed data from the 2013 National Health Interview Survey. In a second example, Avvisati<sup>37</sup> reviewed measures and indicators of SES in PISA, a large-scale international assessment of student achievement. Their findings explore the basis for the construction of a composite measure of SES in PISA and other large-scale surveys. Furthermore, SES is a crucial factor affecting psychological well-being and academic achievement and should be considered when designing educational programs and curriculum (Navarro-Carrillo *et al.*, 2020).<sup>38</sup> For these reasons, increased research on the correlations between SES and education is essential.

As this work was exploratory in nature, our aim was to examine what factors emerged from the analyses of the seven surveys described in this work to explore the extent to which these factors could be used as proxy measures for SES. Further research could capitalize on these results and examine alpha reliability tests of the factors that were included. The benefit of an alpha reliability analysis is to strengthen the argument that certain items act as a proxy measure for SES. An SES proxy instrument using alpha reliability analysis would help indicate the degree to which lower-level factors are impacting the overall factor. Consequently, lower-level factors could potentially be dropped, and the alpha reliability analysis further examined.

By adhering to the definition of SES provided by Hauser and Warren, we were better placed to clearly link the variables explored in the NHES surveys and SES. This is because some studies may use narrow operationalizations, which emphasizes the relevance of choosing a model representing the complexity of socioeconomic factors. Future studies on SES should be done with continued attention to this multidimensionality, for it accommodates a better grasp of the effects of these indicators on educational outcomes.

Our results support the argument that conceptually and psychometrically rigorous measures of SES can be developed in most of the NHES surveys. However, the ECPP and PFI 2019 surveys demonstrated a weak ability to act as proxy measures of SES due to less factor loadings. The results, while preliminary, suggest that rigorous measures of SES can be constructed based on Hauser and Warren's<sup>1</sup> model which used income, education, and occupation. The availability of such measures should enhance inferences based on these data and simultaneously increase the appeal of this database to the educational research community.

## REFERENCES

1. Hauser, R. M., and Warren, J. R. (1997) 4. Socioeconomic indexes for occupations: A review, update, and critique. *Sociological Methodology* 27(1), 177-298. <https://doi.org/10.1111/1467-9531.271028>
2. Nesler, M. S., and Gunnarson, R. (2000) Adult education in the 1990s: An analysis of the 1995 National Household Education Survey Database. Paper presented at the Bridges to the Future: Building Linkages for Institutional Research. <http://www.bates.edu/Prebuilt/papers2000.pdf#page=149>. Pittsburgh, PA.
3. May, H. (2002) Development and evaluation of an internationally comparable scale of student socioeconomic status using survey data from TIMSS [Doctoral Dissertation]. University of Pennsylvania (AAT 3073031).
4. Silva, T., Cahalan, M., and Lacireno-Paquet, N. (1998). Adult education participation decisions and barriers: Review of conceptual frameworks and empirical studies [Unpublished manuscript]. <http://nces.ed.gov/pubs98/9810.pdf>
5. Zill, N., Loomis, L. S., and West, J. (1997) The elementary school performance and adjustment of children who enter kindergarten late or repeat kindergarten: Findings from national surveys. *National Household Education Survey. Statistical Analysis Report*. National Center for Educational Statistics. <http://eric.ed.gov/PDFS/ED414076.pdf>
6. Mueller, R. O. (1996) *Mathematical Statistics and Data Analysis*. Cengage Learning, Boston.
7. Kurki, A., Boyle, A., and Aladjem, D. K. (2005) *Beyond free lunch: Alternative poverty measures in educational research and program evaluation*. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada.
8. West, C. E., Goldschmidt, A. B., Mason, S. M., and Neumark-Sztainer, D. (2019) Differences in risk factors for binge eating by socioeconomic status in a community-based sample of adolescents: Findings from Project EAT. *International Journal of Eating Disorders* 52(6), 659-668. <https://doi.org/10.1002/eat.23079>
9. Carroll, S. J., Dale, M. J., Niyonsenga, T., Taylor, A. W., and Daniel, M. (2020) Associations between area socioeconomic status, individual mental health, physical activity, diet and change in cardiometabolic risk amongst a cohort of Australian adults: A longitudinal path analysis. *PLOS ONE* 15(5), e0233793. <https://doi.org/10.1371/journal.pone.0233793>
10. Reiss, F., Meyrose, A. K., Otto, C., Lampert, T., Klasen, F., and Ravens-Sieberer, U. (2019) Socioeconomic status, stressful life situations and mental health problems in children and adolescents: Results of the German BELLA cohort-study. *PLOS ONE* 14(3), e0213700. <https://doi.org/10.1371/journal.pone.0213700>
11. Breen, R. (2005) Foundations of Class Analysis in the Weberian Tradition, in *Approaches in Class Analysis*, 31-50.
12. Harwell, M. R., and LeBeau, B. (2010) Student eligibility for a free lunch as an SES measure in education research. *Educational Researcher* 39(2), 120-131. <https://doi.org/10.3102/0013189X10362578>

13. Heppt, B., Olczyk, M., and Volodina, A. (2022) Number of books at home as an indicator of socioeconomic status: Examining its extensions and their incremental validity for academic achievement. *Social Psychology of Education* 25(4), 903-928. <https://doi.org/10.1007/s11218-022-09704-8>
14. Tan, C. Y., Lyu, M., and Peng, B. (2020) Academic benefits from parental involvement are stratified by parental socioeconomic status: A meta-analysis. *Parenting* 20(4), 241-287. <https://doi.org/10.1080/15295192.2019.1694836>
15. Bollen, K. A., Glanville, J. L., and Stecklov, G. (2001) Socioeconomic status and class in studies of fertility and health in developing countries. *Annual Review of Sociology* 27(1), 153-185. <https://doi.org/10.1146/annurev.soc.27.1.153>
16. Oswald, D. P., Zaidi, H. B., Cheatham, D. S., and Brody, K. G. D. (2018) Correlates of parent involvement in students' learning: Examination of a national data set. *Journal of Child and Family Studies* 27(1), 316-323. <https://doi.org/10.1007/s10826-017-0876-4>
17. Mueller, C. W., and Parcel, T. L. (1981) Measures of socioeconomic status: Alternatives and recommendations. *Child Development* 52(1), 15-30. <https://doi.org/10.2307/1129211>
18. Yerdelen-Damar, S., and Peşman, H. (2013) Relations of gender and socioeconomic status to physics through metacognition and self-efficacy. *Journal of Educational Research* 106(4), 280-289. <https://doi.org/10.1080/00220671.2012.692729>
19. Gorusch, R. L. (1983) *Factor Analysis* (2<sup>nd</sup> Ed.). Lawrence Erlbaum Associates, New Jersey.
20. Bilder, R. M., Mukherjee, S., Rieder, R. O., and Pandurang, A. K. (1985) Symptomatic and neuropsychological components of defect states. *Schizophrenia Bulletin* 11(3), 409-419. <https://doi.org/10.1093/schbul/11.3.409>
21. Lyden, P. D., Lu, M., Levine, S. R., Brott, T. G., Broderick, J., and the NINDS rtPA Stroke Study Group. (2001). A modified national institutes of health stroke scale for use in stroke clinical trials: Preliminary reliability and validity. *Stroke*, 32(6), 1310-1317. <https://doi.org/10.1161/01.str.32.6.1310>
22. Henson, R. K., and Roberts, J. K. (2006) Use of exploratory factor analysis in published research: Common errors and some comment on improved practice. *Educational and Psychological Measurement* 66(3), 393-416. <https://doi.org/10.1177/0013164405282485>
23. Costello, A. B., and Osborne, J. W. (2005) Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment and Evaluation* 10(7), 1-9.
24. Fabrigar, L. R., Wenener, D. T., MacCallum, R. C., and Strahan, E. J. (1999) Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods* 4(3), 272-299. <https://doi.org/10.1037/1082-989X.4.3.272>
25. Velicer, W. F., and Jackson, D. N. (1990) Component analysis versus common factor analysis: Some issues in selecting an appropriate procedure. *Multivariate Behavioral Research* 25(1), 1-28. [https://doi.org/10.1207/s15327906mbr2501\\_1](https://doi.org/10.1207/s15327906mbr2501_1)
26. Nunnally, J. C. (1978) *Psychometric theory*, 2nd ed., McGraw-Hill.
27. Comrey, A. L., and Lee, H. B. (1992) *A first course in factor analysis*, 2nd ed., Lawrence Erlbaum Associates.
28. Floyd, F. J., and Widaman, K. F. (1995) Factor analysis in the development and refinement of clinical assessment instruments, *Psychological Assessment* 7, 286-299.
29. Tabachnick, B. G., and Fidell, L. S. (2001) *Using Multivariate Analysis*, Allyn & Bacon, Boston.
30. UCLA: Statistical Consulting Group. (2021) *A Practical Introduction to Factor Analysis: Exploratory Factor Analysis*. OARC Stats. <https://stats.oarc.ucla.edu/spss/seminars/introduction-to-factor-analysis/a-practical-introduction-to-factor-analysis/>
31. IBM Corp. (2011). *IBM SPSS Statistics for Windows* (Versions 20 and 26). Armonk, NY: IBM Corp.
32. Breit-Smith, A., Cabell, S. Q., and Justice, L. M. (2010) Home literacy experiences and early childhood disability: A descriptive study using the National Household Education Surveys (NHES) program database. *Language, Speech, and Hearing Services in Schools* 41(1), 96-107. [https://doi.org/10.1044/0161-1461\(2009/08-0048\)](https://doi.org/10.1044/0161-1461(2009/08-0048))
33. Park, H, Lin, C.-H., Liu, C., and Tabb, K. M. (2015) The relationships between after-school programs, academic outcomes, and behavioral developmental outcomes of Latino children from immigrant families: Findings from the 2005 National Household Education Surveys Program. *Children and Youth Services Review* 53, 77-83. <https://doi.org/10.1016/j.childyouth.2015.03.019>
34. Hartlep, N. D., & Ellis, A. L. (2010) Are household income, gender, and race important in shaping parental involvement in children's education? *Journal of Educational Research and Practice* 2, 1-24. <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=1898d6e0dcb204d14cbf9c6deefe10df5eb3e046>



35. Costanzo, M. A., & Magnuson, K. (2019) How does disability influence child care arrangements for young children? An examination using the NHES ECPP. *Children and Youth Services Review* 99, 210–225.  
<https://linkinghub.elsevier.com/retrieve/pii/S0190740918306741>
36. Martinez, S. A., Beebe, L. A., Thompson, D. M., Wagener, T. L., Terrell, D. R., and Campbell, J. E. (2018) A structural equation modelling approach to understanding pathways that connect socioeconomic status and smoking. *PLOS ONE* 13(2), e0192451. <https://doi.org/10.1371/journal.pone.0192451>
37. Avvisati, F. (2020) The measure of socio-economic status in PISA: A review and some suggested improvements. *Large-Scale Assessments in Education* 8(1), 1-37.
38. Navarro-Carrillo, G., Alonso-Ferres, M., Moya, M., and Valor-Segura, J. (2020) Socioeconomic status and psychological well-being: Revisiting the role of subjective socioeconomic status. *Frontiers in Psychology* 11, 1303.  
<https://doi.org/10.3389/fpsyg.2020.01303>

### ABOUT THE STUDENT AUTHORS

Amanda Delore is an alum of Dean College with a Bachelor of Arts in Psychology and Theatre. They are a recent graduate of Lesley University, where they received a Master's in Drama Therapy and Mental Health Counseling. Amanda began this research when they were an undergraduate student at Dean College. This is Amanda's first publication.

Ashley Jeanty graduated from Dean College with a Bachelor of Arts in Psychology. Ashley began this research when she was an undergraduate student at Dean College. Her research interests include different factors of socioeconomic status, child education, and decreasing challenging behavior in young children. This is Ashley's first publication.

### PRESS SUMMARY

Researchers have successfully developed robust measures of socioeconomic status (SES) for the National Household Educational Surveys (NHES). By employing rigorous methods, they found that measures based on income, education, and occupation enhance the understanding of educational dynamics. This research promises to deepen insights into diverse educational phenomena, from early childhood programs to adult learning. By bridging theory with empirical analysis, this work amplifies the significance of NHES data for educational research, offering valuable tools to explore the intricate relationship between SES and educational outcomes.