Use of the Multiple Lens Approach to Multiple Regression Findings with a National Dataset

Laura L. Nathans University of North Texas
Kim Nimon University of North Texas
David A. Walker Northern Illinois University

This study builds upon the work of Nathans, Oswald, and Nimon (2012) to demonstrate how to apply variable importance metrics to an analysis of a national dataset. The paper (a) reviews the variable importance metrics; (b) details the strengths and weaknesses of national datasets as well as issues the analyst must confront when using them; (c) presents an integrated write-up of how to interpret variable importance metrics in analysis of fall of kindergarten math achievement from the Early Childhood Longitudinal Study – Kindergarten Cohort dataset; and (d) discusses how variable importance metrics inform the extant substantive literature on kindergarten math achievement.

Multiple regression (MR) is a commonly employed statistical technique in social science research. Its purpose is to determine the "predictive power" of individual, as well as sets of independent variables (IVs), in explaining variance in a single dependent variable (DV) (Pedhazur, 1997). Researchers commonly rely upon beta weights when assessing the contributions of specific IVs to variance in the DV (Nimon, Gavrilova, & Roberts, 2010). As Nathans, Oswald, and Nimon (2012) and Zumbo (2012) pointed out, sole reliance on beta weights in interpreting MR findings provides limited, and potentially inaccurate, information regarding IVs' contributions to regression models. Shared variance between variables, a statistical phenomenon referred to as multicollinearity, may distort or minimize the contribution of an IV to an MR equation as reflected in beta weight values (Courville & Thompson, 2001; Nunnally & Bernstein, 1994). Accordingly, researchers need to rely on other measures of IVs' contributions to MR models to "yield the richest and most complete picture regarding the relationships between independent variables and the dependent variable" (Nathans et al., p. 2).

Nathans et al. (2012) highlighted several measures of variable importance in their guidebook. They noted that variable importance must be operationalized in terms of a specific measure of a variable's contribution to a regression equation, rather than employed as a "blanket term" (p. 2) that has general applicability when interpreting regression findings. As such, a brief outline of variable importance metrics utilized in said guidebook will be applied to the example in the current study.

Metrics for Variable Importance

Standardized Beta Weight. The first measure discussed is the standardized beta weight, which quantifies in standard deviation units, the expected change in the DV given a one standard deviation unit increase in the corresponding IV (Pedhazur, 1997). Beta weights are best used as a "beginning step" in the analysis of regression findings. However, because beta weights do not present a complete picture of how shared variance impacts a regression equation, integration of other variable importance measures in regression analyses is warranted.

Zero-Order Correlation. The second measure included in the guidebook is the zero-order correlation, which reflects the magnitude and direction of the bivariate relationship between the IV and the DV without controlling for other IV contributions to the regression equation. Squaring the zero-order correlation reflects the amount of variance shared between the IV and the DV.

Pratt Index. The third measure is Pratt's (1987) product measure, which partitions the regression effect by multiplying the zero-order correlation and beta weight for each IV. By partitioning the regression effect, the sum of all product measure values, unlike beta weight values, is equivalent to the amount of variance explained by the regression equation. The product measure is the most easily computed method of partitioning regression equation variance in the presence of correlated predictors, although not necessarily the most accurate, particularly when the value of the zero-order correlation or beta weight is negative.

Structure Coefficient. The fourth measure is variable importance, or the structure coefficient, which reflects the bivariate correlation between each IV and the predicted (y-hat) value resulting from the MR model (Courville & Thompson, 2001). We concur with Courville and Thompson's assertion that structure coefficients should always be computed in addition to beta weights when predictors are correlated, as they help identify where shared variance is distributed when beta weights are computed. If an IV has a large squared structure coefficient and a small beta weight, this discrepancy informs the researcher that variance shared between this IV and another IV was assigned to another IV's beta weight.

Commonality Analysis. The fifth measure presented is commonality analysis. Commonality analysis partitions the regression effect into variance that is (a) unique to each variable (i.e., unique effects) and (b) shared between all possible combinations of variables (i.e., common effects) (Rowell, 1996). When reporting commonality analysis results, the researcher can sum the common effects for each variable and compare them with the unique effect in order to determine whether a variable contributes more to a regression equation through variance it shares with other IVs versus uniquely contributes to the regression equation.

Dominance Analysis. The sixth measure employed is dominance analysis, which entails comparison of unique variance contributions of all pairs of IVs to regression equations involving all possible subsets of IVs (Azen & Budescu, 2003). An IV is referred to as "dominating" another IV if the former IV shows a greater unique variance contribution than the latter IV when it is entered last into regression equations containing all possible subsets of predictors. There are three hierarchical levels of dominance analysis described in this paper: (a) complete dominance - the strongest level of dominance that is demonstrated when one IV shows a higher unique variance contribution than another IV across all possible subsets of IVs; (b) conditional dominance - a weaker level of dominance that is shown when one IV contributes more unique variance than another IV on average across all subset sizes; and (c) the weakest level of dominance, termed "general dominance," which is represented when an IV contributes more unique variance to the regression effect than another IV on average across all IVs. It is important to note that dominance relationships can be established at weaker levels, if not higher levels, in the hierarchy of dominance relationships.

Relative Weight. The final measure integrated in this study is relative weights, which partition the regression effect through a computational method that creates IVs' uncorrelated "counterparts" (Johnson, 2000). Relative weights determine how much variance each IV contributes to the DV as a joint function of the relationships between (a) each IV and its uncorrelated counterpart and (b) the uncorrelated counterpart and the DV. Arguably, relative weights are the optimal method of partitioning regression equation variance, as they account for all explained variance while minimizing the impact of shared variance on the regression equation (Johnson & LeBreton, 2004).

Nathans et al. (2012) presented a guidebook of variable importance measures and applied them to a practical example. The current study extends upon this previous work by explicating how multiple measures of IVs' contributions to an MR equation/effect can be employed in analysis of a national dataset—the Early Childhood Longitudinal Study - Kindergarten Cohort (ECLS-K). National datasets are increasingly available to researchers through such national organizations as the National Center for Education Statistics (NCES) and the National Science Foundation (Hahs-Vaughan, 2005). Knowledge of how to apply analyses of multiple measures of IV contributions to a regression equation will facilitate more statistically "savvy" and methodologically accurate analyses of national datasets.

National Dataset Analyses

Strengths. Unfortunately, secondary data analysis, which includes analysis of national datasets, is an often untapped resource in the field of educational research. For example, Smith (2008) noted that less than half of papers published in British education journals employed analyses of secondary datasets. Because such datasets allow for reanalysis of existing datasets with new theoretical perspectives and/or more methodologically sophisticated analyses (Greenhoot & Dowsett, 2012; Smith), more frequent, in-depth analyses of secondary datasets (i.e., national) are warranted in the educational research literature. Researchers can confirm pre-existing findings through replication (Duncan, 1991), refine hypotheses (Castle, 2003), or conduct exploratory work that tests new ideas, theories, and research designs with national datasets (Smith). By eliminating the often logistically cumbersome step of data collection, researchers are freed to devote added "time and energy" to employing more thorough and advanced theoretical frameworks and analytic techniques that shed new light on old research questions (Greenhoot

& Dowsett; Smith) such as predictors of early mathematics achievement (i.e., the analysis conducted here). National dataset analysis conserves resources and is cost-effective, and the analyst does not have to rely on other researchers to collect data (Castle; Hofferth, 2005). Another advantage of such datasets is that they generally provide wide arrays of variables collected across large numbers of subjects (Greenhoot & Dowsett). As such, their samples are more representative and results have greater generalizability to broader populations, such as the U.S. as a whole, than studies that are primarily based on original data collection from circumscribed locations (Duncan; Greenhoot & Dowsett; Hofferth; Smith). Duncan pointed out that selection bias is largely eliminated in national datasets because of their broad, rather than purposive, sampling procedures. Additionally, data are typically collected with highly monitored, standardized surveys and assessments; thus, it is of the highest quality (Smith). Sophisticated statistical controls for multiple covariates can be employed with national datasets for infrequently sampled events such as poverty that would not be as likely to be reflected in smaller samples (Duncan). Overall, researchers can pose questions with national datasets that they frequently lack resources and subjects to otherwise explore (Greenhoot & Dowsett).

Weaknesses. Despite the many advantages of using national datasets, there are limitations and challenges inherent in the process. First, there is scant methodological literature addressing how to effectively analyze such datasets (Hahs-Vaughan, 2005), which the current study attempts to remedy. Second, because of the often large numbers of variables found in such datasets, the researcher must spend fairly extensive time familiarizing him/herself with the dataset prior to analysis (Castle, 2003; Greenhoot & Dowsett, 2012; Smith, 2008). Third, the data may have been collected with a different purpose or analytic approach from that used by the current data analyst (Castle). As Greenhoot and Dowsett pointed out, "the researcher has no control over who was sampled, what constructs were measured, or how they were measured" (p. 5) in secondary data analysis. The population and measures used in the dataset are typically based on others' decisions regarding data collection; as such, the researcher cannot potentially use the most validated or extensive measures for a particular analysis (Castle). For example, in the analyses outlined in this study, a measure such as the Child Behavior Checklist (Achenbach & Rescorla, 2001) would constitute a more in-depth measure of internalizing and externalizing behaviors than the Social Rating Scale used in the ECLS-K, which includes a limited four or five items reflecting these behavioral constructs.

There are several issues that must be addressed when analyzing national datasets. First, the dataset chosen for analysis must be congruent with the research questions posed by the analyst (Hofferth, 2005). Second, a guiding theoretical or conceptual framework must be relied on to effectively sift through the "mounds" of data generally found in such datasets (Castle, 2003). Additionally, a frequent issue that the researcher must confront when analyzing national or secondary datasets is that the population from whom the data was collected must be generalized to the broader population the dataset represents (Hahs-Vaughn, 2005). Hahs-Vaughan explicated in detail the process of weighting variables in order to generalize findings. She noted that national datasets generally obtain subjects through cluster, stratified, or multistage rather than simple random sampling. Accordingly, data must be weighted in order to address issues of selection probabilities, non-response and non-coverage, and post-stratification and effectively generalize findings beyond the population included in the original sample. It is important that researchers familiarize themselves with the literature that has involved prior analysis of the same dataset in order to effectively build on existing research (Hofferth). Lastly, such datasets often contain significant amounts of missing data. It is recommended, and used in this study, that researchers employ multiple imputation procedures to estimate missing data prior to analyses (Greenhoot & Dowsett, 2012).

Complex Sampling. Another significant issue to consider when analyzing national datasets is the design effect for the study. Cluster sampling methods are the most prevalent form of data collection for national datasets (Garson, 2012). Clusters are made up of natural groupings of individuals (Crespi, 2012) and have thus been described as, "a convenient aggregation of observations" (Davern & Strief, 2013, pp. 1-2). There are two types of cluster sampling: (a) a one-stage cluster sample in which clusters are selected and then individuals within clusters are sampled and (b) a multistage cluster sample in which certain locations, such as geographic areas, are selected and then clusters (e.g., households) within these areas are sampled. In the second stage of sampling, units (e.g., individuals within households) are sampled from which to collect data (Garson).

Because individuals or units within clusters are generally more similar (or homogeneous) than individuals selected with simple random sampling methods from a population, the statistical assumption of independence of observations is violated when data are collected with cluster sampling methods (Garson, 2012; Osborne, 2011; Wears, 2002). Observations within clusters are correlated (Zelin & Stubbs, 2005) for several reasons, including (a) the fact that individuals may select the cluster to which they belong; (b) cluster-level variables simultaneously affect all members within a cluster; (c) individuals within a cluster interact with and influence each other; and (d) individuals within a cluster share similar demographic characteristics such as socioeconomic status (Johnson & Elliot, 1998; Wears). Greater internal homogeneity of clusters is related to greater variability between clusters than would be found if the population was sampled randomly (Thomas & Heck, 2001). Increased between-cluster variance results in less precise and efficient statistical measurements of the population (Zelin & Stubbs) because of increased standard errors of the estimate for statistics measured with cluster sampling (Crespi, 2012; Hocking & Carlin, 1999). As measurements are less precise, they also show (a) wider confidence intervals (Wears; Zelin & Stubbs), (b) less reliability (Zelin & Stubbs), and (c) increased p values (Wears). The statistical power for the sample is proportionately reduced with increasing internal homogeneity within clusters (Wears).

Design Effect. Kish (1965) developed the design effect to quantify the proportionate reduction in effective sample size obtained with cluster sampling. The design effect has been described as a "correction factor" for the increased between-cluster heterogeneity that results in increased overall sample variance relative to what would be obtained with simple random sampling (Standardized Monitoring and Assessment of Relief and Transitions (SMART), 2012). Zelin and Stubbs (2005) explicate that the design effect is a ratio that is computed by dividing the variance from the actual sample by the variance that would be obtained if the sample had been collected using simple random sampling. It is both a ratio of the true to theoretical variances for a particular sample and the actual to effective sample sizes for the collected sample. It has also been referred to as the "variance inflation factor" because it quantifies the increase in sample size needed for a design to have the same statistical power as a sample obtained randomly (Wears, 2002). For example, if the value of a design effect is 2, the sample size would need to be twice as large as that obtained through simple random sampling to obtain adequate statistical power (SMART). Unfortunately, most statistical packages do not account for design effects. Their analyses are based on the assumptions that data were collected using simple random sampling methods (Herring & Liu, 1998) and; therefore, that observations are independent of each other (Osborne, 2011). Ignoring the design effect underestimates standard errors, thereby rendering statistical tests inappropriately sensitive and inflating Type I errors (rejecting the null hypothesis when it is true) (Osborne; Thomas & Heck, 2001). Most standard statistical packages, such as SPSS and SAS, do not account for design effects and thus fall prey to these statistical weaknesses (Herring & Liu).

The design effect (deff) can be computed with this formula: $\text{deff} = 1 + \rho(n - 1)$, where ρ represents the intracluster correlation and n quantifies the number of clusters (Wears, 2002). This formula demonstrates that design effects (a) increase with increased cluster size and (b) decrease with increasing values of the intracluster correlation (Davern & Strief, 2013). The intracluster correlation is a measure of both (a) the correlation between pairs of observations within a cluster (Crespi, 2012) and (b) the between-cluster variance (Crespi; Zelin & Stubbs, 2005). Due to incorporation of the intracluster correlation, the design effect largely reflects between-cluster variability (Zelin & Stubbs). If the intracluster correlation is high, the variance within the clusters is smaller than the variance of a sample obtained through simple random sampling, which increases the design effect (Wears). Intracluster correlations are higher when cluster membership relates to the outcome variable or members of the cluster influence each other, increasing internal homogeneity and; thus, external heterogeneity of clusters (Wears). In contrast, Zelin and Stubbs explain that the more internally heterogeneous and externally homogeneous the composition of clusters is, the less the between-cluster variability impacts the sample and the closer the design effect is to 1. As the number of clusters is multiplied by the intracluster correlation when computing the design effect, increasing the number of clusters augments the required sample size. Small differences in intracluster correlation values can result in large differences in sample sizes required for adequate statistical power (Crespi). Importantly, intracluster correlations can vary between clusters (Crespi).

Weighting. Use of sample weights also impacts design effects. Herring and Liu (1998) explicate that weights may increase or decrease the design effect depending on the correlation of the values of the weights with the standard deviations that are used in computations of descriptive statistics. Thus, the researcher should be aware of the design effect and the weighting process that is involved in analysis of a national dataset. Neglecting to factor design effects into analyses of national datasets results in errors of inference and inaccurate parameter estimates (Osborne, 2011).

Methods

Weighted Sample. The ECLS-K study involved data collection for 21,260 children throughout the United States. The purpose for data collection was to address questions regarding contributing factors to the developmental status of children in the United States at school entry and how it is affected by teacher, school, and parent characteristics (NCES, 1999). Child, teacher, and parent interviews were obtained. The first wave of data was collected in the fall of 1998 (i.e., the start of kindergarten year). Schools were sampled within geographic units and then students were sampled within schools. The ECLS-K data were weighted to adjust for differential selection probabilities at each sampling stage and for the effects of non-response to ECLS-K questionnaires (NCES). The data for the sample analyzed for this example write-up were weighted using a weight that accounted for concurrent collection of child, parent, and teacher data yielding a weighted sample size of 3,842,961 subjects. Prior to weighting, missing data were addressed through multiple imputation procedures in LISREL.

Variables. An MR model with four independent variables was employed for this study. The dependent variable in the MR model was the child's math score. In the fall of kindergarten, children were given a cognitive assessment that included both multiple choice and free response items. The items on the mathematics assessment measured mathematical abilities in the areas of conceptual knowledge, procedural knowledge, and problem solving (NCES, 1999). All four IVs were obtained from teacher and parent versions of the Social Rating Scale (SRS); a rating scale of kindergarteners' social development. From the teacher version of the SRS, two IVs were used in the MR equation: (a) the Externalizing Behavior Problems Scale, which reflects acting out behaviors such as arguments, fights, and impulsivity and (b) the Internalizing Behavior Problems Scale, which reflects the presence of anxiety, loneliness, low self-esteem, and sadness. From the parent version of the SRS, two IVs were also entered into the MR equation: (a) the Approaches to Learning Scale, which rates the child's ability to engage appropriately and consistently in classroom learning activities and (b) the Social Interaction Scale, which reflects pro-social behaviors as well as the child's ability to build relationships with peers and adults.

Results and Discussion

Table 1 presents the predictor metrics for the MR model employed in this study. While the coefficients that involved bivariate correlations, (r, r_s, r_s^2, CD:0 [i.e., r^2]) identified the relative importance of predictors as *Learning*, *Internal*, *Social*, and *External*, the remaining coefficients, including the beta weights, suggested a relative predictor importance order of *Learning*, *Internal*, *External*, and *Social*. These findings first underscored the importance of using multiple variable importance metrics as, when viewed in concert, some yielded different rank orderings of IVs in terms of their contributions to the regression equation. Even within the metrics that yielded identical rank orderings (i.e., all three variance partitioning statistics), there were differences in magnitudes of IV variance contributions to the regression equation. For example, it is important to note that relative weights and general dominance weights were identical in value although they are based on different variance partitioning methods. However, Pratt's (1987) product measure findings indicated that *Social* contributed almost half as much variance as *External* to the regression equation, while the other two variance partitioning methods suggested that it contributed nearly the same amount of variance as *Social* (i.e., .16 versus .13, respectively). Thus, results generally suggested that the researcher needed to conduct detailed comparisons across the relative importance metrics in order to more clearly understand the roles of shared variance and the variance assignment process when the regression equation was computed.

A first notable area of divergence between beta weight results and results from bivariate, correlational metrics was that beta weights suggested that *Internal* contributed substantially more to the regression equation than did *External* and *Social*. Additionally, the beta weight for *Externalizing* was almost double in magnitude when compared with the beta weight for *Social*. However, zero-order correlations for

Table 1. Predictor Metrics for ECL Study

Variable	β	*r*	r_s	r_s^2	Unique	Common	CD:0	CD:1	CD:2	CD:3	GDW	Pratt	RLW
Extern	-.098	-.154	-.500	.250	.009	.015	.024	.017	.013	.009	.016	.015	.016
Intern	-.136	-.185	-.600	.360	.017	.017	.034	.026	.020	.017	.024	.025	.024
Learn	.193	.241	.782	.611	.029	.029	.058	.046	.036	.029	.042	.047	.042
Social	.052	.160	.519	.269	.002	.023	.026	.015	.008	.002	.013	.008	.013
Total	NA	NA	NA	1.49	.057	.084	.142	.104	.077	.057	.095	.095	.095

Note.. Extern = Externalizing Behavior Problems Intern = Internalizing Behavior Problems Learn = Approaches to Learning Social = Social Interaction

External and *Social* were (a) nearly identical in magnitude (-.15 and .16, respectively) and (b) were nearly as large as the zero-order correlation for *Internal* (-.185). Additionally, the structure coefficient findings paralleled these findings. Structure coefficients for *External* and *Social* were (a) nearly identical in magnitude (-.50 and .52, respectively) and (b) reflected that both of these variables shared approximately a fourth of their variance, and only 10% less variance than *Internal*, with the predicted y-hat scores for *Math Achievement*. The discrepancy between the beta weights and structure coefficient results suggested that multicollinearity was operating between these two variables and the two most significant IVs (i.e., *Learning* and *Internal*) with some shared variance from *Externalizing* and *Social* being assigned to *Learning* and *Internal*.

The researcher needed to consult commonality analysis to determine the patterns of shared versus unique variance in this equation. The columns Unique and Common in Table 1 demonstrate that *External* contributed nearly twice as much shared as unique variance to the regression equation, and *Social* contributed almost solely shared variance to the regression equation. These findings highlighted the above supposition that the researcher should not “rule out” these two variables as insignificant in the regression equation; rather, these two variables played a role through the variance they shared with *Learning* and *Internal*.

The researchers then needed to explore Table 2 to confirm that these two variables shared variance with *Learning* and *Internal*. Table 2 illustrated that *External* shared 8% of the variance in the regression effect with *Internal* and 4.3% of the variance in the regression effect with *Learning* in second-order commonalities. *Social* shared 1.7% of the variance in the regression effect with *Internal* and 17.3% of the variance with *Learning* in second-order commonalities. Third-order commonalities indicated that *External* shared 1.9% of regression effect variance with that shared between *Internal* and *Learning* and Social shared 4.2% of the variance in the regression effect with that shared between *Internal* and *Learning*. Overall, commonality analysis findings supported the inclusion of all four variables in this equation due to the role of shared variance. However, it is important to note in Table 2 that the majority of the regression effect was explained by variance unique to *Internal* (18%), variance unique to *Learning* (31%), and variance that was common to *Learning* and *Internal* (17%); thus, supporting the larger role these two variables played in the regression equation.

The D_{ij} values for general dominance (see Table 3) reflect the order of general dominance weights presented in Table 1 and indicate that *Learning* generally dominates *Internal*, which generally dominates *External*, which generally dominates *Social*. The D_{ij} values for conditional dominance reflect the order of conditional dominance weights presented in Table 1 and indicate that *Learning* conditionally dominates *Internal*, which conditionally dominates *External*. Conditional dominance cannot be established between *External* and *Social* as the average incremental variance that the variables produce are not consistently higher or lower across regression models of different sizes (i.e., .026 > .024, .015 < .017, .008 < .013, .009 > .002). The D_{ij} values for complete dominance indicate that *Learning* completely dominates all remaining predictors and that *Internal* completely dominates *External*.

Complete dominance could not be established between *External* and *Social* because in the case of the predictor set containing only *Internal*, the incremental variance provided by *Social* (.018) was greater than the incremental variance provide by *External* (.016) (see Table 2). These findings highlighted the divergence between relative importance metrics in providing rank orderings of IVs.

Table 2. APS Related Metrics for ECL Study

Variable Set	Commonality	% Total	R^2	Extern.Inc	Intern.Inc	Learn.Inc	Social.Inc
Extern	.009	.093	.024	NA	.023	.05	.024
Intern	.017	.180	.034	.013	NA	.05	.018
Learn	.029	.309	.058	.016	.026	NA	.003
Social	.002	.023	.026	.022	.027	.036	NA
Extern,Intern	.008	.080	.047	NA	NA	.046	.019
Extern,Learn	.004	.043	.074	NA	.019	NA	.004
Intern,Learn	.001	.006	.084	.008	NA	NA	.002
Extern,Social	.000	-.004	.048	NA	.018	.03	NA
Intern,Social	.002	.017	.053	.013	NA	.033	NA
Learn,Social	.016	.173	.061	.016	.025	NA	NA
Extern,Intern,Learn	.002	.019	.093	NA	NA	NA	.002
Extern,Intern,Social	.000	.000	.066	NA	NA	.029	NA
Extern,Learn,Social	.000	.003	.078	NA	.017	NA	NA
Intern,Learn,Social	.004	.042	.086	.009	NA	NA	NA
Extern,Intern,Learn,Social	.001	.016	.095	NA	NA	NA	NA
Total	.095	1.000	NA	NA	NA	NA	NA

Note. Extern = Externalizing Behavior Problems Intern = Internalizing Behavior Problems Learn = Approaches to Learning Social = Social Interaction

Conclusions

Our study has illustrated the importance of analyzing MR results with large scale national datasets with the "multiple lens" approach. Thus, we hope this study serves as a model that will encourage researchers to analyze MR findings with national datasets in more depth, which will shed light on research questions that have generalizability intent to broader samples (i.e., such as the U.S. population in the ECLS-K). Researchers and practitioners will; thus, be better informed when generating recommendations for intervention and practice that may benefit large segments of populations tapped by national dataset samples.

Substantively, our findings support research that highlights the importance of learning-related social skills (McClelland, Morrison, & Holmes, 2000; Yen, Konold, & McDermott, 2004) and internalizing behaviors (Dobbs, Doctoroff, Fisher, & Arnold, 2006; Normandeau & Guay, 1998) in predicting early childhood math achievement. They add to the current body of literature on kindergarten math achievement by demonstrating that externalizing behaviors and social skills do not contribute unique variance to prediction of math achievement when accounting for learning-related social skills and internalizing behaviors. These findings contrast literature suggesting the unique importance of both externalizing behaviors (Pagani, Fitzpatrick, Archambault, & Janosz, 2010) and general social skills (Agostin & Bain, 1997; Downer & Pianta, 2006) to early childhood math achievement. Previous research has generally studied these variables in isolation rather

Table 3. D_{ij} Values for ECL Study

X_i	X_j	Complete	Conditional	General
Extern	Intern	0	0	0
Extern	Learn	0	0	0
Extern	Social	.5	.5	1
Intern	Learn	0	0	0
Intern	Social	.5	1	1
Learn	Social	1	1	1

Note. *Dij* = 1 denotes that X_i dominates X_j. *Dij* = 0 denotes that X_j dominates
X_i. *Dij* = .5 dominance cannot be established (cf. Azen & Buscescu, 2003). Extern = Externalizing Behavior Problems Intern = Internalizing Behavior Problems. Learn = Approaches to Learning. Social = Social Interaction

than in concert; thus, our study is the first to shed light through use of multiple MR "lenses" on patterns of unique and shared variance between these variables in predicting math achievement in kindergarten.

Without using the multiple lens approach to this national dataset, the practical importance of these findings would not have been discovered. Specifically, we would not have been able to determine how *External* and *Social* work in concert with *Learning* and *Internal* in impacting math achievement in the fall of kindergarten without comparison of beta weights with structure coefficients followed by commonality analysis. Because of its substantially smaller beta weight, we may have "ruled out" *Social* as playing a role in early kindergarten math achievement, despite its role in sharing variance with other variables in the equation. We also would not have understood that *External* does not necessarily play a larger role than *Social* in this regression by relying on beta weights alone, rather than dominance analysis and relative weights as well.

References

Achenbach, T.M., & Rescorla, L. A. (2001). *Manual for the ASEBA School-Age Forms andProfiles.* Burlington, VT: University of Vermont, Research Center for Children, Youth, and Families.

Agostin, T. M., & Bain, S. K. (1997). Predicting early school success with developmental and social skills screeners. *Psychology in the Schools*, *34*(3), 219-228.

Azen, R., & Budescu, D. V. (2003). The dominance analysis approach for comparing predictors in multiple regression. *Psychological Methods, 8(*2), 129-148.

Castle, J. E. (2003). Maximizing research opportunities: Secondary data analysis. *Journal of Neuroscience Nursing*, *35*(5), 287-290.

Courville, T., & Thompson, B. (2001). Use of structure coefficients in published multiple regression articles: ß is not enough. *Educational and Psychological Measurement, 61*, 229-248.

Crespi, K. (2012). Methodological challenges for the design and analysis of cluster randomized trials. Retrieved from http://cas.stat.ucla.edu/page_attachments/0000/0078/Crespi_nov13_2012.pdf.

Davern, M., & Strief, J. (2013). IPUMS user note: Issues concerning the calculation of standard errors (i.e., variance estimation) using IPUMS data products. Retrieved from https://international.ipums.org/international/resources/misc_docs/user_note_variance.pdf.

Dobbs, J., Doctoroff, G. L., Fisher, P. H., & Arnold, D. H. (2006). The association between preschool children's socio-emotional functioning and their mathematical skills. *Applied Developmental Psychology, 37*, 97-108.

Downer, J. T., & Pianta, R. C. (2006). Academic and cognitive functioning in first grade: Associations with earlier home and child care predictors and with concurrent home and classroom experiences. *School Psychology Review, 35*(1), 11-30.

Duncan, G. J. (1991). Made in heaven: Secondary data analysis and interdisciplinary collaborators. *Developmental Psychology, 27*(6), 949-951.

Garson, G. D. (2012). *Sampling*. Raleigh, NC: Statistical Associates Publishing.

Greenhoot, A. F., & Dowsett, C. J. (2012). Secondary data analysis: An important tool for addressing developmental questions. *Journal of Cognition and Development, 13*(1), 2-18.

Hahs-Vaughan, D. L. (2005). A primer for using and understanding weights with national datasets. *Journal of Experimental Education, 73*(3), 221-248.

Herring, S. G., & Liu, J. (1998). Complex sample design effects and inference for mental health survey data. *International Journal of Methods in Psychiatric Research, 7*(1), 56-65.

Hocking, J., & Carlin, J. B. (1999). Design of cross-sectional surveys using cluster sampling: An overview with Australian case-studies. *Australian and New Zealand Journal of Public Health, 23*(5), 546-51.

Hofferth, S. L. (2005). Secondary data analysis in family research. *Journal of Marriage and Family, 67*, 891-907.

Johnson, J. W. (2000). A heuristic method for estimating the relative weight of predictor variables in multiple regression. *Multivariate Behavioral Research, 35*(1), 1-19.

Johnson, D. R., & Elliott, L. A. (1998). Sampling design effects: Do they affect the analyses of data from the National Survey of Families and Households? *Journal of Marriage and the Family, 60*(4), 993-1001.

Johnson, J. W., & LeBreton, J. M. (2004). History and use of relative importance indices in organizational research. *Organizational Research Methods, 7*(3), 238-257.

Kish, L. (1965). *Survey sampling*. New York: Wiley.

McClelland, M. M., Morrison, F. J., & Holmes, D. L. (2000). Children at risk for early academic problems: The role of learning-related social skills. *Early Childhood Research Quarterly, 15*(3), 307 – 329.

Nathans, L. L., Oswald, F. L., & Nimon, K. (2012). Interpreting multiple linear regression: A guidebook of variable importance. *Practical Assessment, Research, and Evaluation, 17*(9), 1-19.

National Center for Education Statistics. (1999). *ECLS-K base year data files and electronic codebook.* Rockville, MD: Westat.

Nimon, K., Gavrilova, M., & Roberts, J. K. (2010). Regression results in human resource development research: Are we reporting enough? In C. Graham & K. Dirani (Eds.), *Proceedings of the Human Resource Development 2010 International Conference* (pp. 803-812), Knoxville, TN: AHRD.

Normandeau, S., & Guay, F. (1998). Preschool behavior and first-grade school achievement: The mediational role of cognitive self-control. *Journal of Educational Psychology, 90*(1), 111-121.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill, Inc.

Osborne, J. W. (2011). Best practices in using large, complex samples: The importance of using appropriate weights and design effect compensation. *Practical Assessment, Research, and Evaluation, 16*(12), 1-7.

Pagani, L.S., Fitzpatrick, C., Archambault, I., & Janosz, M. (2010). School readiness and later achievement: A French Canadian replication and later extension. *Developmental Psychology, 46*(5), 984-994.

Pedhazur, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction* (3rd ed.). Stamford, CT: Thompson Learning.

Pratt, J. W. (1987). Dividing the indivisible: Using simple symmetry to partition variance explained. In T. Pukkila & S. Puntanen (Eds.), *Proceedings of the second international Tampere conference in statistics* (pp. 245-260). Tampere, Finland: University of Tampere.

Rowell, R. K. (1996). Partitioning predicted variance into constituent parts: How to conduct regression commonality analysis. In B. Thompson (Ed.), *Advances in social science methodology* (Vol. 4, pp. 33-43). Greenwich, CT: JAI Press.

Standardized Monitoring and Assessment of Relief and Transitions. (2012). Sampling methods and sample size for the SMART methodology. Retrieved from http://www.smartmethodology.org/documents/sampling/Sampling_Paper_June_2012.pdf.

Smith, E. (2008). Pitfalls and promises: The use of secondary data analysis in educational research. *British Journal of Educational Studies, 56*(3), 323-329.

Thomas, S. L., & Heck, R. H. (2001). Analysis of large-scale secondary data in higher education research: Potential pitfalls associated with complex sampling designs. *Research in Higher Education, 42*(5), 517-540.

Wears, R. L. (2002). Advanced statistics: Statistical methods for analyzing cluster and cluster randomized trials. *Academic Emergency Medicine, 9*(4), 330-341.

Yen, C., Konold, T. R., & McDermott, P. A. (2004). Does learning behavior augment cognitive behavior as an indicator of academic achievement? *Journal of School Psychology, 42*, 157-169.

Zelin, A., & Stubbs, R. (2005). Cluster sampling: A false economy? *International Journal of Market Research, 47*(5), 503-524.

Zumbo, B. D. (2012, April). *Measuring relative importance of explanatory variables with applications to linear, logistic, and latent variable regression*. Paper presented at the Annual Meeting of the American Educational Research Association, Vancouver, B.C., Canada.

Send correspondence to: Laura L. Nathans
University of North Texas
Email: laura_nathans@yahoo.com
