Student Engagement and Achievement on High-Stakes Tests: A HLM Analysis across 68 Middle Schools

Jerry W. Valentine
Professor Emeritus, University of Missouri

Justin A. Collins
Post-Doctoral Education--Law Student
University of Missouri

Abstract

This study analyzed the relationship between the nature and levels of student engagement in Missouri middle schools with the standardized test performance levels of the schools. More than 10,000 classroom observations were collected over a multi-year period of time and sorted according to the six categories of the Instructional Practices Inventory Process for profiling student engagement. Of particular interest in this study were the relationships between student disengagement, student engagement in lower-order/surface learning, and engagement in higher-order/deeper learning experiences, with student achievement as measured by the Missouri high-stakes assessment as the dependent variable. Engagement data, achievement data and demographic data from the schools were analyzed with demographic data from their districts and the nine regional professional development centers of the state to establish two-level and three-level analyses using Hierarchical Linear Modeling.

Statistical relationships between the levels of student engagement and standardized achievement were evident. While higher-order student engagement enhancements lead to marginal increases in standardized achievement, disengagement levels detrimentally impact student learning at considerably more pronounced rates. Indeed, the findings reveal that the student disengagement levels can deleteriously impact student achievement levels at a magnitude three times as great as the benefit of higher-order student engagement on student achievement levels.

The general public and school leaders have become fixated on test score data with the quality of schools and their leadership teams largely evaluated based upon their schools' standardized test score performance (Caldas & Bankston, 1999). Concerned community members and state accountability mandates demand that public schools demonstrate impressive growth in test performance in successive years. Not surprisingly, therefore, the metric upon which the work of school leaders is gauged is largely predicated upon their ability to improve student achievement scores (Halverson et al., 2007). Such heightened expectations offer yet another justification for classroom walkthroughs that facilitate school leaders in assessing their progress toward providing instructional environments that ensure healthy test performance levels. It should also be noted that the observations and data gleaned from classroom walkthroughs can enable more than simple administrative oversight. Indeed, the information gleaned from walkthroughs "can be used to generate conversations with teachers regarding student learning and their use of best educational practices" (Skretta, 2007).

In today's educational environment, classroom walkthroughs have become a common phenomenon in the school reform process, but they are not a new educational concept (Kachur, Stout, & Edwards, 2010). Kachur and his colleagues describe walkthroughs as providing "snapshots of instructional decisions and student learning, that, over time, create an album" (p. 1) of the school's practices. Walkthroughs are not infallible cure-alls for school improvement. Highly disconcerting to many is the validity with which walk-through processes measure instructional practices and student engagement. Just as disconcerting is the extent to which teachers may mistake student activity within the classroom for actual academic progress (Skretta, 2007). Classroom observations can provide teachers with accurate and relevant data on both the quality of their instruction (Skretta, 2007) and of the students' classroom contributions and learning (Valentine, 2005, 2009; Kachur, et al., 2010). As such, classroom observations, and the assessment of student learning, can allow school leaders to better ascertain the quality of student learning within a school. Halverson et al. (2007) noted that collecting data solely to possess such information will be an exercise in futility, however, because "a central assumption across our schools was that data fuel the improvement process: the data acquired must have the potential to inform teaching and learning, but it need not be limited to test score results" (Halverson et al., 2007, p.167). Indeed, appropriately collected and synthesized data may assume many forms, but will ultimately provide school leaders with a comprehensive picture of the quality of classroom instruction that students receive (Halverson et al., 2007).

Assessing the nature and vigor of student learning in schools is an important component of improving school performance. Indeed, school reform initiatives that seek to target and ameliorate certain operational deficiencies must first identify such areas that warrant improvement. Data collected from classroom-walkthroughs that capture student engagement and the teachers' pedagogical practices that foster such student participation can prove to be a valuable data source (Valentine, 2005; 2009). A properly designed and performed walk-through process can entail identifying both the nature and extent of student engagement within classrooms, as well as collecting information that can later be used to facilitate productive faculty discussion on how to best improve the instructional environment within the school (Valentine, 2005; 2009). Skretta (2007) speaks to the need and importance of classroom walkthroughs, noting that "continuous examination of data…is important for sustaining the professional learning community…Data should be collected throughout the year on a regular basis" (p. 16). Teacher collaboration can be facilitated by walk-through data collection. Further, faculty reflection on this data can structure their focus on those building-level practices related to curriculum, instruction, and professional development. Such efforts may lead to higher test score levels (Goddard, Goddard, & Tschannen-Moran, 2007). In fact, teacher collaboration appears to be a significant predictor of achievement variability between schools (Goddard et al., 2007).

The importance of data reflection within schools has been well documented, and teacher-led study sessions can serve to inform schools of their current instructional practices (Valentine, 2005; 2009). Faculty sessions can be conducted in a fashion that does not intimidate or overwhelm participants with reams of data, but rather presents a clear and telling representation of the nature of instruction that is provided across the school. Such information can be empowering to teachers, as it not only informs best practices, but also spurs enhanced collaboration with colleagues and administrators (2005; 2009). Valentine's contention is consistent with the assertion that "the value of local data reflection sessions seems to lie not in the sophistication of the statistical analysis but in the frank discussions of practice" (Halverson et al, 2007, p. 172). As school leaders incorporate such data into their faculty discussions, they can form decisions about how to best proceed in improving teacher instruction, and subsequently, student learning and engagement. Such data can also be used to better inform a school's attempts at program alignment, in which school activities are tailored to function in accord with the instructional goals of the school (Halverson et al., 2007). The extent to which data are successfully incorporated into practice may vary depending upon the level of fragmentation or misdirection of such programs, which of course, varies considerably from school to school.

The data collected in walkthroughs can provide more than a singular assessment metric of the quality of a school's instructional provision. These data can also be used by schools to facilitate community outreach efforts (Halverson et al., 2007). Enhanced and continuous data collection can also assist school leaders as they craft blueprints that chart the direction of a school's instructional provision over time, after the changes are introduced and begin to take effect.

Teachers and Teaching

While school leaders are responsible for undertaking the school improvement initiatives required to provide the most suitable academic environment for students, it is the teachers who are ultimately tasked with executing such initiatives within their classrooms. Much like determining a school's contribution to student achievement can be exceedingly useful, so too can studying teacher effects on student learning (Druian & Butler, 1987). Student learning, after all, is largely impacted by the quality of teacher instruction.

Quality teaching appears to be largely related to teacher efficacy. Vanosdall et al. (2007) found, for instance, that "when teachers lacked the confidence and experience in using investigative modeling, they tended to resort to more familiar 'teacher-directed' instruction over the intent of the materials" (p.5). Consequently, offering any prescriptions for quality classroom instruction first requires an understanding of what impacts these teacher efficacy levels. More specifically, Thompson (2002) found that lowered teacher expectations and reduced levels of teacher efficacy were common among those teachers who possessed the following characteristics: 1) a belief that differences in student performance based on race were deficits that were not the teachers responsibility to bridge, 2) a belief that diversity of families and cultures was viewed as a deficit rather than a strength, 3) lowered expectations and a lack of confidence in students, and a low determination to teach minority children. Conversely, Thompson (2002) found that highly efficacious teachers felt innately drawn to the teaching profession.

The practical knowledge of teachers can also greatly impact their teaching ability (Marland, 1986). Research suggests that such knowledge may be incomplete or misguided. More specifically, Marland (1986) conducted a series of teacher interviews and learned that: 1) teachers believed they knew what various student cues meant, 2) teachers rarely checked the accuracy of their interpretations against student expectations, 3) teachers relied heavily on recall, and 4) teacher responses were individualized to the particular student. The teachers interviewed were also asked to reflect upon the problems that they most anticipated within the classroom (Marland, 1986). The common respondent concerns, in order of their salience to school leaders, were: 1) classroom disruption/student inattentiveness/disinterest, 2) irrelevant comments, 4) lesson disorganization, and 5) difficulty with seatwork.

Empirically studying the characteristics that are common to superior classroom performance allows the researcher to better understand how teachers in these settings establish environments that are highly successful. Studying what constitutes good teaching is far from a trivial undertaking, however. Lewis (1978) finds, for example, that "Good teachers can assess and evaluate instructional needs and progress, create exciting learning environments where discovery can occur, stimulate high-level thinking with provocative questions and activities, and thereby promote cognitive growth" (p. 264). Similarly, Waxman, Huang, Abderson, and Weinstein (1997) sought to delineate the differences between effective and efficient schools (E/E) and those schools that are inefficient/ineffective (I/I). The effectiveness of classroom instruction might be related to deeper school cultural issues (Thompson, 2002). Waxman et al. (1997) studied effective schools based upon student activity and learning within classrooms. In the inefficient and ineffective classrooms, teachers interacted with students 47% of the time, while there was no teacher interaction with students 40% of the time (Waxman et al., 1997). In effective and efficient schools, on the other hand, teachers interacted with students 70% of the time, while no student-teacher interaction with students occurred only 21% of the time (Waxman et al., 1997). From Waxman et al.'s (1997) study of respondents within efficient and effective schools, these authors were able to determine the following about the nature of these instructional environments:

1) 72% of time was dedicated to whole class instruction,
2) 20% was independent student work,
3) 98% of the students were working on task,
4) 40% of the time students were watching/listening to the teachers,
5) 29% of classroom instruction was comprised of students engaged in writing activities, and
6) 5% of classroom instructional time consists of students' interactions with one another.

Classrooms that prepare students to excel academically can be studied to determine the traits and mannerisms of excellent teachers. The results of such studies reveal that in inefficient and ineffective classrooms, 81% of class time was dedicated to whole class instruction, 14% of class time was dedicated to independent work, and only 5% of instructional time was dedicated to engaging students in small group work (Waxman et al., 1997). Such knowledge would prove helpful in preparing teachers to provide their students with quality educational opportunities. Thum and Bhattacharya (2001) note certain traits and skill-sets that excellent teachers often possess that appear to be very much related to the work of Waxman et al. (1997):

1) a proactive attitude about discipline,
2) an active working relationship with parents,
3) an assumption of responsibility to motivate students,
4) knowledge of students on a personal level,
5) a willingness to use background knowledge about students, and
6) a proclivity to assign homework that students can complete in class.

Studies also demonstrate that it is commonly the case that accomplished teachers are confident about themselves as teacher leaders, are engaged in leadership roles, yet lack the training needed for the roles they are asked to assume (Dozier, 2007) . Accomplished teachers also appear to be desirous of new leadership roles, training assistance, and becoming acquainted with the methods that would further deepen their knowledge base (Dozier, 2007).

There exists an undeniable disconnect between the knowledge of best practices, which include higher-order pedagogy, and the actual instructional delivery provision of higher-order subject matter engagement of students. Brophy (1990) argues that “it is important that students connect knowledge with beliefs and action. To do that, thinking skills can be developed systematically throughout the years of formal schooling” (Brophy, 1990, p. 364). For instance, Lewis cites a Davis and Tinsley study in which half of the teachers studied asked no application, analysis or synthesis questions (Lewis, 1978). The resulting incongruence between appropriate pedagogical practices and actual classroom instruction can result in learning environments that inadequately prepare students for academic success.

Public school students within the same school district or school are not guaranteed access to equal learning opportunities. Indeed, not all teachers possess the same competency, and a student’s situation within effective or ineffective classrooms can materially affect the nature of his or her learning experience. Students in effective and efficiently run classrooms oftentimes work in individualized settings, yet are afforded ample opportunity to work in small groups where they collaborate with their peers, have healthy levels of interaction with their teachers, and work on written assignments in class (Waxman et al., 1997). Furthermore, the students enrolled in effective and efficient classrooms possessed heightened perceptions of self-concept, affiliation, teacher support, and student aspirations (Waxman et al., 1997). Additionally, the effective and efficient classrooms were housed in schools that had established more positive learning environments (Waxman et al., 1997). Also, teachers in the effective and efficient classrooms were credentialed at disproportionately higher rates than in the ineffective and inefficient schools (Waxman et al., 1997). Research further demonstrates that students enrolled at successful, effective schools felt empowered to exert wider latitude in dictating the nature of their learning experiences (Wilson, 2007).

Not surprisingly, the nature and quality of instruction in high and low-track classrooms starkly differ. In low-track classrooms, for instance, open classroom discussion averaged 3.7 minutes, while open discussion in high-track classrooms averaged 14.5 minutes (Applebee et al., 2003). Such differences are far from superficial, and can dictate the extent to which students are able to effectively pursue future academic endeavors. Indeed, effective preparation and a quality educational experience can empower students to then actively explore their own academic interests. Applebee et al. (2003) argue that “when student’s classroom experiences emphasize high academic demands and discussion-based approaches to the development of understanding, students internalize the knowledge and skills necessary to engage in challenging literacy tasks on their own” (p. 723).

Increased expectations are a sound starting point in ensuring excellent educational environments. Beyond demanding educational excellence for and from students, teachers should provide students with the skills needed to engage in meaningful learning, which includes: 1) pursuing learning that is integrated/interdisciplinary, 2) building deep conceptual knowledge, 3) ensuring that students learn problem solving/learning strategies, 4) ensuring that student learning is contextualized and authentic, and 5) challenging students with various forms of problem solving/mysteries (Wilson, 2007). Authentic learning that involves the construction of knowledge, disciplined inquiry, and learning beyond the classroom has been demonstrated to produce heightened levels of student engagement and learning (Wilson, 2007).

Students who are probed and challenged not only by their instructors, but also by their peers, benefit appreciably from such interaction (Lewis, 1978). Indeed, Lewis (1978) argues that these social interaction processes between not only the teacher and pupil but also between peers are important components of the students' learning experiences (Lewis, 1978). Teachers who stress relevance in their curriculum are more likely to produce active student engagement in the learning process (Brophy, 1990). The quality of instruction is only as sound as the curriculum upon which it is based, however. A dynamic and sound curriculum, as defined by Brophy (1990), includes cohesive content that presents knowledge as a means and not an end, and that includes the value-laden components of decision-making (Brophy 1979; 1990). This allows for the students' assessment of such values, and enables students to develop the skills needed to acquire knowledge and to further explore their values.

Instruction that incorporates teacher questioning is more effective than that which does not (Cotton et al., 1989). The purpose of teacher questioning includes the attempt to motivate students to develop their interests, to evaluate student preparation, to develop critical thinking skills, to assess instructional goals and objectives, and to motivate students to pursue their own interests (Nickerson, 1988). Brophy's (1990) research findings comport with Cotton's contention that not all teacher-directed instruction is undesirable, as he writes that: "I see some value in a certain amount of direct instruction designed to ensure that students know about the cognitive strategies used for learning and applying social studies content" (Brophy, 1990, p. 399). The results of Brophy's (1990) research suggest that some teacher-directed instruction is not only necessary, but desirable.

The ultimate effectiveness of teacher instruction is as influenced by the student's initiative as it is by teacher guidance (Vanosdall et al., 2007). Nevertheless, guided inquiry is often employed in an attempt to actively engage and stimulate student learning (Vanosdall et al., 2007). As the delivery of guided inquiry can be challenging, and teachers and students might lack certain competencies that would render such instruction effective, instructional scaffolds can be helpful (Vanosdall et al., 2007). Such guided inquiry has been shown to lead to substantial test score increases for students when compared to similar cohorts that are exposed to traditional textbook inquiry (Vanosdall et al., 2007).

Higher-Order/Deeper Engagement

Critical thinking, a desirable practice that is often the target of instructional initiatives, involves the judgments and evaluations of arguments that individuals make while solving problems (Geertsen, 2003; Lewis, 1978; Lewis & Smith, 1993). There is an individualized, differential component to critical thinking, in which students relate their personal experiences to the substantive facts of their authentic studies (Pogonowski, 1987). Classroom activities, such as extrapolation or inquiry, that require more abstract mental processing constitutes critical thinking. As such, critical thinking entails “thinking about your thinking while you’re thinking in order to make your thinking better” (Geertsen, 2003).

It is important for educators to make concerted efforts to integrate factual content with the taught thinking habits (Geertsen, 2003). As students become competent in transferring learning skills, this involves extrapolation, which can be facilitated by problem-based learning (Geertsen, 2003). Students who begin to focus on creative thinking must also focus on the learning process itself (Webster, 1990). Teachers who encourage divergent learning strategies, in which multiple and nuanced answers exist, facilitate such critical-thinking skills to a greater extent than those teachers who commonly incorporate activities which generate the convergent thinking commonly associated with standardized test preparation (Webster, 1990).

Engaging students in critical thinking stimulates student learning and prepares them for subsequent educational endeavors (Pogonowski, 1987). The simple acquisition of knowledge that informs students’ information base is a necessary but insufficient component of appropriate instruction, as students should also be engaged in higher-order thinking (Underbakke, Borg & Peterson, 1993). The study of higher-order student thinking and learning requires consideration of how higher-order thinking is to be defined.

There exists no singular, clear-cut metric that one can employ to unambiguously determine whether higher-order thinking is occurring within classrooms. A consideration and definition of the fundamental characteristics of higher-order thinking can guide research and practice in fostering such behavior. Higher-order thinking involves the manipulation of information to arrive at certain outcomes and conclusions (Underbakke, Borg & Peterson, 1993). Higher-order thinking also involves non-algorithmic learning, which is characterized as uncertain inquiry that is effortful and intellectually taxing (Kowalchuk, 1999). Lewis and Smith (1993) offer perhaps the most exhaustive definition of higher-order thinking. In short, they argue that “higher-order thinking occurs when a person takes new information and information stored in memory and interrelates and/or rearranges and extends this information to achieve a purpose or find possible answers to perplexing situations” (p. 136).

To gain a better appreciation of the nature of higher-order thinking, it is useful to contrast it with the lower-order learning that is commonplace within classrooms. Cotton et al. (1989) write that “lower cognitive questions are those which ask the student merely to recall verbatim or in his or her own words material previously read or taught by the teacher” (Cotton et al., 1989). Lower-order thinking does not require student judgment or interpretation, as this lower-order problem solving is largely intuitive and obvious. As such, lower-order thinking requires only basic cognitive skills such as description, explanation, and illustration with examples (Daniel, Lafortune, Pallascio, & Schleifer, 1999; Lewis & Smith, 1993). Conversely, higher-order cognitive questions are “those which ask the student to mentally manipulate bits of information previously learned to create an answer or to support an answer with logically reasoned evidence” (Cotton et al., 1989). Furthermore, higher-order thinking can be characterized by several defining features, which may be evidenced in the form of students’ responding to lectures in a complex way, justifying these responses, expressing a nuance, familiarizing a question, developing logical relationships, hypothesizing, and criticizing (Daniel et al., 1999; Lewis 1978; 1993).

Higher-order thinking is an intellectual practice that actively promotes student learning (Brophy, 1990; Kauffman et al., 1990; Underbakke, Borg, & Peterson, 1993; Freeman, 1989; Kowalchuk, 1999). The educational history and knowledge base of the learner matters; consequently, teachers must be cognizant of a student's previous exposure to certain content material if they are to effectively engage students in appropriate learning (Lewis, 1978). It should be cautioned that the assumption that students must master basic skills before moving on to higher-order skills can lead to inequitable educational experiences for students (Freeman, 1989). Furthermore, state education officials are reluctant to incorporate components of higher-order thinking in standardized tests, as higher-order thinking is hard to test (Freeman, 1989). The challenges associated with an active incorporation of higher-order thinking into the educational curriculum should, therefore, not be underestimated.

Ultimately, it is the teacher that is the mediator of a community of inquiry (Daniel et al., 1999). The teacher's role in establishing higher-order thinking and learning within classrooms is irreplaceable. Actively engaging students in higher-order thinking enables students to more effectively and actively process information (Underbakke, Borg, & Peterson, 1993). A study conducted by Cotton et al. (1989) that involved the collection of classroom observation data to glean the nature of teacher pedagogy found that during an average recitation, 60% of questions were found to be lower cognitive, 20% were higher cognitive and 20% were procedural (Cotton et al., 1989). This is not to suggest that higher cognitive questions are categorically better than lower ones, however (Cotton et al., 1989). It is the case that a certain level of teacher directed pedagogy that provides students with an appropriate knowledge base is both necessary and desirable (Valentine, 2005).

There is no singular or superior way to provide higher-order instruction and thinking to students. It is common, however, for pedagogy that is intended to engage students in higher-order thinking to incorporate instructional methods that encourage students to engage in the following behaviors: 1) hypothesizing and testing, 2) assessing arguments, 3) solving interpersonal problems, and 4) thinking in probabilistic terms (Kowalchuk, 1999). Kowalchuk (1999) argues that teachers must actively consider the extent to which they give due attention to content knowledge, the development of student learning skills, and the awareness of those student dispositions that lead to thoughtfulness and understanding associated with meaningful learning (Kowalchuk, 1999).

Teachers who encourage students to entertain comprehension questions, such as asking them to translate, interpret, and extrapolate, will more successfully engage their students in meaningful higher-order thinking and learning. Enhanced levels of higher-order thinking are often generated when teacher discourse is focused on the application of such information while also relating this information to students' experiences (Kowalchuk, 1999). Indeed, studies have found that simple data gathering leads to smaller student achievement gains, whereas the application of higher-order learning leads to more robustly enhanced student performance (Brophy, 1990). Studies have also demonstrated that students tend to view classes that require elevated levels of higher-order thinking as more challenging, and also as more interesting and engaging (Brophy, 1990).

Research suggests that the nature of student thinking can be influenced by the nature of teachers' pedagogical mannerisms (Marzano, 1993). Teachers most commonly question students as they attempt to enhance student thinking (Marzano, 1993). Underbakke, Borg, and Peterson (1993) note that "while critical thinking skills apparently do not develop spontaneously, a number of research studies have demonstrated that students can learn these skills if they are taught" (p. 141). The teaching of thinking need not be explicit nor mechanical; rather, it can involve enculturation, in which teachers create a culture of thinking in the classroom by 1) "acting as exemplars of metacognition," 2) providing opportunities for students to interact with one another, and 3) providing students with direct instruction in metacognitive practices and activities (Tishman & Perkins, 1993; Underbakke, Borg, & Peterson, 1993). Teachers more commonly employ constructivist strategies, in which meaning is constructed by the learner. Such a practice underscores the vital role that students exhibit as they are taught to learn (Marzano, 1993).

Research suggests that effective lectures are those in which teachers provide overviews before beginning the lecture, use examples to explain and elucidate concepts, and allow time for students' questions (Cotton et al., 1989). Such practices increase the likelihood that students will link new information to their current knowledge base (Cotton et al., 1989). Education researchers have also studied the mechanics associated with the pedagogical processes employed in an effort to teach students certain thinking skills. Studies have found that veteran teachers want to know how to teach students to think (Fogarty & McTighe, 1993). Fogarty and McTighe (1993) investigated the issue and found that teaching thinking skills to students ultimately involve: 1) providing students with sufficient opportunity for skill acquisition, 2) enabling students to make meaning of instructional content, as teachers transition from direct instruction to encouraging students to actively construct knowledge and meaning of the information that they process, and 3) encouraging the transfer and application of knowledge, as students become increasingly aware of how they are learning. The importance of metacognition should not be understated, as "the importance of teaching students to be conscious of the mental processing they use when comprehending information, solving problems, researching topics, communicating with others or making decisions" has been well-documented in the literature (Brophy, 1990, p. 379).

For students to attain deeper levels of content understanding, learners must actively process new information by thinking about and constructing meaning from such information (Nickerson, 1988). Webster et al. (1995) suggest that thinking skills "develop as a consequence of provoked encounters with situations" (Webster et al., 1995, p. 173). The standardized test movement, however, has redirected instructional attention away from such situational learning and toward test-taking strategies that will foster exclusively content learning (Nickerson, 1988).

Higher-order and critical thinking are not the inevitable byproducts of the complex task assignments associated with teacher instruction (Marzano, 1993). Instead, critical thinking results from the engagement of students in tasks where the answer is not readily apparent, where teachers push the limits of student knowledge, and where students are asked to generate new and unconventional ways of viewing situations (Marzano, 1993). Engaging students in abstracting strategies, for instance, requires that a student "links literal information to that which does not appear related at a surface level" (Marzano, 1993, p.158).

The current challenge for teachers is not simply to teach thinking, but rather to teach good thinking (Nickerson, 1988). Educators often forget that teaching students how to learn is different from simply teaching them how to take tests (Cooper, 1989). The narrow focus on test preparation has encouraged teachers to promote rote learning. While the future demands placed upon students are likely to involve enhanced critical thinking skills, public schools are not properly equipping students to meet such challenges. Greeno (1997) convincingly argues that appropriate thinking skills are not being taught in schools, as he notes that "the activities of school learning are mainly organized so that students can accumulate the skills that they need to think with rather than presenting them with problems that present challenges for complex thinking for which they are assumed not to be ready" (p. 88).

The traditional focus of classroom instruction has changed little, even as the wider political and accountability environment is remarkably different than in years past. Cooper (1989) reports that 80-95% of classroom work is derived from published instructional material. While this represents a most expeditious way of preparing students for high-stakes testing, it may fail to enhance students' critical thinking and reasoning abilities. Indeed, it is not uncommon for instruction to involve teaching lower-order, basic skills so that students perform satisfactorily on standardized tests, such as the NAEP (Cooper, 1989). Meaningful student learning and effective test preparation need not be mutually exclusive undertakings, however. To effectively facilitate student learning, teachers must gain the commitment from their students (Cooper, 1989). Furthermore, the classroom instruction should be interactive and interdisciplinary, as opposed to simply providing students with the materials that will enable them to produce recall information on standardized tests (Cooper, 1989). Although the inquiry-based instruction that produces reasoning gains for students can be more time consuming and demanding of teachers, the resulting gains to student learning make the extra efforts justifiable.

Methods

The Instructional Practices Inventory (IPI) process for codifying student engagement coding rubric utilizes "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, 2009). The IPI instrumentation allows a trained classroom observer to collect approximately 100-150 observational codes that capture student engagement behaviors for each school over a specified period of time, usually a full school day. 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. A copy of the coding rubric and related descriptions of each category are provided in Appendix A.

The IPI process focuses on student engagement and cognitive thinking rather than teacher or student behavior. In Classroom Walkthroughs to Improve Teaching and Learning, Kachur, Stout, and Edwards, the IPI process was described, along with twelve other "most commonly used walk-through observations processes" as a viable system for defining student engagement. The IPI process was also recommended as a school improvement process in the National Association of Secondary School Principals' *Breaking Ranks in the Middle (NASSP, 2006).* Through a grant from the Institute of Educational Sciences, the Southeast Regional Educational Laboratory recently published a report entitled *Measuring Engagement in Upper Elementary through High School: A Description of 21 Instruments* (REL Southeast, 2011). Among the 21 instruments described in the report, four were observation instruments, two of which, including the IPI, were classroom observation processes. Three of the four observational instruments focused on classroom/students' behavioral and emotional engagement. The IPI was the only process selected for the report that was observational of cognitive engagement the classroom learning setting. The authors of the report objectively described the IPI Process and noted that it was a school improvement process designed to foster faculty collaborative study of the IPI data profiles with the purpose of shaping the nature of engagement across the school.

Valentine (2009) reports that the IPI profile data can be used to foster teacher engagement in whole-faculty and small-group collaborative analysis, reflection, and decision-making. The IPI instrumentation, and the accompanying building-level instructional processes, can ultimately provide telling and comprehensive school-wide data that allow teachers and administrators to continuously monitor and refine their pedagogical practices. These components of the IPI process support continuous change and collectively foster organizational learning.

Undoubtedly, there exists a multitude of factors whose impact on student learning are noteworthy. This exploratory study was designed to glean the extent to which student engagement levels may or may not be associated with gains in the standardized achievement performance of public school students. The study is constructed in a manner whereby the researcher is able not only to offer dichotomous "yes/no" conclusions about such a relationship, but also to expound upon the magnitude with which different forms of student engagement ultimately impact students' abilities to perform at or above the proficiency levels on the Missouri Assessment Program (MAP) standardized tests.

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-order 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, and categorized accordingly. The IPI strikes a methodologically appropriate balance between meaningfully sorting student engagement into 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, the researcher believes that 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 across six categories that collectively account for the spectrum of student engagement that instructors can expect to find in any given classroom at a particular moment.

An explanation of the six coding categories is provided in Appendix A. It is important to note that the higher-order categories ("5" and "6") represent highly desirable forms of student learning, whereas the student disengagement category ("1") and the teacher disengagement category ("2") represent less effective and generally undesirable, indefensible forms of disengagement and engagement. It is not always possible, nor desirable, however, for students to be engaged solely in higher-order activities. As such, categories "3" and "4" account for those moments during classroom instructional time when the teacher is primarily involved in informing, directing, and supporting lower-order learning in commonly passive, inactive learning experiences. 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.

Engagement data for the 68 study schools represented more than 10,000 classroom observation codes. The descriptive data for the variables analyzed in this study are presented in Table One. Student disengagement in all classes (Category T1) averaged 4.55% across the 68 schools and disengagement in only core classes (Category C1) averaged 4.16%. The percent of observations coded as student lower-order engagement when the teacher was not engaged with, attentive to, or supportive of the students (Category 2) was slightly more than 10% for all classes (T2) and core classes (C2). The percentage of observations in which student engagement was observed to be higher-order was considerably greater for student active engaged learning (T6) than for higher-order student discussions (T5). In total, higher-order thinking (Categories 5 and 6 combined) comprised 20% of classroom observations in the study's middle schools.

<u>Descriptive Data</u>

Demographically the 68 public middle schools contained a relatively diverse population sample of Missouri public school students. The average percentage of free and reduced lunch rate (FRL) was nearly 45%, while the percent of minority students in each school averaged nearly 19%.

The 68 middle schools within the study were linked with their 58 respective school districts. The descriptive data for the districts are provided in Appendix B. The average per pupil expenditure was just over $7900, while both the Communication Arts and Mathematics proficiency rates were approximately 45%. Seventy-six percent of students' families had retained a residence within the county for the last five years, while nearly 58% of students' belonged to households of married parents.

The Three-level models tested within the study contained a regional level as the third level of the models. Missouri is comprised of nine Regional Professional Development Center's (RPDC's). The data associated with these nine RPDC's are presented in Appendix C. At the regional level, the per pupil expenditure average was $8147.27, the average percent of minority students was 12.42%, and the average percent of the students eligible for free-and-reduced lunch (FRL) was 45%.

A Realistic Application of the Findings

The essence of the overall IPI process is to foster collaborative study of the data and for the collaborative study to promote school-wide changes in the nature of student engagement throughout the school. After an initial IPI data collection, school administrators and teacher leaders immediately become aware of their school's current student engagement profile in percentage terms. Quantifying student engagement behaviors is not only diagnostically meaningful, but presents the opportunity for more healthy and constructive goal setting at the building level. Indeed, principals and teachers are then empowered to chart a more data-driven course for engagement. The designated benchmarks, in turn, will augment standardized test passage rates according to the findings fleshed out in this study.

Teacher leaders and administrators can approach the process of change with meaningful, valid data about engagement. Estimations and conjectures about engagement can lead to arbitrary mandates by school leaders and high levels of resentment and dissent among faculty. With the IPI process, a well-designed plan for addressing engagement can unfold. For instance, a school may set a goal of achieving a specific level of higher-order/deeper engagement (Categories 5 and 6 combined) and establish action plans to move toward that goal in a steady progressive manner that reflect changes in engagement throughout the school year, not merely on those "data collection" days when profiles are developed. Such progress must be purposeful, effectively conceived and designed by faculty, and implemented with faculty support, all of which is viable when faculty are empowered to collect the data, lead the study of the data, develop meaning from the study of the data, and design plans to enhance engagement.

A pre-designated benchmark of engagement levels enables principals and teacher leaders to more purposively, confidently, and competently attack lagging higher-order student engagement levels and rising levels of student disengagement. While the optimal level of total disengagement (Category 1) is, of course, zero percent, faculty members are also wise to appreciate the devastating effect that ballooning student disengagement levels can exact on both teacher morale and standardized achievement. Consider a level of 25% disengagement across a school. To attain such a level, the school (principals and teachers collectively) would have to have been asleep at the wheel for many school years. Such levels seldom occur in a short period of time, and commonly are the result of years of negligence, essentially educational malpractice.

Also important is the temporal design of the IPI process. The IPI process is not a quick fix or shock treatment meant to remedy all that ails a school's instructional health instantaneously. Instead, the IPI process demands from faculties a sustained commitment to altering their pedagogical techniques and practices over a sustained time horizon. It is in this vein that teacher leaders and principals can employ the student engagement benchmarks not as a punitive or heavy-handed oversight metric, but as attainable building-level guideposts that signify faculty growth, commitment, and instructional excellence. Presently, for instance, higher order student engagement levels hover around 20% of all student engagement activity within a school. Enhancing the present level of higher-order thinking to a considerably more ambitious 40-60% of all student classroom behavior could seem to be an unduly formidable obstacle for a school's faculty. Likewise, such growth could also require an exceptionally forceful probing from a nervously reluctant principal and/or central office administrator. Viewing the higher-order gap in current and theoretically optimal higher-order thinking levels over a four year time horizon fundamentally simplifies the process. Deconstructing the desired amount of change into 4-5% changes over time makes for a more palatable plan. The incremental approach established highly desirable and attainable goals without precipitating a highly stressful and hopeless perspective about the desired change. Redefining this optimal goal as short, digestible semi-annual or annual goals enable principals and teacher leaders to guide the faculty to the "Pedagogical Promised Land" of higher-order thinking optimality with considerable momentum and manageable resistance. Indeed, mapping out for a faulty a two or three-year plan with incremental benchmarks is a strategy that would likely be accompanied by annual standardized achievement gains of between 3-5%. As such, faculty morale would swell, their pedagogical techniques would become noticeably more expert, and this transformative effect would provide the requisite propulsion needed to thrust the school's higher-order thinking levels toward the desired goal.

HLM Analyses

The population of middle schools considered in the present study had a level of student and teacher disengagement (Categories 1 and 2) of approximately 15% for all observations. Higher-order/deeper engagement (Categories 5 and 6) consisted of approximately 20% of all observations. The findings suggest that were these middle schools to create educational settings in which student and teacher disengagement were nonexistent, and where higher-order student engagement was a more robust 60% of all reported observations, standardized achievement gains might be considerably enhanced. The Mathematics engagement-achievement relationship for student disengagement appears to be almost identical to that of Communication Arts, but the output capturing the relationship between Mathematics and higher-order student engagement levels was not found to be statistically significant.

Two-Level Findings:

Engagement Construct: Communication Arts proficiency levels comprised one dependent variable of interest in the study (Table Two). The engagement construct for student disengagement within core-classrooms ("C1") yielded a slope of -.51, and -.43 for student disengagement in all classrooms ("T1"). When the mathematics proficiency rate was assigned as the dependent variable, a core classroom disengagement slope of -.47 was evidenced. The higher-order thinking slope was a more diminished .18, however. Such a finding is particularly valuable, as it suggests that disengagement can detrimentally effect standardized achievement to an extent that is nearly three times more impactful than higher-order student engagement.

Racial/Socioeconomic: The greatest magnitude for the percentage of minority student independent variable assigned to models with the corresponding Communication Arts dependent variable was -.15, while mathematics was a more augmented -.27. Elevated regression coefficients for level-one FRL independent variables were offset by Level-two FRL predictors with dueling directionality. When the aggregate slope values are computed, slopes as great as -.25 are found for the Communication Arts dependent variable, while the level-two FRL slopes for the mathematics outcome models were found to be statistically insignificant, precluding a like computation for that dependent variable.

Three-Level Findings

Engagement Construct: The higher and lower-order regression coefficients in the three-level models (Table Three) were smaller than in the two-level models. More specifically, the slope for the percentage of student disengagement within core classrooms ("C1") was found to be -.44 for the Communication Arts dependent variable models, while the higher-order thinking construct for core classrooms in the model was a more tempered .22 ("C56"). As for the mathematics dependent variable models, teacher disengagement in core classrooms ("C2") was -.26 and -.28 in all classrooms ("T2").

Racial/Socioeconomic: At the district level, the slope associated with the percentage of minority students within schools was as high as -.11 for Communication Arts and -.23 for Mathematics. The regional level (level three) percent minority construct was found to be -.29 for Communication Arts, leaving an aggregate slope of -.40 for Communication Arts. The regional level percentage minority slope for the Mathematics dependent variable was not found to be statistically significant in any of the tested models. The free and reduced lunch rate was also not found to be statistically significant across levels. At level one, however, the slope was as high as -.47 for Communication Arts and -.57 for Mathematics. Statistically significant offsetting effects might be found at the district (level-two) and regional levels (level-three). Accordingly, a FRL analysis based solely upon the level-one findings may be neither straightforward nor accurate.

Practical "Takeaways" for Principals and Teacher Leaders

The findings of this study inform school leaders as to the effects that enhanced or diminished higher and/or lower-order thinking can exhibit on standardized achievement levels. To illustrate how the findings may be applicable to building-level faculty and administrators, consider a scenario under which student disengagement in core classes (category C1) and teacher disengagement in all classes (category T2) increases of 25% of all coded observational data. In such a scenario the Communication Arts pass rates would decrease 10.63 percentage points, while Mathematics pass rates would decrease by 9.79 percentage points, according to the two-level model findings.

More elaborate three-level statistical models evidence similar results to the two-level models. The Communication Arts pass rates would decrease 9.17 percentage points. Mathematics pass rates would decrease by 4.11 percentage points.

Exploring whether higher-order student engagement levels within classrooms can appreciably enhance student achievement levels can be tested by manipulating the proportion of higher-order student engagement within classrooms from its current level of 18.93% to a robust 60% of all coded classroom behavior. The results of the heightened higher-order thinking levels are compelling evidence of the importance of higher-order/deeper thinking across the classrooms of a school. The increase of higher-order engagement by approximately 40 percent would produce Communication Arts achievement pass rates increases of 7.93% for the two-level model and 9.04% for the three-level model. Interestingly, the negative effect of increasing student and teacher disengagement by roughly 20% produces almost identical results to enhancing higher-order/deeper engagement by approximately 40%. Reducing disengagement while simultaneously increasing deeper engagement provides the fast-track to student growth.

Clearly, the impact of such changes provide promise for educators that improving student achievement is, to a significant degree, within the realm of the classroom teacher. As student engagement in higher-order/deeper learning goes up, so do achievement scores. As levels of disengagement go down, achievement levels rise. Prior studies (Collins and Valentine, 2010) document that the ratio of disengagement to achievement is approximately two to one. In other words, as disengagement moves in one direction two percentage points, achievement move in a counter direction one percent pass rate on the Missouri high stakes test. A school-side focus on deeper engagement and disengagement is a powerful weapon to increase not only achievement scores on high-stakes tests, but more importantly the overall capacity of students to think more deeply, reason more effectively, and thus become more capable life-long learners.

According to the US Office of Education, the average length of a school day in the United States is six and one-half hours. Students are engaged in higher-order/deeper learning approximately 40 minutes out of that time. It takes very little effort for a teacher, or a team of teachers, to ratchet up the opportunity for students to engage in deeper learning and make a significant difference for students. If a student has six teachers during the course of the school day and three of those teachers add one additional five-to-six minute higher-order engagement learning experience, deeper learning would then be 55 minutes, an increase of nearly 40% more higher-order/deeper learning time. The result of an increased diet of higher-order thinking could produce significant gains in student achievement. A steady diet of cumulative deeper thinking would be to the mind what healthy nutrition is to the body, producing life-long mental gains for the students. The true value of such gains does not lie in the increased test scores, but in the long-term effects for students and society.

Table One: Descriptive Statistics of the 68 Study Middle Schools (Level One)

VARIABLE NAME	N	MEAN	SD	MINIMUM	MAXIMUM
DISCIPLI	68	2.51	3.19	0.00	16.90
FRL	68	44.75	16.57	16.20	85.90
PCT_MIN	68	18.74	24.87	0.00	100.00
COMM_07	68	43.98	10.80	12.60	63.10
MATH_07	68	44.73	12.53	7.80	64.70
AV_T1	68	4.55	3.52	0.00	16.50
AV_T2	68	10.33	5.95	1.00	34.50
AV_T5	68	4.31	3.46	0.00	17.00
AV_T6	68	15.36	7.22	2.00	38.00
AV_C1	68	4.16	3.36	0.00	16.50
AV_C2	68	10.42	6.39	0.00	34.50
AV_C5	68	4.61	3.58	0.00	16.67
AV_C6	68	14.32	7.38	1.00	33.00

Table Two: Two-Level Output

DV	Engagement Category	Teachers with Masters	FRL 1	Student: Teacher Ratio	Engage Slope	Per Pupil Expenditure	Percent of Minority Students	FRL 2	Percent Not Transient	Percent Married Households
Comm	C1	.12*	-.44***	.14	-.51**	.01	-.11*	.20**	-.33**	.19
Comm	C2	.13*	-.46***	.07	-.09	.03	-.15**	.21**	-.33**	.20
Comm	C56	.12*	-.38***	.07	.18**	.01	-.13**	.15*	-.39**	.23**
Comm	T1	.12*	-.43***	.13	-.43**	.00	-.11*	.19**	-.34**	.20
Comm	T2	.13*	-.46***	.08	-.13	.04	-.15**	.21**	-.33**	.19
Comm	T56	.13*	-.41***	.08	.08	.02	-.14**	.19*	-.36**	.25**
Math	C1	.16**	-.50***	.63*	-.47*	.12	-.23***	.16	-.09	.14
Math	C2	.17**	-.53***	.50	.26	.15*	-.26***	.17	-.10	.08
math	C56	.16	-.48***	.52	.09	.12	-.25***	.16	-.12	.18
Math	T1	.16**	-.50***	.62	-.40	.11	-.23***	.17	-.10	.15
Math	T2	.17**	-.53***	.52	-.28	.16*	-.27***	.18	-.10	.08
Math	T56	.17**	-.48***	.52	.11	.11	-.25***	.17	-.13	.20

Table Three: Three-level Model Output

DV	Engagement Category	Teachers with Masters	FRL2	FRL1	Student-Teacher Ration	Engagement	Per Pupil Expenditure	Percent of Minority Students	FRL3	Percent of Minority Students	Percent Married Households
Comm.	T2	.14**	.16	-.47***	-.09	-.12	.02	-.11*	.45*	-.29**	.16
Comm.	C1	.13**	.15	-.45***	-.02	-.44*	.00	-.08	.41	-.29**	.17
Comm.	C2	.14**	.16	-.47***	-.11	-.10	.02	-.11*	.47*	-.29**	.16
Comm.	C56	.13**	.09	-.38***	-.14	.22**	-.02	-.08	.56**	-.36***	.19
Comm.	T56	.14**	.12	-.42***	-.12	.12	-.01	-.09	.53*	-.34**	.21
Math	C1	.17**	.15	-.52***	.51	-.40	.11	-.21***	.28	-.06	.13
Math	C2	.18***	.16	-.57***	.31	-.26*	.13	-.23***	.34	-.09	.04
Math	T1	.17**	.15	-.53***	.47	-.34	.11	-.20***	.30	-.07	.14
Math	T2	.18***	.17	-.51***	.34	-.28*	.14	-.23***	.32	-.09	.05
Math	T56	.18***	.13	-.51***	.34	.16	.10	-.21***	.42	-.12	.18
Math	C56	.17**	.13	-.51***	.36	.11	.10	-.21***	.35	-.11	.16
Comm.	T1	.12**	.13	-.44***	-.03	-.38	-.01	.08	.43	-.30*	.18

Appendix A: The Instructional Practices Data Coding Rubric

Student Active Engaged Learning (6)	Students are engaged in higher-order thinking and developing deeper understanding through analysis, problem solving, critical thinking, creativity, and/or synthesis. Engagement in learning is not driven by verbal interaction with peers, even in a group setting. Examples of classroom practices commonly associated with higher-order/deeper Active Engaged Learning include: inquiry-based approaches such as project-based and problem-based learning; research and discovery/exploratory learning; authentic demonstrations; independent metacognition, reflective journaling, and self-assessment; and, higher-order responses to higher-order questions.	**Student Engagement in Higher-Order Deeper Learning**
Student Verbal Learning Conversations (5)	Students are engaged in higher-order thinking and developing deeper understanding through analysis, problem solving, critical thinking, creativity, and/or synthesis. The higher-order/deeper thinking is driven by peer verbal interaction. Examples of classroom practices commonly associated with higher-order/deeper Verbal Learning Conversations include: collaborative or cooperative learning; peer tutoring, debate, and questioning; partner research and discovery/exploratory learning; Socratic learning; and, small group or whole class analysis and problem solving, metacognition, reflective journaling, and self-assessment. Conversations may be teacher stimulated but are not teacher dominated.	
Teacher-Led Instruction (4)	Students are attentive to teacher-led instruction as the teacher leads the learning experience by disseminating the appropriate content knowledge and/or directions for learning. The teacher provides basic content explanations, tells or explains new information or skills, and verbally directs the learning. Examples of classroom practices commonly associated with Teacher-Led Instruction include: teacher dominated question/answer; teacher lecture or verbal explanations; teacher direction giving; and, teacher demonstrations. Discussions may occur, but instruction and ideas come primarily from the teacher. Student higher order/deeper learning is not evident.	**Student Engagement in Knowledge and Skill Development**
Student Work with Teacher Engaged (3)	Students are engaged in independent or group work designed to build basic understanding, new knowledge, and/or pertinent skills. Examples of classroom practices commonly associated with Student Work with Teacher Engaged include: basic fact finding; building skill or understanding through practice, "seatwork," worksheets, chapter review questions; and multi-media with teacher viewing media with students. The teacher is attentive to, engaged with, or supportive of the students. Student higher-order/deeper learning is not evident.	
Student Work with Teacher not Engaged (2)	This category is the same as Category 3 except the teacher is not attentive to, engaged with, or supportive of the students. The teacher may be out of the room, working at the computer, grading papers, or in some form engaged in work not directly associated with the students' learning. Student higher-order/deeper learning is not evident.	
Student Disengagement (1)	Students are not engaged in learning directly related to the curriculum.	**Students Not Engaged**

IPI coding is not based on the type of activity in which the student is engaged, but rather how the student is engaging cognitively in the activity. Examples provided above are only examples often associated with that category. The Instructional Practices Inventory categories were developed by Bryan Painter and Jerry Valentine in 1996. Valentine refined the descriptions of the categories (2002, 2005, 2007, and 2010) in an effort to more effectively communicate their meaning.

<u>The IPI was developed to profile school-wide student engaged learning and was not designed for, nor should it be used for, personnel evaluation.</u>

Appendix B: Level Two Descriptive Statistics for the 58 School Districts used in the Two and Three-Level HLM Models

VARIABLE NAME	N	MEAN	SD	MINIMUM	MAXIMUM
PER_PUP	58	79.39	15.62	60.58	143.95
PCT_MIN	58	16.43	23.45	0.40	99.50
FRL	58	43.41	13.53	13.10	78.50
COMM07	58	44.75	7.93	18.60	64.70
MATH07	58	45.56	9.55	14.40	63.90
PCT_NOT	58	75.89	7.46	49.50	87.60
NOW_MARR	58	57.73	7.61	32.70	68.90

Appendix C: Level Three Descriptive Statistics for the 9 Regional Professional Development Centers used in the Three-Level HLM Model

VARIABLE NAME	N	MEAN	SD	MINIMUM	MAXIMUM
PER_PUP	9	8147.27	877.06	7270.11	9415.83
PCT_MIN	9	12.42	11.02	3.61	32.43
FRL	9	44.68	4.55	38.24	53.01
COMM07	9	43.97	1.32	41.56	45.79
MATH07	9	45.10	1.92	41.60	47.68
PCT_NOT	9	77.78	2.49	73.98	82.80
NOW_MARR	9	59.36	2.65	55.90	63.39

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.

Brophy, J. (1990). Teaching social studies for understanding and higher-order applications. *The Elementary School Journal, 90*(4), 351-417.

Brophy, J. E. (1979). Teacher praise: A functional analysis. Lansing, MI: The Institute for Research on Teaching, Michigan State Univ.

Caldas, S. J., & Bankston III, C. L. (1999). Multilevel examination of student, school, and district-level effects on academic achievement. *The Journal of Educational Research, 93*(2), 91-100.

Collins, J., & Valentine, J. (2010). Testing the impact of student engagement on standardized achievement: An empirical study of the influence of classroom engagement on test scores across school types. Paper presented at the Annual Conference, University Council of Educational Administration, New Orleans, LA. (October 30, 2010).

Cooper, E. J. (1989). Toward a new mainstream of instruction for American schools. *The Journal of Negro Education, 58*(1), 102-116.

Cotton, K. (1989). Critical thinking across the curriculum: Aloha High School. *School Improvement Research, Series III*, 57-61.

Daniel, M.-F., Lafortune, L., Pallascio, R., & Schleifer, M. (1999). Philosophical reflection and cooperative practices in an elementary school mathematics classroom. *Canadian Journal of Education, 24*(4), 426-440.

Druian, G., & Butler, J. A. (1987). *Effective schooling practices and at-risk youth: What the research shows*. Portland, OR: Northwest Regional Educational Laboratory.

Fogarty, R., & McTighe, J. (1993). Educating teachers for higher order thinking: The three-story intellect. *Theory into Practice, 32*(3), 161-169.

Freeman, D. J. (1989). State guidelines promoting teaching for understanding and thinking in elementary schools: A 50-state survey. *Educational Evaluation and Policy Analysis, 11*(4), 417-429.

Geertsen, H. R. (2003). Rethinking thinking about higher-level thinking. *Teaching Sociology, 31*(1), 1-19.

Goddard, Y. L., Goddard, R. D., & Tschannen-Moran, M. (2007). A theoretical and empirical investigation of teacher collaboration for school improvement and student achievement in public elementary schools. *Teachers College Record, 109*(4), 877-896.

Greeno, J. G. (1997). Theories and practices of thinking and learning to think. *American Journal of Education, 106*(1), 85-126.

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

Kachur, D., Stout, J., & Edwards, C. (2010). Classroom walkthroughs to improve teaching and learning. Larchmont, NY: Eye on Education.

Kowalchuk, E. (1999). Recognizing and using higher order approaches to art. *Art Education, 52*(6), 13-18.

Lewis, E. S. (1978). Teacher influence on pupil cognitive performance. *Peabody Journal of Education*, *55*(3), 252-264.

Lewis, A., & Smith, D. (1993). Defining higher order thinking. *Theory into Practice, 32*(3), 131-137.

Marzano, R. (1993). How classroom teachers approach the teaching of thinking. *Theory into Practice, 32*(3), 154-160.

National Association of Secondary School Principals. (2006). Breaking ranks in the middle. Reston, VA: National Association of Secondary School Principals.

Nickerson, R. S. (1988). On improving thinking through instruction. *Review of Research in Education, 15*, 3-57.

Pogonowski, L. (1987). Developing skills in critical thinking and problem solving. *Music Educator Journal, 73*(6), 37-41.

Regional Education Lab Southeast (2011). Measuring student engagement in upper elementary through high school: A description of 21 instruments. Greensboro, NC: Regional Educational Laboratory at SERVE Center. (January, 2011; document #: 098)

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

Thompson, G. L. (2002). African American students in schools: Research and effective instructional practices. *Education Horizons, 80*(3), 105-116.

Thompson, G. L. (2002). Elementary teachers. *Education Horizons, 80*(3), 147-152.

Thum, Y. M., & Bhattacharya, S. K. (2001). Detecting a change in school performance: A Bayesian analysis for multilevel join point problem. *Journal of Educational and Behavioral Statistics, 26*(4), 443-468. 419

Tishman, S., Jay, E., & Perkins, D. N. (1993). Teaching thinking dispositions: From transformation to enculturation. *Theory into Practice, 32*(3), 146-153.

Underbakke, M., Borg, J. M., & Peterson, D. (1993). Researching and developing the knowledge base for teaching higher order thinking. *Theory into Practice, 32*(3), 138-146.

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. (2009). *The instructional practices inventory: Using a student learning assessment to foster organizational learning*. National Staff Development Council Annual Conference, St. Louis, MO (December, 2009)

Vanosdall, R., Klentschy, M., Hedges, L. V., & Weisbaum, K. S. (2007). *A randomized study of the effects of scaffolded guided-inquiry instruction on student achievement in science*. Paper presented at the American Educational Research Association, Chicago, IL

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*, 420 49-59.

Webster, P. R. (1990). Creativity as creative thinking. *Music Educators Journal, 76*(9), 22-28.

Wilson, L. (2007). Great American schools: The power and culture of passion. *Education Horizon, 86* (1), 35-44.