The present study attempted to statistically test whether a nexus exists between student achievement and engagement levels within schools that elected to incorporate the Instructional Practices Inventory (IPI). A desirable feature of the Four-Level HLM methodology involves its potential to most aptly capture the realities associated with school improvement initiatives.

The findings from the study suggest that considerable growth in higher-order student engagement can result within schools that adopt the IPI process within 5-7 years of its implementation. The practices and processes which school leaders employ subsequent to IPI adoption also matter, as the findings suggest that higher-order thinking levels can be influenced by as much as 18 percentage points on the basis of IPI implementation fidelity levels.

Introduction

The passage of the No Child Left Behind (NCLB) Act has ushered in an era of unprecedented accountability standards for K-12 education (Kirsh, Braun, & Yamaota, 2007). While this landmark legislation follows decades of accountability reforms and calls for fundamental reconfigurations of the nation's public education systems, not since the Elementary and Secondary Education Act of 1964 has the nation seen such sweeping, ambitious federal legislation (Kirsh et al., 2007). In an era of pronounced accountability, school administrators and policymakers possess a voracious appetite for test score data and information on the factors that produce such test scores (Halverson, Grigg, Pritchett, & Thomas, 2007; Skretta, 2007). Standardized test performance and accountability initiatives are highly salient to policymakers and educators alike (Halverson, et al., 2007; Kirsch, 2007; Skretta, 2007). The current educational literature is surprisingly bereft of studies that examine the extent to which a relationship exists between the introduction of school improvement initiatives with schools and resulting classroom experiences and behaviors. Furthermore, the level of student engagement and higher-order thinking that occurs within schools and the standardized test performance of students within those schools is a relationship that has been left rarely tested to date.

Hierarchical Linear Modeling (HLM) has been proven as a trusted method in educational research to capture the complex environment context in which student learning occurs. Oftentimes, HLM can best address the complex and tiered data structuring that represents the reality of student engagement and learning within the wider educational environment. Willms (1999) underscores the promise of HLM, as he writes that “these methods allow one to systematically ask new questions about how policies and practices affect students’ outcomes. The advantage of HLM is that it allows the analyst to explicitly examine the effects on student outcomes of policy-relevant variables, such as class size or the implementation of a particular reform” (p. 475). Such an observation is a valuable contribution to the literature, and informs the methodology associated with the present research efforts.

The HLM methodology becomes amenable to domain specific analyses that can determine implications for the differential effects that the instructional treatment might exact in disparate geographic regions (Willms, 1999). Willms (1999) elaborates on the shortcomings of the many conventional approaches to studying instructional treatments, as he states that “many monitoring systems have emphasized simple comparisons among schools in their average levels of performance, without paying attention to the social context of the schools or to the factors that lead to improved performance. When analysts appreciate the possibilities afforded by HLM, they tend to shift their focus away from this good-schools/bad schools paradigm and design better monitoring programs with better data…" (p. 491). The four level HLM models employed in the present study are meant to represent a corrective measure to prior modeling efforts. Indeed a four-level approach to the study's research questions is offered as exploratory measure with a potentially stronger explanatory power. An exhaustive, holistic accounting of the wider environmental context which is so often neglected in the education research was a primary purpose which underlies the statistical models constructed in this paper.

Indeed, the current study takes into account Wech and Heck (2004)'s observation that "organizations are a multilevel, hierarchical phenomena," a qualitative observation that represented the very consideration for incorporating Hierarchical Linear Modeling in the present study (p. 1). Wech and Heck (2004) also reveal that "in large organizations, departments or facilities may be embedded in divisions that are, in turn, nested in organizations" (p. 1). Taking note of such complexities, Wech and Heck (2004) extend their endorsement of HLM as a statistical technique that offers "…many advantages in interpreting data within and across groups than other typically used techniques, such as ordinary least squares" (p. 7). The enthusiasm in the educational research community for HLM appears to be warranted on methodological grounds, therefore, and is actively incorporated into the present study to more completely advance the inquiry of school effectiveness and improvement initiatives.

Methods

Purpose of the Study

The purpose of this study is to ascertain whether the levels of higher-order thinking and active student engagement within schools, which are affected by such wider accountability mandates, were statistically related to the practices and processes encouraged by the IPI. These student engagement and IPI treatment processes were tested for evidence of a statistical relationship with the standardized test score performance of these schools, as well. While traditional demographic and input variables have been demonstrated to be highly correlated with, and related to, test score performance, few studies to date have empirically tested the relationship between standardized test performance rates and the level of student engagement within classrooms. This study ultimately requires more than a static consideration of the nature of student participation within classrooms. Operating from a conceptual framework of organizational learning, this study instead considered the extent to which the nature of pedagogical instruction evolves over time as instructional and curricular treatments are incorporated and employed within schools.

The Instructional Practices Inventory

The Instructional Practices Inventory (IPI) is a process employed by classroom observers to ascertain the nature of student engagement across classrooms within a school. The IPI is comprised of "a set of observational categories complex enough to provide substantive data grounded in the knowledge of best practice (valid) yet easily understood and interpreted" (Valentine, 2007). The IPI instrumentation allows a trained classroom observer to collect scores of observational codes that capture student engagement behaviors for each school. The observation categories included in the IPI observation protocol are: (1) student disengagement, (2) student engagement in non-higher order activity without teacher participation, (3) student engagement in non-higher order activity with teacher support, (4) teacher-directed instruction,

(5) student engagement in higher-order classroom discussion, and (6) all other higher-order student learning.

One of the more complex methodological challenges presented by the present study is not formulaic in nature, but rather involves the adequate and accurate definition of student engagement levels and what constitutes higher and lower ordering thinking. Such attempts to delineate meaningfully nuanced distinctions between various types of student engagement can quickly become obscured and fruitless if student engagement behaviors are hyper-parsed. The Instructional Practices Inventory (IPI) strikes a methodologically appropriate balance between meaningfully categorizing student engagement categories without deconstructing the classroom environment to such an extent that coding the minutia of student behavior becomes an untenable task for the classroom observer. More importantly, as the categories become more numerous (and indistinguishable), the reliability of such classroom observations can become greatly diminished. With this in mind, the IPI categorizes student engagement levels on a continuum from 1 to 6, which is designed to account for the spectrum of student engagement that one can expect to find in any given classroom at a particular moment.

Table One offers an explanation of each of the six coding categories. It is important to note that the higher-order categories ("5" and "6") represent desirable forms of student learning, whereas the lower-order categories ("1" and "2") represent less effective and generally indefensible forms of student activity within classrooms. It is not always possible, nor desirable, for students to be engaged solely in higher-order activities, however. As such, categories "3" and "4" account for those moments during classroom instructional time when the teacher is primarily involved in informing and directing the students' activities in the classroom, as student engagement becomes mostly passive and inactive. This might come in the form of teachers informing students of certain tasks or logistical considerations or teacher-directed learning, both of which are inevitable components of effective teacher pedagogy and student learning.

Insert Table 1 approx. here

The Instructional Practices Inventory Survey

The Instructional Practices Inventory (IPI) survey questionnaire will be the primary instrument by which to capture data about the nature of the implementation of the IPI process. The IPI survey was constructed in collaboration with the developer of the IPI process to more concretely measure several environmental factors demonstrated to directly affect student performance. More specifically, the IPI survey seeks to ascertain the perceived levels of school trust, collective teacher efficacy, teacher commitment, and the self-reported levels of importance that are placed on academic achievement (Valentine, 2005; 2007; 2008).

Survey responses suggestive of school environments in which school leaders are mindful and prospective in their mission, and desirous of improving the operational effectiveness and academic excellence of the school, served as a proxy of faculty trust. Question 14 of the IPI survey, for instance, while not restricted solely to such a consideration, captures elements of the processes and practices in the public school environment. The level of faculty receptivity can also be gleaned from several of the survey questions (Valentine, 2005; 2007; 2008). Finally, the extent to which the survey respondents were convinced that the IPI was effectively being employed within the schools and ultimately yielded material gains to both the quality of teacher pedagogy and subsequent student learning serves as a sound proxy for teacher collective efficacy (Question 14) (Valentine, 2007; 2008).

The survey questionnaire further captures the number of times and the duration that the IPI practices were executed within a given school (Question 5) (Valentine, 2007; 2008). The frequency and duration of IPI practices undertaken within schools serves as a proxy for the extent to which the schools value academic achievement (Hoy, Tarter, & Hoy, 2006). For the purposes of the present study, the extent to which the IPI is implemented with integrity is assumed to be a robustly telling proxy that manifests other features of the school's culture and climate (Valentine, 2007; 2008). More specifically, a good faith IPI implementation effort can be equated with a school's desire to ensure that a challenging climate of academic excellence exists at the building level (Hoy, Tarter, & Hoy, 2006). Furthermore, a consideration of the priorities of the respondents, as they enumerate their objectives on the IPI questionnaire (Questions 14), serves as a telling indicator of the extent to which classroom instruction and student achievement is valued by the school, and can be considered to be an apt proxy for the fidelity of IPI treatment implementation (Hoy, Tarter, & Hoy, 2006).

The IPI Survey: Constituent Parts Tested in the Present Study

The IPI survey questions employed in the present study are as follows:

IPI Question 3: This survey question was constructed to facilitate the researcher in better understanding the nature of data analysis within IPI-treated schools, including the mechanics and organization of faculty meetings.

IPI Question 5: Survey respondents are asked to indicate the number of times IPI observational data was analyzed by the faculty. The frequency and duration of IPI practices undertaken within schools serves as a proxy for the extent to which the schools value academic achievement (Hoy, Tarter, & Hoy, 2006). For the purposes of the present study, the extent to which the IPI is implemented with integrity is assumed to be a robustly telling proxy that manifests other features of the school's culture and climate (Valentine, 2007; 2008). More specifically, a good faith IPI implementation effort can be equated with a school's desire to ensure that a challenging climate of academic excellence exists at the building level (Hoy, Tarter, & Hoy, 2006).

IPI Question 6b IPI Survey Question 6b (Q6b_s), or the faculty's initial receptivity of IPI usage within their schools, represents a proxy for the likelihood of IPI implementation with sufficient fidelity. Additionally, such a construct may quantify school leaders accumulated facility with the IPI instructional treatment.   Accordingly, wider components surrounding the faculties' abilities to spearhead and accommodate change initiatives within their schools are accounted for, if only in part, but this survey construct.

IPI Question 10:   Survey Question 10 asked respondents to elaborate upon the physical configuration of the IPI faculty sessions. The researcher coded these responses on the basis of small groups or a whole group faculty meeting, and whether the underlying procedure of such meetings were structured in a manner that is amenable to facilitating the school improvement process.

IPI Question 12: This question was constructed with the purpose of enabling the researcher to quantify the outcome of the first IPI analysis, in terms of lessons learned by school leaders (more retrospective analysis), as well as future goals and plans established by the faculty (a more prospective vision).

IPI Question 14: The final question of the IPI survey was designed to capture the perceived benefit of the IPI to school improvement and effectiveness efforts. Question 14 represents a distinctive component of the mechanics that underlay IPI implementation. The question is not, therefore, a raw aggregation of the remaining thirteen constituent questions contained within the IPI survey. Additionally, Question 14 captures elements of the processes and practices in the public school environment. Furthermore, it might be the case that the IPI adoption could serve as a meaningful proxy of schools that are proactive in their efforts to anticipate future challenges by focusing on the current operational deficiencies within buildings.

The Interplay of Survey Questions

These survey responses provided the quantitative data needed to ascertain whether a statistical relationship existed between effective teaching and administrative practices, without assaulting the respondents' with a barrage of Likert-style questions. By questioning respondents as to whether teachers were able to work in teams (IPI Survey Question 10), and whether these teachers were empowered to determine the tone and direction of the meetings (Question 12), an empirical determination of efficacy levels can be constructed by the researcher. The frequency and duration of the undertaken IPI practices serves as a proxy for the extent to which the schools value academic achievement (Hoy, Tarter, & Hoy, 2006). For the purposes of the present study, the extent to which the IPI is implemented with integrity is assumed to be a robustly telling proxy that manifests other features of the school's culture and climate (Valentine, 2007; 2008). More specifically, a good faith IPI implementation effort can be equated with a school's desire to ensure that a challenging climate of academic excellence exists at the building level (Hoy, Tarter, & Hoy, 2006). Furthermore, a consideration of the priorities of the respondents, as they enumerate their objectives on the IPI questionnaire (Questions 3, 12, and 14), serves as a telling indicator of the extent to which classroom instruction and student achievement is valued by the school. As such, an apt proxy for the fidelity of IPI treatment implementation is generated (Hoy, Tarter, & Hoy, 2006). The importance of the academic optimism factor (captured by Questions 6b and 11) must not be understated, as such academic emphasis can explain mathematics and reading achievement scores despite markedly differing SES levels of a school population (Hoy, Tarter, & Hoy, 2006).

A Brief Explanation of the IPI Practices and Processes Fidelity Levels

As the practices and processes associated with the IPI have been dissected and elaborated upon, it is important for the reader to appreciate that the IPI treatment fidelity is defined as the measurable level of such key practices and processes. The IPI survey responses crafted to coax specific building level information regarding the IPI implementation can be coded. The resulting raw, numerical gains in such coded survey responses can then be subjected to HLM model testing, just like any other quantifiable input data.

Statistical Model Configurations

For the purposes of the present study, the classroom will be used as a starting point to accumulate the data needed to address the extent to which student engagement levels are altered as a result of IPI implementation. The study also explored the student engagement and standardized test performance relationship. To adequately account for the nesting of students and their classroom behaviors in a greater environmental context, the introduction of the district and regional levels (levels three and four, respectively) can additionally be considered. This enables the researcher to better account for the structure inherent in student learning. Furthermore, knowledge of the most elaborately constructed HLM models allows for the immediate evaluation of the proportion of variance explained among each of these levels. From such data, the researcher can then determine whether a parsimonious pruning of entire an entire level (or levels) of the HLM Models is warranted.

<u>Level-One Classrooms</u>: Level-One of the Four-level HLM models represents the student engagement levels within schools' classrooms. Raw percentage breakdowns are computed for each of the 241 schools that provided IPI classroom data profiles, in the form of singular lower order codes for core and total classrooms ("C123" and "C2"), higher order and an aggregated metric of distinctive higher order categories "C5" and "C6" ("C56"- in core content classrooms), and "T5" and "T6" ("T56"- across all classrooms). Multiple classrooms observations are coded for each classroom within the building over the course of a school day. As a result, a statistically representative depiction of student engagement levels within the school can be introduced into the multilevel statistical study at Level One. Student engagement levels were assigned as dependent variables in the model. The outcome variables were then tested against the corresponding IPI practices and processes, as captured by coded IPI survey responses.

Level Two-Building Processes: While classroom observational data remains a reliable, quantifiable metric upon which to gauge the nature of student behavior within classrooms, the wider practices and processes associated with the IPI treatment represent a phenomenon within which classrooms are considered to be nested. As such, level two of the Survey incorporates survey response predictors from 144 schools, and in some instances, additional student engagement data, as well.

Level-Three School Districts: School districts comprise the third level of the multilevel statistical study. The schools that provided IPI classroom data were located within 105 distinctive school districts across the state. While not categorically the case, there are sound theoretical reasons to believe that school districts will exhibit a pronounced and inescapable influence on the health and effectiveness of the schools that operate within them. The demographic data provided by Missouri's Department of Elementary and Secondary Education is quite exhaustive. For the purposes of this study, traditional socioeconomic, and controllable and uncontrollable educational resources and input factors, were collected and recorded for the corresponding school districts within which the schools were situated.

Level-Four Regional Professional Development Centers: Not unlike many states across the nation, Missouri is comprised of several disparate regions. Impoverished urban centers within Kansas City and St. Louis are surrounded by more affluent and effective suburban districts. As these districts were nicely dispersed across the region, the resulting averages amount to a statistically representative average of regional demographic, controllable and uncontrollable inputs, and student achievement, as opposed to more hollow arithmetic averages of only a few concentrated districts within an RPDC.

To determine whether a nexus exists between student engagement levels and the practices and processes associated with the IPI, the data collected from IPI walkthrough observations were subjected to Hierarchical Linear Modeling (HLM) analysis employing HLM 6.4 software. Such student engagement levels represent a variable that the schools are able to directly influence (Lee & Weimer, 2006). These statistical models will account for across classroom (Level-1), within school environment (Level-2), district-level (Level-3), and regional-level (Level-4) data nesting effects.

The student engagement outcome of greatest interest to the researchers, the percentage of core content area classrooms that were coded a "5" or a "6" ("AV_C56"), was assigned as the dependent variable in both the two and three level models. It is further theorized that non higher-order student engagement levels ("C1," "C2", "C3") should not represent an identically influential relationship or a proportionately inverse relationship to higher-order engagement levels. Total higher order thinking data ("AV_T56") was also assigned as the dependent variable to be tested by HLM modeling.

To test this theoretical proposition, the percentage of classrooms coded as either a "1", "2", or a "3" (student disengagement/teacher inattentiveness or passive seatwork "AV_C123", within core classrooms) was assigned to be the dependent variable in HLM Model. An isolated lower-order thinking construct, teacher disengagement with non-higher order student engagement in core classrooms, ("AV_C2") was also designated as a dependent variable to determine the effects of this undesirable pedagogical practice. Ultimately, the student engagement and achievement relationship can be more exhaustively explored by testing data under the HLM statistical framework. Simply put, the HLM models enable the researcher to determine the extent to which the IPI more directly influenced student engagement levels. Such fluctuations might, in turn, also exhibit influence on standardize test score levels of schools.

Explanation of Population Sample and Descriptive Data

In 2005, numerous schools across Missouri and the nation began to conduct IPI classroom observations. These data collections enable the level of student engagement in each classroom within a school to be recorded by a trained observer. It is now estimated that between 300-400 schools within the state of Missouri utilize the Instructional Practices Inventory with some degree of fidelity. The eventual population size of 152 schools offers evidence of a robust response rate to the electronic survey.

Data Collection Procedure

Two points are to be stressed as it relates to the trained IPI classroom observers who conduct classroom walkthroughs: First, teacher and school leaders other than principals are designated as data collectors to diminish the possibility of bias or punitive evaluation associated with the data collection. Second, as all IPI data collectors are trained in a workshop after passing an IPI reliability test, the inter-rater reliability of such coding must be at least .90. These observations provide a comprehensive, empirical representation of the nature and level of student engagement that was evidenced within the population sample of schools in the study. For the purposes of this study, these coded student engagement percentages were analyzed and incorporated as measurable independent variable metrics, introduced in the form of predictors in the HLM models.

Findings

Descriptive Statistics Associated With the IPI

Table Seven provides a summary of the descriptive statistics associated with the IPI-treated schools that served as the population sample of the four-level Hierarchical Linear Modeling (HLM) that captures every school which provided data for the current study. As is evidenced from the HLM descriptive output, 241 schools contributed data that was treated to be nested within the 105-district level variables, which were in turn nested within 9 RPDCS. Several variables included in the first level are particularly noteworthy and warrant further explanation. The average percentage of students within a school's population who receive free and reduced lunch (FRL) of the 241 schools included in the IPI study was found to be 46.79%, a value slightly above the state's average 41.8%. The percentage of minority students within the 243 school sample set was 16.72, less than the statewide average of 23.6%. The average percentage of students within these schools who demonstrated proficiency on the 2007 Mathematics and 2007 Communications Arts segments of Missouri MAP tests were 42.18% and 42.57%, respectively. The state average mathematics passage rate was 35.8%, however, while 42.9% of Missouri students demonstrated proficiency on the Communications Arts section of the MAP test. At the district level, 64.33% of students within the 105 district sample populations went on to pursue some form of higher education. Furthermore, 9.64 percent of families within these districts were impoverished. Not surprisingly, and nearly in complete accord with federal mandates, 97.26% of teachers within these districts were certificated. Finally, 77.6% of families had remained in the same county within the last five years. A good many of the RDPC variables are consistent with the national averages for the corresponding categories. Two variables appear to deviate somewhat from such averages, however as the average teacher salary within the nine RPDC's was computed to be $39,330.87, while the average per pupil expenditure was $8,147.27.

It is also worthwhile to compare the student engagement levels within the population of the present study with both the successful and ineffective schools considered in 2004-2005, as presented in Table Two.

Insert Table 2 approx. here

The findings from the four-level HLM Models that tested higher-order student engagement as outcome variable were telling. Table Three below displays the findings for these models, in which the average percentage of higher order thinking within core classrooms "C56" was assigned as the dependent variable. Interestingly, the percentage of students who receive free and reduced lunch with a school's population ("FRL") was found to be statistically significant at level one in the model that solely employed FRL predictors, and evidenced a rather unsubstantial slope. The fixed effects coefficients of the survey questions were considerably more elevated, however. Indeed Survey Questions 5, (the number of data analyses) 6b (the initial receptivity to IPI process), and 14 (the perceived benefits of IPI to the school setting) were all found to materially and considerably impact higher-order student engagement levels across core classrooms.

Perhaps the most considerable findings associated with free-and-reduced lunch rate ("FRL") of the student populations of IPI-treated schools occurred at the district and regional levels (levels three and four, respectively).   Indeed, the FRL rate at the district level was found to be twice as impactful on student engagement as it was at the school level, while the regional level FRL rate was found to be almost three times as great at the district level and fully five times as great as at the school level.

The variance explained by the higher-order thinking across core classrooms, as incorporated in these HLM models, varied quite markedly. Still, only 0-2% of level one variance was explained by the assigned predictors. The second level of the models appeared to be more adequate in accounting for model variance, explaining between 0-56.2% of the variance. The district level predictors (Level three) accounted for between 0-72.7% of the variance, while four of the five models' level four predictors accounted for nearly all the variance associated with the models at the regional level.

Insert Table 3 approx. here

As expected, the findings for the total higher order thinking averages across all classroom HLM Models ("AV_T56") were found to be strikingly similar to those for core classrooms only (the "AV_C56" HLM models). Again it was the case that the magnified of the FRL fixed effects coefficient was found to muted, while the coefficients associated with the survey question predictors were moderately elevated. The FRL rate at the district level in this model was found to be statistically insignificant, however. The FRL fixed effects coefficients at the regional level nearly mirrored their counterparts in the "AV_C56" HLM models, however, evidencing magnitudes of four to five times greater than were found at the school level. The HLM Models that assigned "AV_T56" as the dependent variable also appeared to explain slightly more variance at levels one and two, while explaining considerably less variance at levels three and four. Nevertheless, only one "AV_T56" model was unable to account for any of the computed variance, as this occurred at a single level (Level four), and in only one instance. Consequently, the explanatory consistency of these models is more noteworthy than was the case with the "AV_C56" HLM models. These results, along with other important output data, are provided in Table Four below.

Insert Table 4 approx. here

Also of interest was the extent to which lower-order thinking levels were statistically related to IPI practices and processes. Furthermore, an empirical consideration of the lower-order fixed effects coefficients compared with their higher-order thinking counterparts provided valuable findings. These results are provided in Table Five below. The FRL fixed effects coefficients in the lower-order thinking models were quite similar to the higher-order thinking HLM models, while the regional level (level four) fixed effects coefficient magnitudes were slightly less robust than was the case in the higher order thinking models. Of greater interest is a finding at level three of the HLM models: at this district level, all three models evidenced positive slopes twice as great than at the school level, suggesting that district-level poverty might enhance lower order thinking across school districts. The negative fixed effects coefficients at the school and regional levels, however, evidence not only a counteracting effect, but also suggest that the FRL rate of a student population is not highly impactful in dictating the extent of lower order thinking across IPI-treated schools.

Insert Table 5 approx. here

The study's standardized achievement models were comprised of distinctly-configured HLM models that designated the 2007 Communication Arts and Mathematics Missouri MAP test passage rates as the dependent variables. The output associated with these models is provided in Table Six below. Worth mentioning is the statistically significant FRL fixed effect coefficients, whose magnitudes were fully four times greater than in student engagement models. Additionally, the FRL predicators at level two were also found to be statistically significant, although the magnitudes were slighter than those at level one of the models. The district-level FRL fixed effects coefficients, while greater in some instances, widely mirrored those at level two. At level four, the magnitudes of the fixed-effects coefficients were slightly greater than their higher and lower-order thinking counterparts.

The student engagement fixed effects coefficients, while moderately elevated, were all found to be in the expected direction for both the core ("C56") and total classroom higher and lower order dependent variable ("T56"). It was also determined that assigning both IPI survey question predictors and student engagement predictors simultaneously within a single level of the HLM model does not impact the magnitude of the student engagement fixed effects coefficient.

The variance explained at level one of the achievement HLM models was insignificant. More specifically, between 2-7% of the level one variance was explained by the assigned predictors of the achievement models.   An exception was found in the model that did not incorporate a survey predictor, but instead assigned the “AV_C56” predictor at level one, which explained 85% of level one variance. The variance at the second level was not accounted for in three of the six HLM models, while the remaining three models accounted for between 11.5-59.5% of the variance at this level. Both levels three and four of the HLM Achievement Models were highly successful in explaining the models’ variance, with 62.6-98.1% of level three variance explained by the assigned predictors. Additionally, five of the six HLM Achievement models explained all or nearly all of the region-level variance.

Insert Table 6 approx. here

Discussion

This study tested the statistical relationship of schools’ student engagement levels with various IPI processes. In addition, the student achievement-student engagement relationship was tested at length in this paper. Four-level Hierarchical Linear Modeling (HLM) was employed as the statistical vehicle by which to accomplish this empirical undertaking. The Four-level HLM methodology represented an attractive means by which to account for the nested structure of student engagement and learning within schools’ cultural environments. Ultimately, these relations are themselves embedded within the respective districts and regions where these classrooms are situated. Given the robust accounting for variance at the third and fourth levels of the HLM models, these models do not appear to contain extraneous constructs. While model parsimony is desirable in Hierarchical Linear modeling, so too is a sufficient accounting of the practices and processes that occur not only within the classroom and but across a wider operational context. This study provides compelling findings on several counts. Namely, this more exacting methodological configuration yielded fixed effects coefficients that were generally consistent in magnitude and direction with similarly-constructed two and three level models, while accounting for variance at four distinctive levels.

While the percentage of students who receive free and reduced lunch (FRL) and the percentage of minority students within a school population's relationship to standardized test achievement has been convincingly demonstrated in decades-long research efforts, the relationship between FRL/percentage of minority students and student engagement levels has been demonstrated to be less consequential (Applebee et al., 2003). This study provided additional empirical evidence to corroborate Applebee et al.'s (2003) contention. More specifically, the fixed effects coefficients for FRL were found to be between two and four times greater in those models where student achievement was assigned as the dependent variable than when student engagement was tested as the dependent variable. Such a consideration is especially important in the context of the IPI treatment adoption, which not only represents a negligible cost to schools, but is also predicated upon revamping, or at the very least fine-tuning, collaborative efforts and pedagogical practices. That is, school reformers must act in a matter that will stimulate enhanced levels of higher-order student engagement within schools. The empirical findings of this study support the proposition that the alteration of building level practices and the extent to which students approach content material as they undertake their studies within classroom learning environments do not hinge upon socioeconomic considerations to the same extent as with test performance (Applebee et al., 2003; Waxman, 1997).

The socioeconomic status-achievement nexus was not challenged in the present study; however, the student engagement-student achievement linkage was empirically investigated by incorporating SES metrics where the demographic construct is theorized to exhibit influence within the context of the districts and regions within which the IPI treatment is actually introduced. That the FRL magnitude was most elevated at the district and regional levels deserves some mention. It might seem facially intuitive to presume that the building-level FRL rate would exhibit the greatest influence on student and achievement levels. As the IPI is introduced to school leaders, the quality of the plans for its adoption at the building level occur at the district and regional levels, and must be duly accounted for in the study's methodology. Ultimately, the district and regional-level resources that most aptly capture the capabilities of school leaders as they both guide and exact oversight over IPI implementation may also explain the elevated input magnitudes discovered at levels three and four of the HLM models.

The findings associated with the FRL predictors across the four level HLM models must not overshadow the importance of the findings associated with the dependent variables. Student engagement levels were demonstrated to be statistically related to the various practices and processes associated with the IPI. The growth in higher-order thinking that results from the good faith incorporation of the IPI tenets appeared to be closely related to that of the lower-order thinking models (which exhibited opposite directionality with one another). Furthermore, student engagement levels appear to be statistically related to communication arts and mathematics standardized performance levels. While not unexpected, this finding might serve as a stimulus to promote more active higher-order learning in the public school setting.

The student engagement outcome models yielded several significant findings associated with the IPI practices and processes on the resulting student engagement profile of schools included in the study. Most notable among the findings include the following:

- Were only data analysis be conducted as part of the schools' incorporation of the IPI, optimal levels of higher order student engagement within core classrooms (60% of all coded behavior) would be attainable in 5.6 school years were such data collected quarterly.

- The receptivity with which the IPI is received by faculty can influence higher-order student engagement levels by as much as 12.84 points, while most schools would encounter a 6.42 percentage point growth in higher-order thinking.

- Faculty and administrator perceptions of the current and prospective benefit of the IPI to their schools is a factor closely related to higher-order student engagement levels. Indeed, gains as great as 18.6 percentage points appear to associated with such faculty assessments of the IPI as an optimally effective instructional initiative within their educational settings. Most schools are likely to encounter 9.3 percentage point growth in higher-order thinking were the IPI to be viewed as an optimally effective school improvement initiative within the school.

- Conservative estimates from the multiple IPI survey responses included in the study reveal that school leaders who collect and analyze data, and who then deem the process to be a highly effective instructional initiative, may be able to attain considerably more robust levels of higher-order student engagement (60% of coded behavior) within an average of 6.7 years.

- The practices and processes associated with the IPI adoption may have the potential to influence higher-order student engagement levels across all classrooms to an impressive extent. More specifically, the extent to which the IPI implementation process is discussed and endorsed by a faculty and appears to dictate student engagement levels within schools. The faculty's general perceptions of the IPI process suggest a congruent and ameliorative effect on the school's current pedagogical and engagement behaviors are related to higher-order student engagement enhancements ranging from 10.32 percentage points (based on enthusiasm) to 14.04 percentage points (based on data collection). Additionally, school leaders who most effectively address the underlying rationale and best practices associated with the IPI process before its implementation would enjoy average higher-order thinking gains of 11.64 percentage points.

- Were school leaders to conduct quarterly IPI observations, student non-higher order engagement with teacher disengagement in core classrooms could be entirely eliminated in 4.16 school years. Research suggests that faculty enthusiasm, when coupled with engagement data collection, could eliminate student non-higher order engagement with teacher disengagement in core classrooms "C2" in one year.

- More modest findings were evidenced for all student behavior that is coded as non-higher order in nature. Diminishing non-higher order student engagement to a more desirable level of 20% of all coded classrooms engagement behavior would increase mathematics proficiency rates by 3.51 percentage points. In core classrooms, an identical decrease results in a Mathematics proficiency rate increase of 4.13 percentage points.

The findings associated with standardized achievement level fluctuations based upon student engagement levels within schools were also significant and compelling. School leaders will be encouraged to learn that:

- Communication Arts proficiency rates would increase by 5.48 percentage points were higher-order engagement levels to increase from their current levels to a more optimal 60% of all coded classroom behavior.

- Mathematics proficiency rates increase to an even greater extent were higher-order student engagement levels across all classrooms to increase to 60% of all observed classroom behavior. Were schools in the study to display such growth in their higher-order student engagement levels, mathematics proficiency rates would increase by 7.97 percentage points.

References

Applebee, A. N., Langer, J. A., Nystrand, M., & Gamoran, A. (2003). Discussion-based approaches to developing understanding: Classroom instruction and student performance in middle and high school English. *American Educational Research Journal, 40*(3), 685-730.

Halverson, R., Grigg, J., Pritchett, R., & Thomas, C. (2007). The new instructional leadership: Creating data-driven instructional systems in school. *Journal of School Leadership, 17*(2), 159-194.

Hoy, W. K., Tarter, C. J., & Hoy, A. W. (2006). Academic optimism of schools: A force for student achievement. *American Educational Research Journal, 43*(3), 425-446.

Kirsch, I., Braun, H., & Yamamoto, K. (2007). *America's perfect storm: Three forces changing our nation's future*. Princeton, NJ: Educational Testing Service.

Lee, K., & Weimer, D. (2002). *Building value-added assessment into Michigan's accountability system: Lessons from other states*. Lansing, MI: The Education Policy Center at Michigan State University.

Skretta, J. (2007). Using walk-throughs to gather data for school improvement. *Principal Leadership (Middle School Education), 7*(9), 16-23.

Valentine, J. W. (2005). *Instructional practices inventory: Profiling student engagement for school improvement.* Columbia, MO: Middle Level Leadership Center, University of Missouri (www.MLLC.org).

Valentine, J. W. (2007). *The instructional practices inventory: Using a student learning assessment to foster organizational learning*. Columbia, MO: University of Missouri (www.MLLC.org).

Valentine, J.W. (2008) Middle Level Leadership Center (MLLC) Website material published by Director Jerry W. Valentine. University of Missouri.

Waxman, H. C., Huang, S. Y., Abderson, L., & Weinstein, T. (1997). Classroom process differences in inner-city elementary schools. *The Journal of Educational Research, 91*, 49-59.

Wech BA, Heck AL. An introduction to hierarchical linear modeling for marketing researchers. *Marketing Bulletin*. 2004; 15, Technical Note (1): 1-8.

Willms, J.D. (1999). Basic concepts in hierarchical linear modeling with applications for policy analysis, in Gregory J. Cizek (Ed.). *The Handbook of Educational Policy*. San Diego, CA: Academic Press.