

## Longitudinal Links Between Spanking and Children's Externalizing Behaviors in a National Sample of White, Black, Hispanic, and Asian American Families

Elizabeth T. Gershoff  
*University of Texas at Austin*

Jennifer E. Lansford  
*Duke University*

Holly R. Sexton  
*University of Texas at Austin*

Pamela Davis-Kean  
and Arnold J. Sameroff  
*University of Michigan*

This study examined whether the longitudinal links between mothers' use of spanking and children's externalizing behaviors are moderated by family race/ethnicity, as would be predicted by cultural normativeness theory, once mean differences in frequency of use are controlled. A nationally representative sample of White, Black, Hispanic, and Asian American families ( $n = 11,044$ ) was used to test a cross-lagged path model from 5 to 8 years old. While race/ethnic differences were observed in the frequency of spanking, no differences were found in the associations of spanking and externalizing over time: Early spanking predicted increases in children's externalizing while early child externalizing elicited more spanking over time across all race/ethnic groups.

A large body of research has demonstrated that parents' use of spanking or other forms of physical punishment is associated with high rates of children's externalizing behavior problems, including aggression and antisocial behaviors (Gershoff, 2002). Much of this research has been based on middle-class White American families, and thus the generalizability of its conclusions to other cultures, particularly to other race/ethnic groups within the United States, has been questioned. Deater-Deckard and Dodge (1997) have proposed that the cultural context in which spanking occurs will alter the meaning of discipline to the child, which in turn determines whether positive or negative outcomes result. Spanking, they argued, may have less detrimental effects on child behavior in cultural contexts within which it is normative, and thus more frequent, and to have more detrimental effects on

child behavior in cultures within which spanking is less normative.

Race/ethnic group differences in how frequently parents use spanking have indeed been documented. Compared with White parents, Black parents believe spanking to be more appropriate and effective (Flynn, 1998) and report spanking more often than White parents (Grogan-Kaylor & Otis, 2007; Slade & Wissow, 2004). Hispanic parents report spanking at rates similar to White parents and less than Black parents (Grogan-Kaylor & Otis, 2007; Slade & Wissow, 2004). These differences in usage suggest that normative beliefs about spanking are more positive among Black families than among White or Hispanic families.

In addition to these observed differences in frequency of spanking, several studies have reported differences in the extent to which spanking or other forms of physical punishment are associated with child behavior. Two studies of children in middle childhood found that spanking was associated with more externalizing among White children but associated with less externalizing, or not associated

---

This project was supported by the National Science Foundation (Grant BCS 0818478) through the Developmental Learning Science—IRADS Collaborative on the Analysis of Pathways from Childhood to Adulthood at the University of Michigan.

Correspondence concerning this article should be addressed to Elizabeth T. Gershoff, Department of Human Development and Family Sciences, University of Texas at Austin, 1 University Station A2702, Austin, TX 78712. Electronic mail may be sent to liz.gershoff@austin.utexas.edu.

© 2011 The Authors  
Child Development © 2011 Society for Research in Child Development, Inc.  
All rights reserved. 0009-3920/2012/8303-0008  
DOI: 10.1111/j.1467-8624.2011.01732.x

significantly with externalizing, among Black children (Gunnoe & Mariner, 1997; McLeod, Kruttschnitt, & Dornfield, 1994); two articles (from the same research study) found similar results across the ages of 5–13 years using a composite measure that included spanking, slapping, and hitting (Deater-Deckard, Dodge, Bates, & Pettit, 1996; Lansford, Deater-Deckard, Dodge, Bates, & Pettit, 2004). Yet other research has found negative child behaviors associated with a measure of spanking or a physical punishment composite (including spanking, slapping, and hitting) across White and Black families with children in elementary or high school (Lau, Litrownik, Newton, Black, & Everson, 2006; Pardini, Fite, & Burke, 2008) as well as across White, Black, and Hispanic families with children in elementary or middle school (Grogan-Kaylor, 2005). The contradictory nature of these findings has meant that the question of whether one's ethnicity moderates the effects of spanking remains unresolved. Indeed, whether race/ethnic differences moderate the impact of spanking on children remains central to discussions about the acceptability of spanking as a means of disciplining children (Durrant, 2008).

The present study aims to examine whether there are race/ethnic group differences in how often parents spank and the extent to which spanking is associated with children's externalizing behaviors when examined in a large and nationally representative sample of American children. We first hypothesize that Black parents will report the highest rates of spanking. Second, we hypothesize that once we have accounted for group differences in mean rates, there will be no differences among Blacks, Whites, Hispanics, and Asians in the extent to which spanking is associated with increased child externalizing behavior over time.

This study builds on the previous literature in several ways. First, the study expands the race/ethnic groups considered beyond Black and White to also include Hispanic and Asian American families. Spanking among Hispanics and Asian Americans has been significantly understudied and thus this project contributes to the field's very limited knowledge about these families. Second, the cross-lagged model tested in this article incorporates the transactional nature (Sameroff, 2009) of the associations between spanking and externalizing over time. The model takes into account initial levels of externalizing and spanking and thus focuses on change in the longitudinal associations. Third, the focus of this study is on parents' use of spanking and not of

their use of harsher methods of physical punishment (e.g., using an object, beating) that may vary more across race/ethnic groups in their acceptability. Fourth, this study uses ratings of spanking from mothers and ratings of externalizing from teachers in order to avoid problems of shared method variance. Finally, the model controls for several important indicators of a family's socioeconomic status that might be confounded with race/ethnicity in the United States.

## Method

### Sample

The Early Childhood Longitudinal Study's Kindergarten Cohort of 1998–1999 (ECLS-K) followed a nationally representative sample of 21,260 children who were enrolled in 944 kindergarten programs during the 1998–1999 school year (West, Denton, & Reaney, 2000). For the analyses reported here, we used data from the spring waves of the kindergarten (1999) and third-grade (2002) school years (Rathbun & West, 2004). We restricted the sample to children who: (a) were uniquely identified as being in one of the four main race and ethnic groups of the study, (b) had mothers as the parent respondents for the study, (c) did not have a twin in the sample, and (d) had a third-grade school identifier for clustering purposes. These restrictions resulted in a sample of 11,044 children and families. The excluded sample differed significantly from the included sample on levels of parent education, family income/income-to-needs, father's employment status, and parent's marital status at both kindergarten and third grade. The excluded sample had a significantly higher mean for externalizing behavior at kindergarten but there were no differences in third grade; they did not differ in levels of spanking at kindergarten or at third grade. Additionally, there were significantly more Whites and Hispanics and fewer Blacks and Asians in the included versus the excluded sample. To adjust for these differences, we used a longitudinal non-response weight, clustered by third-grade school ID, and incorporated time-varying covariates, including those significantly different between the samples.

At the kindergarten spring assessment, the mean age of the children in the sample was 6 years and 2 months,  $SD = 4$  months; 49% of the children were female. A total of 64% of the children were White/Non-Hispanic ( $n = 7,057$ ), 19% were Hispanic ( $n = 2,096$ ), 12% were Black/Non-Hispanic ( $n = 1,352$ ),

and 5% were Asian ( $n = 539$ ). Eighty-one percent of the children came from two-parent families. The average age of the mothers at the kindergarten wave was 33 years ( $SD = 6$  years). Ninety-five percent of mothers' race/ethnicity was accurately captured by using child race/ethnicity as a marker; thus, we used child race/ethnicity to indicate family race/ethnicity.

### Measures of Key Constructs

*Spanking.* At both waves, mothers were asked whether they had ever spanked the child, and, "About how many times, if any, have you spanked {CHILD} in the past week?" Responses to this latter question ranged from 0 to 30 times per week. A scale was created such that those mothers who never spanked were coded as 0, those who spanked but had not in the last week were coded as 1, and those who had spanked in the last week were coded with their stated frequency of spanking plus 1. Responses above five times per week (approximately 1% of the sample) were top-coded at 6.

*Child externalizing behavior.* Teachers provided ratings of children's externalizing behavior problems in the spring of each school year using an adapted version of the Social Skills Rating Scale (Gresham & Elliott, 1990). Teachers were asked to report how frequently the child does such things as argue, fight, or get angry on a scale of 1–4, with 1 indicating *never* and 4 indicating *most of the time*. The scale had strong split-half reliability across waves ( $\alpha_s = .90$  in kindergarten and  $.89$  in third grade; Pollack, Atkins-Burnett, Rock, & Weiss, 2005).

### Measures of Control Variables

*Family income-to-needs.* Families' income-to-needs ratios from the spring of kindergarten were included as a continuous control variable; 18% of families had income-to-needs ratios that fell below the poverty thresholds for 1998 (e.g., \$16,530 for a family of two adults and two children), nearly identical to the national poverty rate for children of 18.9% in 1998 (Dalaker, 1999). Income-to-needs ratios were not available at third grade; to approximate the ratio, we included both the categorical family income measure and family size (range = 2–14,  $M = 4.55$ ,  $SD = 1.29$ ) at the time of the third-grade assessment.

*Parent education.* At both waves, the highest level of education attained by either of the child's par-

ents, or by the child's custodial parent, was reported by mothers according to the following ranked categories: 1 = 8th grade or below, 2 = 9th–12th grade, 3 = high school diploma/GED, 4 = vocational school, 5 = some college, 6 = bachelor's degree, 7 = some graduate or professional school, 8 = master's degree, 9 = doctorate or professional degree. The kindergarten mean (4.85,  $SD = 1.94$ ) was between 4 = vocational or technical program and 5 = some college; the third grade mean (5.13,  $SD = 1.94$ ) was at 5 = some college.

*Parents' marital status.* Mothers reported their own marital status at each assessment as married, separated, divorced, widowed, or never married. A dichotomous variable of married (75% at kindergarten, 77% at third) versus not married (25% at kindergarten, 23% at third) was derived for each wave.

*Parents' employment status.* At kindergarten, mothers reported their own and the child's father's work status: not in the labor force (mothers: 29%; fathers: 2%), looking for work (3% and 1%), working part time (less than 35 hr/week: 24% and 3%), and working full time (35 or more hours per week: 45% and 93%). At third grade, the distribution was, for mothers and fathers respectively: not in the labor force (23% and 3%), looking for work (3% and 2%), working part time (26% and 4%), and working full time (49% and 91%). For two-parent families, both parents' work statuses were included as separate covariates at each wave.

### Analyses

Our analyses were conducted using the *Mplus* 6.1 software (Muthén & Muthén, 2010). All models controlled for a set of covariates, namely, child gender, child age (in years), family income-to-needs (or family income and family size), highest parents' education, parents' marital status, mother's and father's employment status, and race/ethnicity. Of these, the time-varying covariates for each wave (child's age, parent's education, income-to-needs, marital status, and employment statuses) were also allowed to covary. Missing data were accounted for through full information maximum likelihood estimation. Robust maximum likelihood estimation was used to account for possible bias in estimates, and the Satorra–Bentler scaled chi-square difference test (Satorra, 2000) was used for model comparisons. To adjust for unequal selection probabilities at both school level and individual child level as well as for nonresponse rates, we used the longitudinal weight C245PW0, which is the base-year parent weight adjusted for nonresponse in third grade

(Tourangeau et al., 2004). All models nested children within their third-grade school to adjust standard errors to account for shared variance at the school level.

## Results

The majority of mothers in the ECLS-K sample (80%) reported at the kindergarten wave that they had used spanking at some point with their children, with some variation by race/ethnic group: 89% of Blacks, 79% of Whites, 80% of Hispanics, and 73% of Asians reported that they had ever spanked their child. At the kindergarten wave, 27% of mothers reported spanking their children in the previous week (40% of Blacks, 28% of Hispanics, 24% of Whites, and 23% of Asians). Fewer mothers (15% of the full sample) reported spanking their children in the previous week at the time of the third-grade interview (23% of Blacks, 21% of Asians, 14% of Whites, and 13% of Hispanics).

Race/ethnic mean differences were found for spanking and externalizing at each wave (see Table 1). Black mothers reported significantly more frequent use of spanking at both kindergarten and third grade than did mothers in any of the three other race/ethnic groups, while Hispanic mothers reported more spanking than White or Asian mothers at kindergarten. Black children were also rated by teachers as having the highest levels of externalizing behavior at kindergarten and third grade. White and Hispanic children were rated as having higher levels of externalizing behavior than Asian children at third grade.

We first tested a cross-lagged path model for the full sample with all race/ethnic groups combined

( $n = 11,044$ ). This model included mothers' use of spanking at kindergarten (hereafter, spanking<sub>K</sub>) predicting child externalizing behaviors at third grade (child externalizing behaviors<sub>3rd</sub>), child externalizing behaviors<sub>K</sub> predicting mothers' use of spanking<sub>3rd</sub>, the autoregressive paths for spanking and child externalizing behaviors, and covariances between the two kindergarten constructs and between the residual variances of the two third-grade constructs. A key aspect of this model is that for each pathway of interest, we have controlled for the initial levels of the dependent variable, and thus the focus is on predicting change in the dependent construct, over and above initial levels.

This model fit data from the full sample well: comparative fit index (CFI) = .98, root mean square error of approximation (RMSEA) = .01, standardized root mean square residual (SRMR) = .01, and  $\chi^2(df = 26) = 83.36, p < .001$ . The standardized path

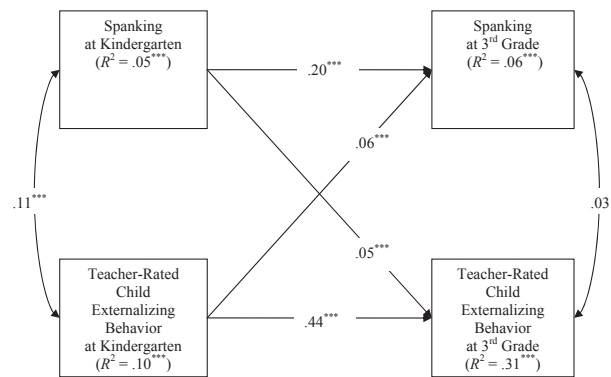


Figure 1. Results from kindergarten to third-grade cross-lag model of spanking and child externalizing behavior for the full sample ( $N = 11,044$ ).

\*\*\* $p < .001$ .

Table 1

Means, Standard Deviations, and Race/Ethnic Group Comparisons for Spanking and Externalizing Behavior at Each Wave

	Family race/ethnicity				ANOVA/chi-square group comparisons <sup>a</sup>
	White ( $n = 7,057$ )	Black ( $n = 1,352$ )	Hispanic ( $n = 2,096$ )	Asian ( $n = 539$ )	
Spanking <sub>K</sub>	1.14 (0.95)	1.59 (1.21)	1.23 (1.01)	1.08 (0.98)	B > W, H, A; H > A, W
Child externalizing behavior <sub>K</sub>	1.59 (0.60)	1.84 (0.72)	1.62 (0.59)	1.44 (0.51)	B > W, H, A; W, H > A
Spanking <sub>3rd</sub>	0.96 (0.71)	1.18 (0.91)	0.93 (0.70)	1.03 (0.80)	B > W, H, A
Child externalizing behavior <sub>3rd</sub>	1.64 (0.56)	1.98 (0.74)	1.68 (0.58)	1.46 (0.48)	B > W, H, A; W, H > A

Note. Standard deviations are presented in parentheses. K = kindergarten; 3rd = third grade. All post hoc comparison  $p$  values have been adjusted for multiple comparisons by Bonferroni corrections. All comparisons noted are significant at  $p < .01$ ; results are available from the first author.

<sup>a</sup>W = White; B = Black; H = Hispanic; A = Asian.



coefficients and  $R^2$ 's for this model are displayed in Figure 1. Each of the cross-lagged paths was significant and were nearly equal to each other in magnitude,  $\beta$  spanking<sub>K</sub>  $\rightarrow$  child externalizing<sub>3rd</sub> = .05,  $p < .001$ , and  $\beta$  child externalizing<sub>K</sub>  $\rightarrow$  spanking<sub>3rd</sub> = .06,  $p < .001$ . To determine whether these paths were statistically equal, we compared this model with a second model in which the two cross-lagged paths were constrained to be equal and found no significant change in  $\chi^2$ , indicating that the parent and child effects could be treated as equal.

We next tested whether the model paths varied by the race/ethnic group of the family. We conducted a multigroup analysis of the cross-lag model, which involved a series of comparative models whereby the cross-lagged model was fit separately for each of the four race/ethnic subsamples and fit with varying numbers of equality constraints for the structural paths of the model. We allowed means and variances to vary freely across groups so that any significant differences among models would isolate group differences in the magnitude or direction of the structural paths. None of the model comparisons was significant, indicating that freeing any or all of the paths to vary across race/ethnic groups did not improve the fit of the model. The fully constrained model, in which all of the paths were set to be equal across the four race/ethnic groups, was thus accepted as the best fitting model: CFI = .97, RMSEA = .01, SRMR = .01, and  $\chi^2(df = 122) = 173.18$ ,  $p < .01$ . Spanking predicted an increase in externalizing behaviors over time, and child externalizing elicited more spanking over time, equally across White, Black, Hispanic, and Asian families.

### Discussion

It is clear from this sample of 11,040 American families with kindergartners that spanking remains a common practice. A large majority of the mothers in this study (80%) reported that they had spanked their children at some point, and 27% of the mothers reported spanking their kindergartner in the week before the interview. Consistent with previous studies (Grogan-Kaylor & Otis, 2007; Slade & Wissow, 2004) and with our first hypothesis, Black mothers in this study reported the most frequent use of spanking at both kindergarten and third grade. In addition, Hispanic mothers spanked more often than White or Asian mothers, but only at kindergarten.

Yet once these subgroup differences in how often mothers spank were controlled, early spanking was

associated with increases in children's externalizing problem behaviors, and externalizing was associated with increases in spanking, for all race/ethnic groups, thus supporting our second hypothesis. The processes examined here do not appear to be bound by race/ethnic group. Rather, the findings that children may be negatively reactive to mothers' spanking and that mothers' spanking may be negatively reactive to children's externalizing suggest that a coercive cycle of parenting (Patterson, 1982) is operating across race/ethnic groups.

We believe that our results differ from the previous studies reporting race/ethnic group differences for four main reasons. First, we used a nationally representative sample that included large subsamples of each of the four race/ethnic groups. Second, we focused on spanking rather than other, more harsh methods of physical punishment; this may account for why our results differ from two studies frequently cited as evidence of a cultural normativeness effect (Deater-Deckard et al., 1996; Lansford et al., 2004). Third, we allowed initial levels of spanking and externalizing to vary across race/ethnic groups so that our test of group differences was focused on change over time controlling for initial levels. Fourth, we simultaneously modeled cross-lagged paths to account for the transactional nature of parent-child interactions (Sameroff, 2009).

The current study is limited in a few ways. First, there is likely to be between-family variability in what is considered to be a "spanking." Families differ both in how a spanking is delivered (i.e., by a hand or by an object) and in when it is delivered (e.g., regularly and deliberately vs. sporadically and emotionally). Second, the effects were small in magnitude and accounted for only a small percentage of the variance in spanking and children's externalizing behavior. However, from an epidemiological standpoint, small yet significant effects can have large impacts when extrapolated to the population as a whole, as we did in this study using sampling weights. Third, as with all of the previous research summarized above, this study focused on children in middle childhood, beyond the peak age of spanking (3 years; Holden, Coleman, & Schmidt, 1995). It will be important for future research to determine if these results replicate with a sample of preschool-aged children. Finally, this study did not measure cultural norms directly; doing so will allow researchers to understand more clearly links between group norms and parenting behavior.

Using a large nationally representative American sample, this study found that the longitudinal associations between spanking and child externalizing

behaviors were the same across several racial and ethnic subgroups. These results suggest that any efforts to reduce both spanking and externalizing behaviors would need to target Blacks, Whites, Hispanics, and Asian Americans. While the approaches would likely vary by cultural group, the message would be the same—spanking children is not associated with better behavior over time.

## References

- Dalaker, J. (1999). *Poverty in the United States: 1998. U.S. Census Bureau, Current Population Reports (P60-270)*. Washington, DC: U.S. Government Printing Office. Retrieved from <http://www.census.gov/hhes/www/poverty/prevcps/p60-207.pdf>
- Deater-Deckard, K., & Dodge, K. A. (1997). Externalizing behavior problems and discipline revisited: Nonlinear effects and variation by culture, context, and gender. *Psychological Inquiry, 8*, 161–175.
- Deater-Deckard, K., Dodge, K. A., Bates, J. E., & Pettit, G. S. (1996). Physical punishment among African American and European American mothers: Links to children's externalizing behaviors. *Developmental Psychology, 32*, 1065–1072.
- Durrant, J. E. (2008). Physical punishment, culture and rights: Current issues for professionals. *Journal of Developmental and Behavioral Pediatrics, 29*, 55–66.
- Flynn, C. P. (1998). To spank or not to spank: The effect of situation and age of child on support for corporal punishment. *Journal of Family Violence, 13*, 21–37.
- Gershoff, E. T. (2002). Corporal punishment by parents and associated child behaviors and experiences: A meta-analytic and theoretical review. *Psychological Bulletin, 128*, 539–579.
- Gresham, F. M., & Elliott, S. N. (1990). *Social Skills Rating System manual*. Circle Pines, MN: American Guidance Service.
- Grogan-Kaylor, A. (2005). Corporal punishment and the growth trajectory of children's antisocial behavior. *Child Maltreatment, 10*, 283–292.
- Grogan-Kaylor, A., & Otis, M. D. (2007). The predictors of parental use of corporal punishment. *Family Relations, 56*, 80–91.
- Gunnoe, M. L., & Mariner, C. L. (1997). Toward a developmental-contextual model of the effects of parental spanking on children's aggression. *Archives of Pediatric and Adolescent Medicine, 151*, 768–775.
- Holden, G. W., Coleman, S. D., & Schmidt, K. L. (1995). Why 3-year-old children get spanked: Parent and child determinants in a sample of college-educated mothers. *Merrill-Palmer Quarterly, 41*, 431–452.
- Lansford, J. E., Deater-Deckard, K., Dodge, K. A., Bates, J. E., & Pettit, G. S. (2004). Ethnic differences in the link between physical discipline and later adolescent externalizing behaviors. *Journal of Child Psychology and Psychiatry, 45*, 801–812.
- Lau, A. S., Litrownik, A. J., Newton, R. R., Black, M. M., & Everson, M. D. (2006). Factors affecting the link between physical discipline and child externalizing problems in Black and White families. *Journal of Community Psychology, 34*, 89–103.
- McLeod, J. D., Kruttschnitt, C., & Dornfield, M. (1994). Does parenting explain the effects of structural conditions on children's antisocial behavior? A comparison of Blacks and Whites. *Social Forces, 73*, 575–604.
- Muthén, L. K., & Muthén, B. O. (2010). *Mplus user's guide* (6th ed.). Los Angeles: Muthén & Muthén.
- Pardini, D. A., Fite, P. J., & Burke, J. D. (2008). Bidirectional associations between parenting practices and conduct problems in boys from childhood to adolescence: The moderating effect of age and African-American ethnicity. *Journal of Abnormal Child Psychology, 36*, 647–662.
- Patterson, G. R. (1982). *Coercive family process*. Eugene, OR: Castalia.
- Pollack, J., Atkins-Burnett, S., Rock, D., & Weiss, M. (2005). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS–K), Psychometric report for the third grade (NCES 2005–062)*. Washington, DC: U.S. Department of Education, National Center for Education Statistics.
- Rathbun, A., & West, J. (2004). *From kindergarten through third grade: Children's beginning school experiences* (NCES 2004-007). U.S. Department of Education, National Center for Education Statistics. Washington, D.C.: Government Printing Office.
- Sameroff, A. J. (2009). The transactional model. In A. J. Sameroff (Ed.), *The transactional model of development: How children and contexts shape each other* (pp. 3–21). Washington, DC: American Psychological Association.
- Satorra, A. (2000). Scaled and adjusted restricted tests in multi-sample analysis of moment structures. In D. D. H. Heijmans, D. S. G. Pollock, & A. Satorra (Eds.), *Innovations in multivariate statistical analysis: A festschrift for Heinz Neudecker* (pp. 233–247). Dordrecht, Netherlands: Kluwer Academic Publishers.
- Slade, E. P., & Wissow, L. S. (2004). Spanking in early childhood and later behavior problems: A prospective study of infants and young toddlers. *Pediatrics, 113*, 1321–1330.
- Tourangeau, K., Brick, M., Byrne, L., Lê, T., Nord, C., West, J., et al. (2004). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS–K) third grade methodology report*. (NCES 2005-018). U.S. Department of Education. Washington, DC: National Center for Education Statistics.
- West, J., Denton, K., & Reaney, L. (2000). *The kindergarten year findings from the Early Childhood Longitudinal Study, kindergarten class of 1998-99*. Washington, DC: National Center for Education Statistics.