

HIGH-QUALITY LITERATURE-BASED INSTRUCTION:
AESTHETIC READER-RESPONSE AND CRITICAL PEDAGOGY
FOR DIVERSE ELEMENTARY SCHOOL STUDENTS

by Tiffany Sabatini

A Project Submitted to the Faculty of
The Evergreen State College
In Partial Fulfillment of the Requirements for the Degree
Master in Teaching

2013

This Project for the Master in Teaching Degree

by

Tiffany Sabatini

has been approved for

The Evergreen State College

by

Jon Davies, Member of the Faculty

June 2013

ACKNOWLEDGEMENTS

I thank my high school English teacher, Mr. Oldham, for helping me know that I can always use literature to grow, change and develop empathy with others. I thank the Institute for Community Leadership for allowing me to be part of the enactment of Paulo Freire's ideas. My gratitude goes to Lester Krupp for helping me articulate my ideas about literacy and for introducing me to the work of Louise Rosenblatt; this paper was a continued exploration of ideas that inspired me in his classes. I thank Terry Ford, Sherry Walton, Jon Davies, Rob Cole, and Sunshine Campbell for helping me learn the social and educational context of the ideas explored in this paper. In particular, I thank Jon for his jokes and his wonderful support throughout the process of writing this paper. I thank my mentor teacher, Deborah Neid, for her interest in enacting the findings of this analysis. I thank the cohort for sharing their ongoing support and passion.

I also thank my friends and family for their inspiration, the sharing of their own experiences of art and school, and their joyful companionship; their energy sustained me in the process of writing this paper. Most of all I thank my dearest friend and husband, Christopher Sabatini, for always reminding me that my ongoing life as an artist, reader, and writer has led to my passion for teaching—I am forever thankful for his perseverance and love.

ABSTRACT

This paper provides a critical review of the professional research examining the effects of literature-based instruction on the literacy of diverse elementary school students. It examines aesthetic reader-response and critical pedagogy as processes that may qualify the effectiveness of literature-based instruction. The research suggests that aesthetic reader-response and critical pedagogy apply to all students regardless of designated ability, and that it contributes to student motivation to use literacy for identity, empowerment, and cultural change. The effectiveness is increased with the integration of metacognitive strategy instruction. Debate still exists as to whether these processes produce equal results on standardized tests as skills-based instruction. This paper also analyzes the implications for classroom practice and suggests directions for further research.

TABLE OF CONTENTS

TITLE PAGE	i
APPROVAL PAGE	ii
ACKNOWLEDGEMENTS	iii
ABSTRACT	iv
CHAPTER 1: INTRODUCTION	1
Introduction	1
Rationale	2
Historical Context	7
Definitions	13
Limitations	16
Summary	17
CHAPTER 2: CRITICAL REVIEW OF THE RESEARCH	18
Introduction	18
Aesthetic Reader-Response	19
Critical Pedagogy	33
Group Discussion	47
Motivation & Engagement	62
Best Practices in Literacy Instruction	78
Strategy Instruction	93
Achievement	106
Summary	131
CHAPTER 3: CONCLUSION	133
Introduction	133

Summary of Findings	133
Implications for Classroom Practice	138
Suggestions for Further Research	143
Conclusion	146
REFERENCES	147

CHAPTER 1: INTRODUCTION

Introduction

In the wake of the 1990s “reading wars” between whole language and skills-based approaches to literacy, educational researchers conclude that more research on whole language is necessary. The same is true for literature-based instruction, which is like whole language but only focuses on literary texts, rather than informational texts in all content areas. Aesthetic reader-response and critical pedagogy are processes that literature-based instruction may or may not include. There is very little research on these distinctions in literature-based instruction for elementary school students.

Tindal & Marston (1996) noted in their study of reading assessment that researchers lack methods of “scientifically sound” measurements that are “psychometrically adequate” to the criteria of whole language (or literature-based) approaches (p. 206). Research on the effectiveness of literature-based instruction is problematic for two reasons. First, different educators define literature-based instruction differently (Stahl, 1999). Secondly, the benefits that advocates of literature-based instruction tout are qualitative, such as social and cognitive ways of using reading, rather than only quantitative benefits like test scores. Unfortunately, the National Reading Panel (2001) left out studies of certain kinds of literacy abilities, favoring studies that focus on technical and measurable aspects of literacy, such as sound-symbol relationships, word recognition, fluency and comprehension (Cummings, 2007; Taylor, Pearson, Peterson, & Rodriguez, 2003; Willis & Harris, 2000). The panel itself stated that

it did not focus on many critical aspects of literacy (2001). This state of affairs remains true in the field of literacy research today.

Two of the missing aspects of literacy from the NRP meta-analysis, among others that have potential to make a difference for the literacy of all learners, are aesthetic reader-response and critical pedagogy (Au & Raphael, 2000; Cummins, 2007; Willis & Harris, 2000). These practices theoretically allow students to create a meaning for how literacy applies to their real life, and thus they accommodate students of diverse backgrounds, and motivate their endeavors to increase their reading ability. This review of the research on these topics may help to better define both literacy and quality literature-based instruction, thus providing a more accurate framework for future research. This paper reviews 37 studies in order to explore the question: What are the effects of aesthetic reader-response and critical pedagogy on the literacy of diverse elementary school students?

Rationale

Due to the Bush Administration's selective response to the National Reading Panel (2001) in the form of No Child Left Behind and the Reading First program, the Department of Education made an unprecedented swing to mandating intensive scripted phonics approaches in schools with low scores on standardized tests (Cummings, 2007; Stahl, 1999; Willis & Harris, 2000). Schools with sufficient achievement scores had more room to choose the kind of literacy instruction they used. Studies were publicized in the media noting that Reading First had no effect and was sometimes detrimental to the reading scores

for low-achieving schools (Glod, 2008). Reading First's funding was cut in 2009 and Obama approved cuts in funding to many other literacy programs as part of a stopgap bill in March 2011 (Klein, 2011); there has been no replacement to date. The question of how to best support literacy engagement and achievement for all students comes back to individual school districts, schools, parents, and teachers. In regard to literature-based instruction, and more specifically aesthetic reader-response and critical pedagogy, educators face controversies when making a decision about how to approach literacy. This paper explores three inter-related controversies.

The first controversy, which was the central issue of the "reading wars," is that some educators believe that literature-based instruction does not include an explicit focus on skills and reading strategies, and that this is particularly detrimental for students who have less cultural experience with literature. Lisa Delpit is a proponent of the socio-cultural approach to literacy (which literature-based instruction is founded on). However, she criticized literature-based educators that did not include skills and explicit focus on code-switching for students who speak dialects other than standard English (2006). In addition, Carol Gilles, pointed out that the "whole" in whole language meant integrating all reading strategies and skills into a meaningful literacy context. To be congruous with its theoretical founding, literature-based instruction must include a focus on skills and strategy work (1988).

Essentially, the dichotomy of the "reading wars" was false. Unfortunately, the NRP report (2001) only fed the fire, as NCLB policy-makers interpreted the

findings as advocating for a skills only approach, supposedly because basic-skills are easily testable. The policy hamstrung teachers into a teach-to-the-test mania (Au & Raphael, 2000; Johnson & Johnson, 2002; Ravitch, 2010; Willis & Harris, 2000). The National Council of Teachers of English re-asserted that a culturally relevant and meaningful context is essential for effective instruction (1998-2012). This paper includes a section on strategy and skills instruction as it is integrated into literature-based instruction in order to bring clarity to the definition of effective literature-based instruction.

A second controversy has to do with which demographics of students benefit from aesthetic reader-response and critical pedagogy. Willis & Harris (2000) asserted that a problem with the assumption that higher level thinking is not for students that scored low on tests is that the dominant culture of “scientism” in research relies on biased reading assessments that have been based on the cultural norms of white middle-class suburban students. This dominant culture takes a deficit approach to student achievement. A deficit model uses the tests to see which students have the skills and which students do not. To focus on problem-solving the achievement of those students that don’t score high, schools use basic skills programs to teach the specific skills needed to pass the test. The idea behind this approach is that basic skills need to come before deep thinking about literature, decoding is the foundation for all reading, and equality is about figuring out how to get students caught up. Those who take the deficit view think that students who don’t have high scores are not capable of complex process like aesthetic reader-response and critical pedagogy, while

students who pass the test are ready for deeper thinking and meaningful reading (Stahl, 1989; Willis & Harris, 2000).

Johnson & Johnson (2002) and Diane Ravitch (2010) link low test-scores to high poverty rather than academic capability. Willis & Harris argue that no single approach to literacy will be the magic bullet for solving the systemic problems that contribute to low-achievement in high-poverty schools.

However, the children who live in poverty, as well as their parents or guardians and teachers, are competent readers of their world, and know that learning to read will not ensure a more equitable or socially just world. They are not swayed by the spin doctors' myths that literacy learning will bring equality; they understand that greater systemic changes are needed in society and education. (p. 83)

In her seminal ethnography, *Ways With Words*, Shirley Brice Heath (2009) suggests that some students come from cultures that use other ways of making meaning rather than reading. In these situations teachers need to form a cultural bridge between the necessities and cultural of home life and literacy. Students may need a deeper justification to be motivated to acquire reading skills than a teacher telling them so.

Aesthetic reader-response and critical pedagogy both focus on the relevance of literature to identity, empowerment, and real life purposes for students who have differing perspectives. Instead of using a deficit model these processes are based in an asset model of education. They focus on the strengths of students' ways of thinking, and seek to build on them, motivating

students to learn skills in the process. Therefore, Willis & Harris strongly support the use of aesthetic response and critical pedagogy as part of creating culturally inclusive and democratic classrooms for diverse student bodies (2000). In order to address this profound controversy, this review critiques studies that explore how students from diverse backgrounds and abilities are able to use aesthetic reader-response and critical pedagogy to develop literacy.

The final controversy is closely related. An idea that is present in society is that elementary students (particularly K-2 students) are not developmentally capable of complex transactions with literature such as aesthetic reader-response and critical pedagogy. Many more studies on these processes exist for middle school and high school students. Others think that educators need to protect children from literature that is unsettling or that has a message that is not fully explained by the author (Hunt, 1991). Aesthetic reader-response and critical pedagogy often use disequilibrating literature to initiate controversial discussions and evolution of identity and purpose in diverse settings. On the other hand, writers like Mike Rose (2006) and Karen Gallas (1994) depict young students vigorously dealing with and debating challenging life issues in context of picture book and poetry study. This review of research also critiques studies including students in grades K-6 to see how they engage in settings that support aesthetic reader-response and critical pedagogy.

As a teacher of literacy, I would like to evaluate the research that has been done so far on aesthetic reader-response and critical pedagogy in elementary schools order to inform my teaching practice. Due to my experiences

and educational training I have a bias for the socio-cultural approach to literacy and the use of aesthetic reader-response and critical pedagogy. But I need to essay my inclinations. This paper represents my attempt to explore these controversies, to critique the credibility and validity of the research, and to gather ideas for classroom teaching practice.

Historical Context

The previously mentioned controversies exist within a socio-cultural and political context that has a long history in the United States. Since the beginning of public education there have been debates about the purpose of public education (Spring, 2011). Aesthetic reader-response and critical pedagogy are aligned with purposes, assumptions, and allies that differ from the historical development of the purposes, assumptions, and allies related to high stakes testing, and “back to basics” ideas today. Choosing how to approach literacy instruction is necessarily a political statement, and all research has bias. In order to explain how culturally controversial aesthetic reader-response and critical pedagogy are in the educational field, this section situates the stance of these processes.

In *The American School*, Joel Spring describes diverse historical perspectives on the purpose of education. In one view the purpose of education is to organize a work force that can maintain America’s competitive power in the world. Inherent in this idea is the need for education to sort the workforce into those who will lead & innovate, those who will manage, and those who will labor. Therefore, students are exposed to norm-referenced tests. This idea is

connected with the idea that there is one true American culture (white and protestant), and that school should teach patriotism to this culture. In practice this means that teachers approach literacy in terms of whether students are right or wrong is ascertaining the message of the author, and that teachers use curriculum that takes the perspective of the one true American culture. School is to orient students of all backgrounds to the opinions of the white, middle class, protestant experts of the past so that they can continue the values of the current order of society (Spring, 2011).

In another view, school is about helping students of different backgrounds gain equal access to jobs. Like the previous view, this perspective would use a back-to-basics remedial approach for students who have received low test scores in order to help them measure up to the standards necessary for college and middle class jobs. This trend also is related to the development of social services in schools. In the past, groups of more affluent citizens provided these services to assuage the protests of the working classes and to reduce crime. Although this perspective takes a stance of equality and service to communities, it still functions on a deficit model: certain students lack an ability, and so schools create remediation (Spring, 2011).

Aesthetic reader-response and critical pedagogy are aligned with a third view of the purpose of education. This view sees education as way to develop the culture and ideas of society. It is sometimes referred to as multicultural education. Multicultural education involves the development of new ideas and a more just culture through the integration of multiple perspectives and assets from

all of America's cultural groups. Spring speaks of an American tradition of marginalized groups of people who self-organized their communities' education in order to advocate for their own cultural voices and access to resources within the United States. He notes that these groups had allies within the dominant culture (white, middle class, protestant), but were not lead by that culture.

Aesthetic reader-response focuses on supporting students to use literature to understand the multiple perspectives of peers and texts; texts are not viewed as authoritative but rather as discourses, therefore, it aligns with the multicultural purpose of education. Instruction that supports aesthetic reader-response sees literature is an opportunity to develop one's identity in empathetic recognition of others' identities. Critical pedagogy goes one step further, students understand themselves in contexts of the larger systems of privilege and power they are part of and they develop their own ideas of what justice would be like. Essentially these processes see education as a venue for students to come up with new integrative ideas and values for society, and the teacher is their ally. It is controversial because it essentially involves changing the current systems of privilege and authority (Johnson, 20006). The debate over the purpose of education (and literacy) comes down to whether education is to maintain the values of the status quo, or to develop the culture's orientation to justice and democracy.

This political history informs an analysis of why the unprecedented swing to intensive skills-based instruction, funded by NCLB, occurred. Stahl (1999) characterizes the swing as a backlash resulting from whole language instruction

that was failing students due to lack of integrated skills instruction. However, NCLB does not advocate for the integration of skills into meaningful literature experiences—it advocates for skills only programs. NCLB facilitated the idea that schools and districts who score low on achievement tests need mandated intensive basic-skills programs, and in many cases restricted the use of literature and complex processes like aesthetic reading and critical pedagogy for those schools (Au & Raphael, 2000). Au & Raphael state the view within the field that it is unethical to strip literacy curriculum in low-achieving, often high-poverty, schools of higher-level meaning-making processes in the interest of teaching only to the skills covered in standardized tests. Willis & Harris (2000) point out that history has a long tradition of elite groups censoring the development of literacy of women and slaves, and equates the skills-based-only approaches used in high-poverty, low-achieving schools with this tradition of providing limited access to the full scope of literacy. Au & Raphael echo this idea when they cite research that shows that low test scores are more an indicator of socio-economic status bias in the test than they are of intelligence. In reality, the backlash may have resulted from all of these views.

The theoretic roots of aesthetic reader-response and critical pedagogy, and those who first defined the terms, adds another lens to interpret the backlash. In the early 1900s John Dewey worked with a new group of progressive educators to dictate and publish children's writings and use them to teach reading (Stahl, 1999). His view of the aim of education was to develop democracy, by facilitating students in developing their strengths in context of a

community of learners (Dewey, 1997). In the 1960s, in response to the desegregation of the Civil Rights Movement, educators used the Language Experience Approach to build literacy on the home language of learners and to chart and discuss literature experiences as a way to develop literacy. Neither of these approaches were in the mainstream of education (Stahl, 1999).

In the 1970s and 80s, Paulo Freire began to write about his work in Brazil using critical pedagogy approaches with illiterate workers to increase their political power. He treated reading as a transaction in which the reader becomes a writer and can rewrite and transform their view of the world through language. In the process readers critically evaluate their experiences of the world and the text in terms of a social empowerment (Freire, 1972). In a military coup that had interests in oppressing the votes of the workers, Freire was imprisoned and exiled (Freire & Horton, 1990). Around the same time, Louise Rosenblatt and others began to define reading in terms of efferent reading and aesthetic reading. Her view of literacy was informed by her background as a daughter of Jewish immigrants and her experiences translating French accounts of the Nazi dictatorship during WWII. She especially emphasized that aesthetic reader-response to literature honors the unique transaction and creation of a new experience between the reader's history and imagination and the text. Rosenblatt said that when readers share such experiences they create valuable opportunities to work across differences and to become self-aware in a process that fuels the progressive development of inclusive societal values (Rosenblatt, 1982).

In the 1980s mainstream education popularized the whole language approach, and in the process whole language was implemented in a variety of ways that were not consistent with its original purpose. Literature-based instruction is an aspect of whole language dealing with literature. However the development of whole language and literature-based instruction rests on the theories of Rosenblatt, Freire and others. Aesthetic reader-response and critical pedagogy are specific kinds of literacy and include specific instructional supports. They may result in a higher quality of literature-based instruction.

Given this history of the development of the theory, it is clear that the founding philosophers did not have views that were in the political mainstream. It suggests that Willis & Harris's assertion of a history of oppression through an incomplete scope of literacy is well founded. Other theorists suggests that, since the majority of teachers are Caucasian and middle class, the oppression is not a conscious effort, but rather a bias against the disruption of traditional authority in society. That is, it is a product of a culture of unbalanced privilege systems and unexamined assumptions (Johnson, 2002).

In the midst of this historical fire, the question still remains as to whether aesthetic reader-response and critical pedagogy do any disservice to students as far as basic-skills acquisition. And the question still remains about what kinds of literacy behaviors students from different backgrounds engage in when teachers facilitate these processes. Is literacy, in its fullest definition, a form of cultural power and a process of cultural change, or is it only decoding & comprehending the message of texts? This review of research explores these questions.

Definitions

One can understand aesthetic reader-response by contrasting it with efferent reading of texts. The purpose of efferent reading is to extract information from the text and arrive at a public take-away message. In contrast, the purpose of aesthetic response is to turn inward during reading and place attention on the personal feelings and values that the text evokes. Reading becomes a lived through experience (including the sounds and images of the language) wherein the reader creates meaning from the transaction of their own past experiences and the world of the text. In reality both efferent and aesthetic stances happen during reading, but as an art form, Rosenblatt suggests that literature should primarily be approached through an aesthetic stance (1982). A second aspect of aesthetic reader-response is related to reading texts within a community of responders. Rosenblatt states, “a collaborative classroom atmosphere may lead to self-awareness and self-criticism, both in the area of the reading process and in the clarification of issues in values” (1985, p. 72). In considering the multiple perspectives of peers, and the experience of the world of the book, students develop empathy, and may integrate other cultural viewpoints into their own narrative of the world (Rosenblatt, 1985). The focus after reading is not on having received a moral or an instruction manual from the authority of the author, but rather on whom the reader has become from the experience.

Critical pedagogy may include an aesthetic process when it involves literature and art, but the primary focus is on understanding oneself in relation to political and social power systems within the world. The intention is social justice

related. Part of the process involves identifying the discourse and culture that the author writes from and how it relates to one's own discourse and culture. Freire describes critical pedagogy as a praxis. The reader starts with their own reading of the world, they read the words of the text—for a time dialoguing with a different perspective on the world, and through the process they leave having re-written the world. But there is an undercurrent of problematizing the current social order and coming back to the problems in one's community with new insights (Freire, 1972). Freire approached teaching literacy by being a self-critical learner, letting the wisdom and views of the community teach him as the community in turn self-criticizes and creates new culture (Freire & Horton, 1990). Freire suggested the need to move beyond a sensual experience of reading, and to include next steps of further research, generalizations, and plans of action for social justice or empowerment in the process—a continued active rewriting of the culture of self and the world (Freire, 1972).

Literature-based instruction is not a process like aesthetic reader-response and critical pedagogy; it is a form of instruction. It may include these processes or it may not. The purpose is to learn reading skills through meaningful experiences of literature, and to be able to use literature for real-life purposes. It usually includes the integration of reading and writing. It includes high quality trade book literature in its entirety, rather than basals formulated to teach skills. Whole language takes the same approach, however it focuses on more than just literature.

Finally, the definition of literacy is highly debatable. A common idea is that literacy is only learning to comprehend and read texts fluently. James Gee offers a broader definition. He says that literate people acquire multiple cultural discourses through language; they also develop a meta-discourse to critique their initial home discourses and the dominant discourses of society. A discourse is the cultural viewpoint, assumptions, and language that a particular group of people uses. In reading literature, a child encounters many discourses. In Lisa Delpit's (2006) terms, students should acquire the codes of power while maintaining home codes. In James Gee's terms, students should also critique the current order. Some educators relate the ability to critique and discuss the meaning of multiple discourses to the phrase, "higher-level thinking." "Higher-level thinking" is a topic of research of the studies in this paper.

Jerome Bruner focuses on literary texts. He suggests that being a literate reader of literature involves creating narratives. Literature has gaps in explanation, paradoxical meanings, and complicated unsettling events; for psychological purposes a reader must make sense of it all. The process of sense-making creates a narrative—this narrative involves a way with words in the world. Literacy involves the act of entering and developing a shared consciousness, culture, and language.

All of these definitions are similar. One question that arises is whether it is useful to integrate aesthetic reader-response and/or critical pedagogy in teaching literacy, and whether the processes improve the quality of the instruction. This review of literature explores that question.

Limitations

This paper does not focus on content literacy or informational reading. It only focuses on the reading of literature. It also does not study the difference between adolescent and adult literacy and elementary school literacy. It only focuses on how young children engage with literature, and the costs and benefits of using literature to approaching literacy.

Due to the nature of the research available, most research on literature-based instruction is conducted from a socio-cultural framework of literacy, not a skills-based framework. In addition the majority of research is qualitative, both because the theory is complex and still being defined, and because some of the literacy abilities related to literature-based instruction are hard to quantify. Therefore it is important to take into account the biased nature of the research, which is one major limitation of this paper.

This paper does not focus on the implementation of skills-based instruction, and aside from a few comparative studies it does not analyze the costs and benefits of that form of instruction. It only focuses on distinctions and outcomes of literature-based instruction.

The paper does very little to explore how skills instruction happens within literature-based instruction, rather it focuses on the kinds of literacy that emerge, some of which might be basic skills. It does, however, contain a section on the influence of incorporating strategy instruction into literature-based instruction and how that relates to skills instruction, but this only to further qualify and define how

the processes of aesthetic reader-response and critical pedagogy are best supported in practice.

Summary

In summary, the NRP (2001) has left out several critical aspects of literacy research in their highly influential report. The report did not study the reading processes of aesthetic reader-response and critical pedagogy. These processes are controversial because of their political stance towards the authority structures of the status quo. Aesthetic reader-response and critical pedagogy are also controversial because they suggest that higher-level literacy skills can be mutually beneficial to technical literacy skills, an idea that directly opposes the dominant view that phonics skills are necessary before higher-level reading. Some educators claim that aesthetic reader-response and critical pedagogy processes may contribute to more inclusive and meaningful literacy experiences for students of all abilities and from all backgrounds. The primary focus of this paper is to review the research on the effects of aesthetic reader-response and critical pedagogy on the kinds of literacy development of diverse elementary school students.

CHAPTER 2—CRITICAL REVIEW OF THE RESEARCH

Introduction

Chapter 1 was a review of the social-historical contexts, and theoretical development, and definitions of aesthetic reader-response and critical pedagogy. It established that aesthetic reader-response and critical pedagogy might be processes that define high-quality literature-based instruction. Chapter 2 contains analyses and critiques of 37 studies on the costs and benefits of literature-based instruction for diverse elementary school students. Some of the studies explicitly relate to aesthetic reader-response or critical pedagogy, while others study the effects of literature-based structures that would support those processes. The chapter uses this mix of perspectives because there is a lack of research available for studies that explicitly relate to aesthetic reader-response and critical pedagogy in elementary school settings. Most of the studies are qualitative and seek to define distinctions in quality literature-based instruction. The critiques are based in Lincoln & Guba's (1985) criteria for credible and valid design.

The sequence of the sections goes from specific to general. In this way, the chapter establishes strong definitions for aesthetic reader-response and critical pedagogy in order to find structures to look for in subsequent studies that would support these processes. Each section groups the studies by similarity of questions, rather than in chronological order. The chapter includes many studies from the 1990s because there was more funding for research into literature-based instruction during that time, and thus more studies are available. The fact

that studies from the 1990s and more recent studies come to similar conclusions justifies this selection and ordering.

The chapter begins with sections explicitly addressing aesthetic reader-response and critical pedagogy; these sections define what elementary students do when they engage in these ways of transacting with literature, the structures that facilitate these processes, and the benefits for students. The analyses of these two sections infer connections between the two processes. Since one major structure that both processes use is discussion, the next section addresses how to implement quality discussions in literature-based settings, and the benefits of discussion for students' literacy. Since the first three sections suggest that motivation and engagement for reading is a major benefit, the fourth section explores studies on motivation. In order to essay whether the structures in the first three chapters coincide with structures recommended by broader studies on literacy, the fifth section addresses best practices studies. The best practices studies suggested that a further qualification of effective literature-based instruction includes the incorporation of explicit strategy instruction, so the sixth section explores this distinction, and the benefits for students. The final section addresses the debate about whether literature-based instruction is beneficial for the basic-skills achievement of all demographics of students.

Aesthetic Reader-Response

This section addresses questions regarding the characteristics of aesthetic reader-response, how it is facilitated, and what kinds of benefits it has for elementary school students. The studies represent students in first through

sixth grade. All of the studies take an ethnographic view, and the majority are action-research studies, as a result the studies only point to areas of further research rather than provide empirical evidence.

This paragraph contains a brief summary of each study that this section will critique in detail. In the first study, Eeds & Wells (1989) conducted an ethnography and found that mixed-ability fifth and sixth grade students used aesthetic reader-response in literature discussion groups to effectively sort out problems in comprehension, and create meaning in context of their peers alternative views. In the second study, Hoffman (2010) conducted a unique case-study of the responses by reviewers, ages eight to 11, to *The BFG*, by Roald Dahl, on a website book review venue and found that children have natural responses to literature that are very similar to responses evoked in reader-response activities at school. Similarly to the first two studies, Parsons (2006) conducted voluntary participant action-research to explore how avid fourth grade readers create, enter, and sustain the world of the story and found that they use a visualization process, and created metaphors to communicate their experience of the story. In the fourth study, Cox & Many (1992) studied responses to a work of literature by fifth graders who were at or above reading level and found that given an open-ended forum for discussion and writing, students achieved the highest levels of understanding as the aesthetic response stance increased. Finally, Lohfink & Loya (2010) studied the influence of culturally relevant picture books on the efferent and aesthetic responses of third grade, bilingual, Mexican American students and found that given a structured open-ended forum, students

focused their responses on aspects of the stories that directly related to their cultural experiences; 60% of their journal responses used an aesthetic stance. As a group, the studies suggest that culturally relevant texts, open-ended discussion, and open-ended writing forums tend to aide aesthetic responses to literature, which in turn deepens comprehension and meaningful uses for reading.

Eeds & Wells (1989) conducted an ethnography and found that mixed-ability fifth and sixth grade students used aesthetic reader-response in literature discussion groups to sort out problems in comprehension, and create meaning in context of their peers' alternative views. After hearing a description of each book students chose which literature circle they wanted to be part of, and the resulting groups were made up of heterogeneous designated ability levels. Fifteen undergraduate students facilitated discussion twice per week for 30 minutes per session, while researchers audio-taped the conversation and kept field notes. Facilitators tried not to ask comprehension questions, but rather elicited the students' responses to the story. Researchers coded themes in the utterances of the transcriptions.

Eeds & Wells (1989) found that the children used the discussions to sort out problems in comprehension, articulate, and revise their perceived meaning in context of their peers' alternative views. In context of comprehending the text they grew in their ability to take risks by sharing their own hypotheses and perspectives. They relate how characters or events in the book made them think or rethink their own experiences, and students often reciprocated each

other's stories with their own. They also used their own and peers' experiences to infer a meaning for parts of the story where the author left unexplained gaps. They developed group inquiry by discussing hypotheses and looking for textual support. They made statements that indicated valuing and evaluating the images, feelings, attitudes, and conclusions they created during a lived-through experience of the literature.

Although Eeds & Wells (1989) study provides many excerpts from the transcriptions to illustrate the thematic findings, it has several problems that harm the credibility of the study. The sample was one of convenience because the researchers had used the classrooms as practicum sights and had relationships with the teachers. They did not describe the process of deriving themes and had no inter-rater reliability, no audit, and no member checking. Because none of these features are present there are no controls for the biases of the researchers. The transferability is also hindered because they provide no details other than the grade levels and reference to ability levels of the students. They provide few examples of what the facilitators did to influence a successful conversation. It is difficult to know which settings this approach would work well in and how to implement it without further research.

None-the-less, Eeds & Wells (1989) study does give an interesting portrait of what the expected outcomes of aesthetic reader-response conversations might look like. It also suggests that group discussion is a structure that supports aesthetic response to literature. Further research should address how these conversations impact the attitudes and measurable literacy abilities of students.

Hoffman (2010) conducted a unique case-study of the responses by reviewers, ages eight to 11, to *The BFG*, by Roald Dahl, on a website book review venue and found that children have natural responses to literature that are very similar to responses evoked in reader-response activities at school. Students could see other children's reviews and structure their review in response to the postings of their peers. With a subscription fee, the site provides a venue for hundreds of schools to use in language arts classes. Hoffman chose the book because it was controversial and continued previous research of children's responses to the book. She surveyed 30 reviews to gain insight as to what children desire to gain from literature, the aspects of the literature they responded to, how they described their experience, and the way they present themselves as readers.

Hoffman (2010) found that children compared and contrasted their own experiences of the book with the views of their peers. They also addressed their reviews using the second person, "you," invited readers into the experience of book that many had partaken in, and described the kind of literary experience "you" could expect from the book. They provided a partial summary of the story, with respect to not giving away the plot. They explicated their feelings and imagined world of the story through making and describing their own illustrations, which often contrasted Quentin Blake's illustrations. Some reviewers tried to describe a lesson for their life that they got from reading the book, or a character they identified with. Children made many comments about how they felt in response to the language and structure of the writing as a purely aesthetic

quality. They also commented on a desire to get better at writing so they could describe their experience more clearly.

The method of Hoffman's (2010) study lacks some criteria of credibility, but the vantage point of the researcher (having no contact with the subjects or control over the framework of their responses) adds to the body of research on aesthetic reader-response. Hoffman did not describe how the themes she elaborates emerged. She uses no inter-rater or audit. These facts harm the credibility and confirmability of the study. However, she does give convincing examples to illustrate her findings, meaning that the themes she names do seem to be present, even if there was room to omit the prevalence of other trends. The transferability is only for students who are motivated to post on the website and are in a school that can afford a subscription. Aside from saying that the participants came from diverse demographics, Hoffman provides no other information on the subjects.

Overall, Hoffman's (2010) study seems to show that children have natural responses to literature that are very similar to responses evoked in reader-response activities at school which they seek to share. This suggests that aesthetic reader-response in the classroom helps initiate students into authentic and natural ways of using books in their lives. The study is supported by other studies that find children relating their own experiences in context of others' experiences as an engaging and meaningful part of reading. It suggests that children choose stories as a ground for shared experiences. It also suggests that part of a child's experience relates to the aesthetic of the writing, and its ability to

invoke images that the child can create in illustrations. It confirms studies that suggest that writing is an essential part of reading. Further studies might focus on the role art and writing play in the transaction with literature, the development of meaning, comprehension, or perception of self as reader.

Parsons (2006) conducted voluntary participant action-research to explore how avid fourth grade readers create, enter and sustain the world of the story. She found that they use a visualization process, and created metaphors to communicate their experience of the story. She sought to compare how children relate their reading process to see if it matched Rosenblatt's theory of transactional reader-response. Parsons was the fourth grade teacher at a mid-western public school, and focused her study within a voluntary "Readers as Researchers Club" that met outside of class time. She collected multiple data points through the children's writing, coding, memory work, discussion, and visual renderings of stories. The children recalled and wrote stories about their experiences with meaningful reading experiences and then they debated, coded, and created graphic organizers to display themes that they all shared. Next they conducted a book group. During the book group they focused on a reading process theme that they all shared, and further defined it. Parsons reports all of the discussions and writings via transcriptions to illustrate the themes the children created.

Parson (2006) and the participants found that avid readers engage in literature in six ways: "being in the book, connecting to characters, experiencing emotions and physical reactions, wondering and predicting, visualizing, and

turning out” (p. 494). During the book club they focused on further defining the visualizing process, as it was the primary mode that they communicated their experience of living through the story. They found that there were three modes of visualizing that represented different perspectives children took while experiencing the story. Picturing takes place as the child enters the story, or when a discrepant event challenges their understanding of the story or their personal experience. It is a still image in the mind that the child can use to think through their understanding of the story. Watching takes place when a reader shadows a character or watches from a particular perspective. Seeing involves a birds eye view as children see the whole of the story and can move in and out of characters as they live the story; they can choose which events to engage in the most. As children go through the visualizing process they use metaphors from their own lives to describe and illustrate how they have experienced events or characters in the story.

Parson’s (2006) method rests on member checking and negotiation among the participants. Because there were no additional researchers documenting the children’s findings there is room for Parson’s bias to editorialize the children’s codes. However she provides many examples of the discussions and student work that illustrate the codes, and she supplies the language the children used, which corresponds with the academic language Parson uses in her report. The transferability tends to be for a small group of highly motivated readers of an undefined community demographic.

Parson's (2006) study offers a unique vantage point; the data comes from the students' own descriptions of their reading processes. The descriptions are very similar to the processes described by Rosenblatt (1978) and Eeds & Wells (1989). This study adds the implication that visualization is a very useful way for children to enter into texts and communicate experiences, as well as understand different perspectives their classmates hold. It suggests that students who use the aesthetic reader-response process are highly engaged, thoughtful readers, who own their own reading process. It is yet unclear what this means for achievement or long-term cognitive processes.

Cox & Many (1992) studied responses to a work of literature by fifth graders who were at or above reading level in order to evaluate them on a continuum of efferent to aesthetic responses as defined by Louise Rosenblatt (1978). They found that given an open-ended forum for discussion and writing, students achieved the highest levels of understanding as the aesthetic response stance increased. The subjects were 38 students who had all scored on reading level or above on the California Achievement Test. Subjects read nine works of literature chosen for quality, and responded in written journals with an open ended prompt to write anything they wanted about the book. Cox & Many analyzed the written responses in order to observe the natural process that the children went through in order to organize responses to a work of literature. They separately analyzed the data through grounded theories of transactional response, types of mental activities involved in response, and the level of

understanding the student achieved. They had an auditor rate the journals as well.

Cox and Many (1992) found that of the 11% of students whose responses were primarily efferent. The efferent responses included naming a message of the author, an analysis of literary structures, an evaluation of the quality of the work, and how they thought it fit into the historical and social contexts of present life. Only 9% of the total responses focused on retelling, which indicates that upper elementary students tend to go beyond a basic comprehension of the story. Students wrote 42% of the responses from a primarily aesthetic stance. These responses included a particular character or event that attracted their attention, and their feelings or judgments about what happened in the event based on personal values or experiences. 20% of students' responses fell into the final category, "most aesthetic response." These responses included describing a moving image of the story, and how they lived through the experience, the feelings evoked, the lingering questions, and the recollections of empathy. They connected detailed accounts of their own lives as a way of trying to understand the motives of characters, making analogies to their world in the process. They noted experiences of characters that expanded their view of possible experiences in the world. They also created alternative endings and explained why the new endings would improve the story. Students achieved the highest levels of personal understanding as the aesthetic response stance increased. Cox & Many define the highest level of understanding as making analogies from inferences, and generalizing about life through the story. Lower

levels of understanding include literal interpretation, retelling, and statement of judgment without explanation.

Cox & Many (1992) make some moves to support credibility, dependability, and confirmability, but do not provide enough information on the subjects for good transferability. They describe how their coding framework was consistent with grounded theories, and how the coding procedures proceeded, giving convincing examples. They establish good inter-rater reliability with each other and the outside auditor. Finally, they name many previous studies that corroborate or contradict their research, which makes for a dependable study. The study is transferable to students who score high on a norm-referenced test of achievement (probably related to basic comprehension) but it is unclear what sort of school and community setting this took place in, or the kind of instruction the teacher used. The last fact means that the response may or may not have been naturalistic, because the teacher may have given encouragement of aesthetic response in the course of teaching.

Overall, the study is consistent with the findings of Eeds & Wells (1989) study, Hoffman's (2010) study, and Parson's (2006) study—students tend to respond aesthetically to texts when given an open-ended forum for discussion or writing. This suggests that aesthetic reader-response may be a natural process to arriving at a complex understanding of the text. It also suggests personal ownership of the reading process. One could make the inference that if teaching primarily focuses on closed efferent questions, then students may never develop deep understanding of the literary work or ownership of their own reading

process. Further studies should address what this means for academic achievement, for long-term cognition, for long-term readership, self-efficacy, cultural advocacy, and establishing a role in society.

Lohfink & Loya (2010) studied the influence of culturally relevant picture books on the efferent and aesthetic responses of third grade, bilingual, Mexican American students. They found that given a structured open-ended forum, students focused their responses on aspects of the stories that directly related to their cultural experiences; 60% of their journal responses used an aesthetic stance. The students were in a public school classroom in rural Kansas. Twenty of the students in the class were Mexican American and bilingual; half were female and half were male. The school received bilingual programs and equity and access programs, but third grade was conducted mainly in English. The teacher noted that the students were all reading at or above grade level. The teacher (Loya) was Mexican American and had been teaching for 12 years in the community. She had experience at choosing books that would match the funds of knowledge of her students.

Loya & Lohfink (2010) engaged the children in three stages of engagement with literature. Before reading, during reading, and after reading. Before reading, children discussed artifacts or stories from their life as they examined the synopsis, cover, and author descriptions. During reading the researcher used two open-ended prompts regarding what students noticed and what it reminded them of. They were also encouraged to make predictions and ask questions. Following reading Loya provided an open forum for children to

share about their experience of the book. They had time to recursively write, draw, and talk, and were encouraged to collaborate to make meaning. Lohfink & Loya (2010) took field notes, videos, made transcriptions, collected student literature journals, and conducted an interview with Loya. Following established research procedures they coded the individual units (intonation or semantic closure of an idea) of students' responses. They examined where the students' stances fell on the efferent to aesthetic continuum and which aspects of the story cued their responses.

They found that 53% of the 970 units involved efferent connections to the text, while 47% has to do with the aesthetic experience of the story. In contrast, 60% of journal units took a primarily aesthetic stance, and 40% efferent. Six foci emerged for the aspects of the literature that cued each child's attention; three of the foci, language, customs, and immigration events, related directly to cultural experiences of the students. The students focused on these cues in both their aesthetic and efferent responses.

Lohfink & Loya (2010) clearly describe and illustrate how they triangulated many data sources and used established coding processes. This establishes credibility, but credibility is also compromised by not describing how access was gained, not naming previous instruction, lack of an inter-rater, and lack of an auditor. None-the-less the ample examples make the study's findings convincing. The transferability is strong for the demographics of the students, but it is unclear how their previous instruction impacted their approach to texts, or how it effected their achievement of skills.

Lohfink & Loya's (2010) study adds another layer to understanding how literacy develops. Students were very engaged, and used their own experiences to understand and experience the text when it matched experiences in their own lives. This suggests that aesthetic reader-response, and literacy development benefit from cultural relevancy. Some advocates of aesthetic reader-response argue that literature is a way to understand experiences very different than one's own as well. This study seems to say that starting with culturally relevant texts may be a productive way to scaffold students into an active meaning-making approach to books. This study also suggests that there may be a cross-over between aesthetic reader-response and critical pedagogy when the issues of culture or power become a central focus of the experience.

In conclusion, the studies in the Aesthetic Reader-Response section all came to similar findings, which strengthens the dependability of all studies. Key insights include the idea that children naturally respond to texts aesthetically when given an open forum. Aesthetic response involves creating an image of a lived through experience; empathy; identifying values; inferring from one's own and others' experiences; asking questions; creating metaphors; and using discussion, art and writing to try to understand disequilibrating story events. These studies found that the process of aesthetic reader-response is a natural way to deepen cognitive processes with reading, generate deeper understanding, take ownership over a meaning-making process, and understand one's self in context of a diverse community. Teachers, websites, and researchers facilitated aesthetic reader-response through open-ended

discussions, writing, art, visualization processes, student choice, student as researcher, culturally relevant literature, as well as quality literature that challenged students ability to empathize. The limitations of these studies include that they do not take necessary measures to establish strong credibility, and the majority of the studies are highly subjective qualitative accounts from teacher action-research. Aside from Lohfink & Loya's (2010) study, the researchers do not sufficiently describe the demographics of the students; this hinders transferability. The value of the studies is that they help define aesthetic reader-response and directions for future research, such as effects on achievement and long-term use of literacy.

Critical Pedagogy

The studies in this section seek to define how critical pedagogy is enacted in elementary school students' literature study, how it is best implemented, and the benefits for students. The studies explore the connection between social justice, awareness of power systems, and literacy. In addition, this section explores the relationship between aesthetic reader-response and critical pedagogy. Most studies find that these two processes are intimately connected. These studies represent multiple age groups and multiple student demographics. All of the studies are ethnographic action-research, and do not represent empirical evidence, or address basic skills achievement. However, they do point to directions for further research.

As a whole, the studies in this section suggest that aesthetic reader-response is a jumping off point for critical pedagogy. The studies suggest that

critical pedagogy in diverse student communities inspires ownership of the reading process as students work together to plan social action. This paragraph contains a brief summary of each study that the section critiques in detail. In the first study, Goldstein (1995) created an action-research and narrative study with low income first and second grade Latino students, designated as LEP, and focused on aesthetic response as a ground for critical pedagogy; Goldstein found that students used culturally relevant texts to expand their social awareness and then extended their experience of the literature into content studies and social action projects characteristic of critical pedagogy. Suggesting that writing is a vehicle for a critical pedagogy, Flint & Laman (2012) conducted an ethnography of second and third graders from multilingual, ELL, high poverty, transitory, and inner-city backgrounds to explore how they responded to integrated reading and writing workshops that focused on multilingual social-justice poetry; they found that literature and writing served as a ground to expand personal experiences, speak up about privilege systems, and learn aesthetic skills in poetry writing. In the third study, Kersten (2006) did action-research with fourth grade students who were bilingual in Cantonese or Spanish and found that critical pedagogy developed as students responded to a culturally encapsulated history textbook; the students pursued alternate accounts that represented their own cultural perspectives and hypothesized about the connection between historical biases and cultural exclusion. Similarly, Souto-Manning (2009) conducted ethnographic action-research with first and second grade students from diverse linguistic, economic, racial, and sexual backgrounds and found that after an introduction

activity to multiple perspectives in literature, students extended multicultural literary experiences to advocating for inclusive inquiry study in their school. Finally, Arce (2000) conducted an ethnography to understand how critical pedagogy could be applied in a culturally, economically, and linguistically diverse first grade classroom; she found that activities that called forth the students own perspectives in relation to social studies and literature resulted in empathy within the community, ownership of the reading process, and civic voice. Next, this section examines these studies in detail.

Goldstein (1995) created an action-research and narrative study with low income first and second grade Latino students, designated as LEP, and focused on aesthetic response as a ground for critical pedagogy; Goldstein found that students used culturally relevant texts to expand their social awareness and then they extended their experience of the literature into content studies and social action projects characteristic of critical pedagogy. Students were in first and second grade in a Californian school that supported bilingual education. Eighty-five percent of students at the school were Latino, 12% African American, 2% Caucasian, and 1% Asian/ Pacific Islander. The school qualified for Chapter 1 funding and most students received free or reduced lunch. Goldstein and her aide were bicultural and bilingual, and facilitated three girls and eight boys in a bilingual program.

Goldstein (1995) taught in Spanish and English. The language arts program was arranged in democratically decided units and integrated content area studies. The literature took the central role of each unit and the students

and teacher chose books based on cultural and social themes that related to students lives. The literature they chose was not structured as having one single message, but rather left room for interpretation and questioning; students had to derive a meaning for themselves. Instruction consisted of shared reading, interactive journals, individual reading, group discussion, and using the writing process to publish original work. Goldstein reports the findings in narrative form, and then summarizes emerging themes.

Goldstein (1995) found that students expanded social awareness through experiencing the central problems of the characters or the society of the book, such as: homelessness; discrimination due to skin color or immigration status; lack of access to doctors and the resulting use of cultural remedies; mistreatment of farm workers; limited access to jobs; and bureaucratic roadblocks to memorializing their culture in their neighborhood. Because the program integrated literature into content areas students came up with plans to extend their literary experiences into finding ways to express their cultures and create social justice in their community. Such projects included: murals, interviews, job research, plant study, letter writing to the city council and more. In the course of this study, students used charts to create vocabulary banks, references for sound/ symbol patterns, and visuals of the structure of the stories. They used the charts as they created illustrations and captions of their experiences.

Narrative action-research studies tend to lack credibility by research standards, but do provide descriptive information that can be used in practice. One way Goldstein (1995) could have increased the credibility of the study would

have been to have more than one person writing narratives in order to make inter-rater comparisons. She also could have collected artifacts and transcriptions for evaluation by an outside auditor. An aspect of the study that does lend credibility is that she reports mostly factual information, such as topics of comments and resulting activities, rather than her own inferences, which she includes separately. This distinction allows for the separation of her bias from the observation of occurrences and makes the study more useful. Additionally, she supplies good information about the subjects, which aids transferability. Clearly the environment of a public school that supported small class sizes, use of aides, and bilingual education all contributed to the work Goldstein was able to do; these advantages are not present in all public schools.

Goldstein's (1995) study is very interesting in context of the central question of this paper. It suggests that when the teacher included culturally relevant texts, depicting social issues the students faced, that aesthetic response allowed them to articulate their experiences of the literature and related life issues, which in turn inspired critical research of the problems in their community. This bridged over into social justice work through content area studies. Aesthetic reader-response to literature may be an effective starting place for critical pedagogy to emerge, especially if it includes a focus on societal issues that students can learn to notice through the experience of their empathy with characters in the book. It also shows a clear connection to the development of ownership of the reading process, cultural advocacy, group efficacy, and a role in society. The integration of skills work suggests that the students grew in

achievement in context a meaningful pursuit, but further studies will need to focus on how this process effected their achievement.

Flint & Laman (2012) conducted an ethnography of second and third graders from multilingual, ELL, high poverty, transitory, and inner-city backgrounds to explore how they responded to integrated reading and writing workshops that focused on multilingual social-justice poetry; they found that literature and writing served as a ground to expand personal experiences, speak up about privilege systems, and learn aesthetic skills in poetry writing. The study took place in two schools (eight classrooms) in the Southeast. One school, located in a mid-sized city, had large numbers of ELL students representing 14 languages, and 55% of students qualified for free or reduced lunch. The second school was located in a transitory neighborhood in large inner-city setting; 98% of students qualified for free or reduced lunch, and 57% of students were designated as ELL.

Flint & Laman (2012) had been in the process of training teachers in writing workshops for four years. The workshops included reading and discussing multilingual multicultural poetry related to social justice issues. They also read non-fiction texts, problematizing issues connected to privilege and power. They used the writing process to create poetry. The study took place for one school year. Flint & Laman triangulated audio and video recorded poetry lessons, writing conferences, share times, writing samples, and interviews of the students and the teachers.

They found that students' use of literacy corroborated previously established theory on four aspects of critical pedagogy. Students use reading and writing to disrupt the common place, interrogate multiple perspectives, focus on sociopolitical issues, and to promote social justice (p. 15). They wrote poetry that intersected the texts with their own experiences. The researchers infer that the aesthetic space of poetry created a safe place where students could express feelings, experiences, and ideas that were risky or disruptive to silence about systems of power that impacted their lives. Teachers and students remarked that they were willing to speak up more in the process of discussing the images, sounds, and voices that emerged in their poetry.

Flint & Laman (2012) do not describe the way they gained access to the subjects or the data analysis procedures they used, and they provide no inter-rater statistics, or outside audit. This harms the credibility of the study. However they do make moves to establish some credibility through use of examples from the transcriptions and student interviews. The transcription excerpts illustrate the aspects of critical pedagogy they were looking for, based on prior research. The interviews served as a way to member-check their inferences. They provide enough evidence to make their findings convincing, even if biases may have allowed them to omit other themes. The transferability is strong due to detailed descriptions of the demographics of the students and communities.

Like Goldstein's (1995) study, Flint & Laman (2012) suggest an intersection between aesthetic process and critical pedagogy. They produce findings that suggest that the appeal of the aesthetics created a ground where a

socially critical voice could emerge. In addition, similarly to Lohfink & Loya (2010), the use of multilingual and multicultural poetry, as well as culturally relevant non-fiction texts inspired engagement and real uses for literacy. Clearly the students gained skills in understanding the elements and structures of poetry, and they had ample time to develop writing skills. Further studies should explore how such workshops effect achievement, long-term engagement with literacy, and how students grow to take a conscious role in society.

Kersten (2006) did action-research with fourth grade students who were bilingual in Cantonese or Spanish and found that critical pedagogy developed as students responded to a culturally encapsulated history textbook; the students pursued alternate accounts that represented their own cultural perspectives and hypothesized about the connection between historical biases and cultural exclusion. Although this study does not include literature study, it does add to a vision of how students engage in critical pedagogy in a reading context; the findings could be applied to literature study. Ninety percent of the fourth grade students had a first language of Cantonese or Spanish, and also spoke fluent English. They were situated in a Californian classroom, in a school that supported multilingual education, and had teachers of the same culture and communities as the students.

Kersten (2006) kept records of the students' study of history. She focused her findings on how students named and criticized the cultural stance of their social studies textbook, and sought out alternative histories that coincided with their own cultures and experiences. She discussed her adoption of critical

pedagogy as a result of the critical direction students naturally took in their studies.

Kersten (2006) found that students initially approached the text as an authority, not to be questioned, as was typical of their educational conditioning in previous years. However, when one student posed the question, “Why’s everyone white?” in relation to a picture in a section entitled “Understanding Culture,” the teacher understood that a critical stance might be a beneficial form of study for her students. With an open questioning format they were eventually willing to voice their criticism of the California history textbook, which took an exclusively Caucasian stance. Students aligned themselves with groups that were held in a critical light, or underrepresented in the text. This personal alignment led to extended research and finding other texts that gave version of history from their own cultures’ perspective. They analyzed why the textbook only represented a perspective of those most privileged in U.S. society while excluding other perspectives. The development of the study largely centered around group discussion that focused on raising questions. The questions led to critique of historical biases.

Kersten’s (2006) study is a personal narrative account of her findings as a teacher. She makes no moves to include an additional narrator or outside auditor, or to establish a coding procedure. Therefore the study is informative, but does not have strong credibility. The inclusion of illustrative conversations does help the credibility, but because Kersten does not say whether or not these are direct transcriptions the examples also are questionable. The study is highly

transferable for a very diverse setting that includes established cultural advocacy. It does not name the income of the students or the community at large.

Kersten's (2006) study suggests that students engaged in more complex cognitive process when they were allowed to bring their own cultural and political experiences, perspectives, and questions to the text. This was largely spurred by naming which groups they related to, why they related, and how their group's treatment in the textbook made them feel. The evaluation of bias in the textbook may have supported the development of students' cultural identity, and articulation of a stance in society at large. These same processes may occur in literature study. As in the previous studies, connection of personal experiences (as is central to aesthetic response) instigated critical pedagogy. Although it is clear that higher-order cognitive processes occurred, future research will need to study how these processes affected achievement scores as a way of addressing the controversy over the use of critical pedagogy.

Souto-Manning (2009) conducted ethnographic action-research with first and second grade students from diverse linguistic, economic, racial, and sexual backgrounds and found that after an introduction activity to multiple perspectives in literature, students extended multicultural literary experiences to advocating for inclusive inquiry study in their school. Students came from living in projects, and living in extremely affluent communities. They had vastly different family structures that often conflicted with the middle-class culture of the school. The school had federally mandated curriculum as a result of NCLB. The students experienced gifted programs and remedial pull-out programs.

Souto-Manning (2009) carried out the study for two years with the same students. With the help of an aide she collected audio-recordings, transcriptions, field notes, and artifacts of student work. In response to the children's concerns about inequality in their school and community, she focused her research on the development of critical pedagogy. Souto-Manning used the constant comparative method to analyze the data and develop themes. The main theme focused on group literature discussions. She developed and revised theories to summarize the data. Instruction included use of multicultural literature, media reports, and the civil rights sections of the textbook. Students were positioned as active agents, and encouraged to compare their own perspectives and peers' perspectives to the power codes of the texts.

Souto-Manning (2009) found that students had initial socialization to a right or wrong perspective of literature, and that this inhibited collaborative group work that honored multiple perspectives. After reading several versions of *The Three Little Pigs*, students expressed ideas about how different people can have different pictures in their mind or use different sounding language for the same story. This developed a framework for further study. The teacher moved into multicultural literature based in civil rights themes. In the conversations, students expressed empathy with the characters, and eventually began to associate the pull-out services in their school with the experiences of segregation they encountered in the stories. This led the students to carry out an investigation of who was included and excluded from different pull out programs, and how students felt in each program. The students then decided to advocate for

inclusive inquiry-based education for all students. The principal, parents, and teacher decided to test the students' idea the following school year. The teacher mentions continued correspondence with students about how the experience of the two years continues to influence their approach to school and society, but does not provide documentation.

Souto-Manning (2009) used triangulation of data, clear descriptions of their data analysis, and provided examples from the transcripts of student talk. Student, parent and administrator interviews provided member-checking. All of this provides some credibility, despite the lack of inter-raters and auditors. The transferability is for a classroom of diverse students with special needs and gifted needs in an inclusion setting. The school climate also resembles schools that are subject to federal mandates.

Like the previous studies in this section, Souto-Manning's (2009) study came to similar findings concerning the relationship between aesthetic reader-response and critical pedagogy processes. In Souto-Manning's study, the aesthetic response to literature seemed to act as a jumping off point for social action and an examination of the privilege systems the students experienced. The teacher facilitated the process by orienting the students to the value of understanding multiple perspectives of a story. The findings suggest that students were able to use their reading to explore inequities in different children's experiences of school. This led to a plan of action to improve their school. It also suggests a long-term effect on student reading and social engagement.

Arce (2000) conducted an ethnography to understand how critical pedagogy could be applied in a culturally, economically, and linguistically diverse first grade classroom; she found that activities that called forth the students own perspectives in relation to social studies and literature resulted in empathy within the community, ownership of the reading process, and civic voice. The subjects were 25 first grade students in a bilingual school. Of these 25 students, 15 were first generation U.S. citizens of Latino descent; of this group the parents worked in service jobs or at home, and lived in the neighborhood that the school served. Five other Latino children were third or fourth generation citizens, and their parents all had college education; the parents chose to send their children to the school and lived in adjacent neighborhoods. The five other students were affluent and of Caucasian decent; their parents chose to send them to the bilingual program and they lived in suburbs outside the city.

Throughout her study, Arce (2000) worked with a teacher, who had no experience with critical pedagogy, and Arce gradually became the facilitator of the literacy activities of the students and teacher. She triangulated field notes, student portfolios, assessments, teacher collaborations, transcriptions, and records of the literacy activities.

Arce (2000) found that integrating social studies with language arts, and culturally relevant literature, was a productive ground for critical pedagogy. Students participated in a unit of study on shelter and what home means from different perspectives. After sharing and depicting their own experiences, they delved into non-fiction texts related to global shelter. In the process they also

read and responded to literary works. The literature studies allowed the children to imagine experiences of shelter other than their own. Discussions throughout the unit allowed children to articulate various value statements regarding human rights and unbalanced power structures. The children crystallized their emerging ideas and feelings by dramatically enacting a community of diverse shelters and ideas of home.

Arce's (2000) findings are presented in narrative form with some examples from transcriptions, and multiple factual observations of the progression of activities and topics of student work. The credibility is reduced due to lack of description of the data analysis, lack of an inter-rater, and lack of an auditor. The transferability is strong due to detailed descriptions of the subjects and the communities they came from.

Arce's (2000) findings concur with the previous studies of this section regarding interplay between the empathy students engaged in through literature, and the work to cultivate and articulate values regarding issue of power and social justice. Her study adds the idea that dramatic reenactment is a useful way to help students experience the lives of characters in literature and crystallize the findings of their efferent research. It is clear that the instruction resulted in collaborative skills, ownership of reading processes, and civic voice, but more research should be done on how such an environment supports language development, skills, strategies, and achievement for the different demographics of students.

In conclusion, the studies in the Critical Pedagogy section have dependability because they come to similar findings. Students develop cultural empathy and self discovery through aesthetic readings of literature. This can then lead to further studies using a process of critical pedagogy. In critical pedagogy students critique and compare discourses in texts and life. In some cases this leads to social justice action. Teachers facilitated critical pedagogy through open-ended discussions, writing venues, integration of non-fiction research, multicultural texts, aesthetic response, role-playing, and community action projects. Students developed the ability to evaluate and critique, as well as respond to oppression. They took ownership of literacy as an important way to develop identity. The limitations of these studies are that they consist of subjective accounts by teachers rather than empirical evidence. They do not focus on skills achievement results. Further research should address how critical pedagogy affects achievement and long-term engagement with literacy.

Group Discussion

In the previous two sections, researchers found that group discussion was a structure teachers used to facilitate aesthetic reader-response and critical pedagogy. This section focuses on defining quality literature-based discussions in order to gain insights for successful facilitation of aesthetic reader-response and critical pedagogy. Some studies address the effects of different kinds of discussion on achievement measures. One study is quantitative, while the others are qualitative. The qualitative studies in this section have more credibility than the action-research studies of the previous two sections; they also confirm

the findings of those sections. In addition, the studies in this section are transferable to a diverse range of student demographics.

A summary of the studies in this section follows. In an inclusion classroom, Berry & Englert (2005) studied the structures that positively influence engagement in group discussion of literature for African American, Hispanic and Caucasian K-2 students from low-income inner-city neighborhood; they found that posted expectations, posted guidelines, regular routines, goal setting, reduced teacher-talk, and student selected topics were all beneficial components. In the second study, Chinn, Anderson & Waggoner (2001) studied the effects of recitation discussions versus collaborative reasoning on the discourse patterns of literature discussions in two socio-economically diverse suburban/ rural fourth grade classrooms; they found that students engaged in collaborative reasoning had significantly more elaborations, predictions, evidence, and alternative perspectives. In the third study, Maloch (2002) researched the relationship between the teacher's role and students' participation in literature discussion groups, in a mixed ability, inclusive, and culturally and economically diverse third grade inner-city classroom, and found that students were able to make the transition to deep student-led discussion of literature when the teacher limited her comments to explicit coaching in the expectations, culture, and language of group discussion. In the fourth study, Goatley (1997) observed the effects of literature-based response groups on racially diverse upper elementary students receiving special education services for LD and EMI and found that after six weeks of strategy instruction, journal writing, and coaching in

book club discussions, students responded to each other's ideas and used discussion to elaborate their journals. Finally, Goatley, Brock & Raphael (1995) conducted an ethnographic study about the experiences of fifth-grade students with learning disabilities, Chapter 1 status, or ESL in a literature-based general education classroom, and found that students naturally took on group roles and shared comprehension strategies; the diversity of the group made the students question more deeply as they worked to understand each other.

Berry & Englert (2005) studied the structures that positively influence engagement in group discussion of literature for African American, Hispanic and Caucasian K-2 students from a low-income inner-city neighborhood; they found that posted expectations, posted guidelines, regular routines, goal setting, reduced teacher-talk, and student selected topics were all beneficial components. The study took place in the Midwest. The subjects were seventeen K-2 students in a multi-age inclusion classroom. The special education students included six students with language delays, five classified as educable mentally impaired, one student with hearing impairment, and ten who were non-disabled. Thirteen were not reading yet, and two were reading at first grade level. Thirteen students were African American, one was Hispanic, seven were Caucasian and two had unknown racial identity. Both the teacher and the special education assistant were Caucasian.

Berry & Englert (2005) conducted a qualitative case study. During the teacher's first year of initiating literature discussion groups, Berry & Englert videotaped and transcribed student discussions in January and June and used

Have's (1998) conversational analytics techniques to interpret the data. They also audio-taped and transcribed teacher commentary about the discussion video-tapes.

Berry & Englert (2005) found that classroom discussion progressed developmentally for the teacher and her students. The teacher set up and adjusted her scaffolding of discussion as the year went on. Each day she would remind the students of the guidelines for body language, turn taking, and sentence starters for expressing agreement or disagreement. She worked with students to develop and post a reference list of the kinds of interpretation students used: text-based responses, reader-based responses, and comments about the author/ style. She read the book aloud, then students and teacher brainstormed a list of topics, and she posted the topics for reference. She kept a consistent routine of teacher-read-aloud followed by discussion of previous goals, then 15 minute discussion (80% lead by students), and a reflection with new goal setting for next time.

One problem that emerged, Berry & Englert (2005) observed, was that students used the topic list to linearly proceed through topics without coherence between group members and with lack of depth. The teacher addressed this by asking student to help each other provide evidence for their ideas from the text, or say their idea again in a different way. She took up the role of paraphrasing students statements, asking others if they agreed or disagreed, and asking them to explain their reasoning as a way of positioning students as authoritative. She worked to reduce her teacher talk from 60% to 40% with all of it focused on

paraphrasing and eliciting coherent comments between students. The teacher removed the topic list, and students began to come up with topics in response to previous topics. Students began to have sustained conversations, such as having 22 turns per topic. By June most students were in the habit of responding, explaining their thinking, and eliciting responses from others. Out of the students receiving special education services, 66% participated regularly; 86% of students with no labeled disability participated regularly.

Berry & Englert (2005) did not describe how they gained entry to the subjects and why they selected the specific group of 17 students. They also did not have an outside audit, which harms the confirmability. Aside from these concerns, the detailed analysis process, examples from transcriptions, triangulation, and member-checking support the credibility of the study. It also concurs with other studies on quality discussion of literature. The study is transferable due to detailed descriptions of each student and detailed descriptions of the strategies the teacher used to scaffold engagement in an inclusion classroom.

Berry & Englert (2005) did not give insight into how responding to multiple perspectives influenced the students' progress in becoming independent readers. However, their study suggests very specific ways to make discussion work. Within the description, students use their lives, their interaction with the text, and their classmates' interaction with the text to learn vital social skills and to build confidence in using literature to reason and think deeply. These are exactly the qualities that aesthetic reader-response, theoretically, cultivates. Their

conversations enable to them to adjust and justify their view of the world, which could easily support the goals of critical pedagogy. Discussion structures, such as those named in this study, may be vital to effective facilitation of aesthetic reader-response and critical pedagogy.

Chinn, Anderson & Waggoner (2001) studied the effects of recitation discussions versus collaborative reasoning on the discourse patterns of literature discussions in two socio-economically diverse suburban/ rural fourth grade classrooms; they found that students engaged in collaborative reasoning had significantly more elaborations, predictions, evidence, and alternative perspectives. Recitation is teacher-controlled and focused primarily on the teacher assessing comprehension, and expecting students to arrive at a “correct” predetermined interpretation; the structure primarily uses conversation between the teacher and the students he or she calls on. Collaborative reasoning involves both aesthetic and analytic response to literature with an emphasis on sensitivity to multiple perspectives of students. It encourages student debates using textual references. Turn-taking is open, there are no predetermined interpretations, and both the teacher and students control the topic during sustained student to student discussion. Students will exit the discussion having developed their own meaning from the text with the help of peers (p. 383).

The subjects of the Chinn and Waggoner’s (2001) study were from two communities. Two classrooms were rural fourth grade students from farms or communities of 1000 or fewer residents with lower income. Two other schools were primarily MSES students from a parochial school in a small city. All

students were Caucasian. The schools were undergoing a switch from recitation to collaborative reasoning discussions as coached by the researchers. Trade books and integration of reading and writing were part of both forms of instruction. The total of 10 small discussion groups also used ability grouping in the formation of the groups. In this 7-week quasi-experimental study, researchers videotaped two initial recitation discussions and compared them to two collaborative reasoning discussions at the end of the study. They coded the discussions for the features of turn-taking, type of question, and cognitive processes. They used t-tests with Pillia's trace to determine significant differences followed by f-statistics to determine which variables were significantly different between groups.

Chinn, Anderson & Waggoner (2001) found that there were significant differences between the recitation discussions and the collaborative reasoning discussions on all codes. F-statistics revealed that students in collaborative reasoning (CR) discussions made 49% more elaborations of text per minute, $SD = .61$, $F = 15.32$, $p < .01$. CR students made 70% more predictions per minute, $SD = .54$, $F = 12.97$, $p < .01$. They provided textual evidence 54% more often, $SD = .74$, $F = 51.80$, $p < .001$. And they articulated alternative perspectives 113% more often than recitation groups, $SD = .91$, $F = 12.50$, $p < .01$. Following the trend of high f-values at high significance and minimal standard deviation, researchers found that teachers asked significantly more questions eliciting ethics, reasoning, and clarifications of student ideas. Additionally, teachers spoke about 20% less often and interrupted runs of student-to-student talk 13%

less often. There were no significant differences between groups of different demographics.

There is little to criticize in Chinn & Waggoner's (2001) study. It uses counterbalancing of texts to ensure that the differences in treatment were not due to higher quality texts. They had high inter-rater reliability on the coding. The only concern for the internal validity is that the researchers do not mention at what time of year the 7-week study took place. If it occurred at the beginning of the year, then a progression of group discussion due to practice and familiarity with each other may have influenced the growth in quality of discussion. A follow up study to control for the influence of student interest in a new format would also have increased the internal validity. The external validity is somewhat compromised by having the researchers present to provide regular coaching to the teachers. However, the fact that the studies findings applied equally to students from lower income rural settings and urban MSES students, who may have had more in-school resources, suggests that the instruction has a large range of generalizability for different demographics.

Overall, Chinn, Anderson & Waggoner's (2001) study suggests that as teachers step back and focus their questions on extending the aesthetic and analytic responses of students, students will take more responsibility for the discussion, offer alternative perspectives, and use each others' ideas to reason about the meaning of the text. In this study the literature-based instruction made use of the features of aesthetic reader-response, and the analytic processes that are characteristic of critical pedagogy. Researchers suggested that this

approach, combined with group-reasoning, greatly increased the depth of cognitive literacy processes students used, which confirms studies in both the Aesthetic Reader-Response and Critical Pedagogy sections of this chapter.

Maloch (2002) researched the relationship between the teacher's role and students' participation in literature discussion groups, in a mixed ability, inclusive, and culturally and economically diverse third grade inner-city classroom, and found that students were able to make the transition to deep student-led discussion of literature when the teacher limited her comments to explicit coaching in the expectations, culture, and language of group discussion. Students were a mix of LSES and MSES and were gender balanced. Fifteen were Caucasian, eleven were African American, and three were Asian Americans. There were four students learning English as a second language, four receiving special education services, and one student designated as gifted and talented. They came from a midsized city and the teacher was Caucasian from a MSES background. In this five-month-long qualitative case study the researcher conducted interviews, made observations, collected artifacts, and took field notes, videos, and audio-recordings of 30 literature discussion groups. Maloch expanded the field notes with methodological and theoretical analysis notes.

Maloch (2002) found that students were able to make the transition to deep student-led discussion of literature when the teacher provided extensive and direct scaffolding in the expectations, culture, and language of group discussion. The teacher sat outside of the discussion groups and provided

comments when necessary. The primary comments of the teacher were metalinguistic comments meant to help students grow in their awareness of different strategies to use to deepen discussion. She used examples of student phrasing to help students see phrases that facilitated the flow of the discussion, she modeled these phrases, she defined roles, and she paraphrased the ideas and evolution of the discussions regularly. The discussion proceeded better when students did an initial retelling of the story, referenced their response logs, and paraphrased each other's ideas. In turn, students used their analyses of discussions to negotiate new roles and rules for the next rotation of discussion groups.

The study meets all criteria of qualitative design as evidenced by triangulation, member checking, outside audit, and clear illustrations of coding procedures and findings. Maloch (2002) states that the study is meant to develop the sociolinguistic perspective of literacy by studying a method of transitioning to a culture of exploratory talk. Maloch does not name specific results for specific students but the study does have transferability for a diverse, inclusive classroom environment.

One angle of critique for the use of literature-based aesthetic reader-response or critical pedagogy may be that students are not accustomed to the format, or come from cultures that do not have familiarity with literature discussions. Maloch's study suggests that the teacher must scaffold the development of such a culture. As the teacher persisted in coaching students and giving them opportunities to develop awareness of group processes, the

students gained authority over the conversation and their own development of meaning from text. Once students developed such a culture they were able to put it to their own uses, which is a feature of someone with a strong literacy identity.

Goatley (1997) observed the effects of literature-based response groups on racially diverse upper elementary students receiving special education services for LD and EMI and found that after six weeks of strategy instruction, journal writing, and coaching in book club discussions, students responded to each other's ideas and used discussion to elaborate their journals. The five students received language arts instruction in a resource room. The school had previously used a skills-based only program in the resource room. In this qualitative study, Goatley observed the skills-based program for three weeks. Goatley collected field notes, a performance based measure of students' awareness of how and what to share in discussions, metacognitive interviews with students, artifacts of written work, and transcriptions of audio and video recordings. The researchers assessed students' responses to literature and use of discussion groups. They repeated the assessment after a 6-week intervention of literature-based instruction.

Goatley (1997) found that students initially were able to use reading logs to pose questions but that students did not respond to each other's questions, elaborate their written responses or sustain conversations. After 6 weeks of literature-based instruction, including strategy instruction, written response, teacher-modeling, and twice weekly student-led book clubs (including explicit

coaching by the teacher), students regularly responded to each other's questions and used the discussion to elaborate their written ideas. Discussions still focused on literal interpretations of the text, meaning that development of personal or critical meaning making would be a next step of development. Students were able to assess their own growth in reading and in discussion skills in the interviews.

There are some problems with the credibility of Goatley's (1997) study. Goatley mentions the use of systematic development of codes, member checking, and triangulation but does not articulate the procedures or the codes that she developed, which reduces the credibility of the study. However, the study is dependable in that it concurs with Maloch's (2002) study, and with Berry & Englert's (2005) study.

Goatley's (1997) study is valuable in that it focuses on students with special needs, a group that is often not mentioned in other studies. Unfortunately, because the study did not extend for longer than six weeks it is hard to know if these students would have developed a deeper capacity to interpret texts based on their own experiences or ideas. Maloch's (2002) study, which was much longer, suggested that the conversations would develop in that direction over time as the students grew in their ability to use the conversation for their literacy needs.

Goatley, Brock & Raphael (1995) conducted an ethnographic study about the experiences of fifth-grade students with learning disabilities, Chapter 1 status, or ESL in a literature-based general education classroom, and found that

students naturally took on group roles and shared comprehension strategies; the diversity of the group made the students question more deeply as they worked to understand each other. The researchers specifically focused on the role of discussion groups in scaffolding opportunities for participation and text interpretation within the group. There were 5 subjects in a book club that took place at the end of the school year; book clubs were voluntary and each one lasted 3 weeks. One student was a female from Vietnam and spoke ESL. One was a male Caucasian and was designated as having behavioral and academic disabilities. One was a female Caucasian who did Chapter 1 part of the day. The fourth and fifth students had no special needs, and were male; one was Caucasian and one was Hispanic. The Midwestern inner-city school had 85% of students receiving free or reduced lunches.

Goatley, Brock & Raphael (1995) triangulated interviews, questionnaires, field notes, audiotapes of discussions, transcriptions, videos, and the written work of students. They each developed categories individually and then compared and negotiated findings, arriving at categories to describe the experiences of the students.

Goatley, Brock & Raphael (1995) found that roles naturally emerged within the group. Since the book was about the Vietnam War memorial, the Vietnamese student acted as a kind of authority, she initiated many questions and topics. Other students invited personal responses and worked to extend the comments of others. Others managed the equity of turn taking and staying focused on the book. Others supported the ideas of their peers and then offered

alternative perspectives. Because there was so much diversity between students the students had to work to clarify each other's perspectives, which in turn developed and deepened their thinking and questions. All students used the group to work out comprehension problems and learn new strategies, including textual cues, inter-textual relationships, and empathizing with characters. All students remarked that book clubs helped their ability to make sense of books, and some students said they read a lot more because of gaining new skills in book club.

Although Goatley, Brock & Raphael (1995) do not name how they gained access to the subjects, the particular group was selected due to being representative of the diversity of students within the school and the classroom, which is close to random selection. They describe the coding process but do not name how much agreement or disagreement occurred between researchers. The detailed transcriptions and descriptions of the students and illustrative examples create both credibility and transferability for a group of this mix of demographics. Because they had no outside audit the study is not confirmable. The study's agreement with the other studies in this section makes it dependable. If the study had included more participants, the transferability would have been more substantial.

The essential idea that this study adds to the collection of studies in this section is that the diversity of the group actually helped the quality of the discussion by creating more perspectives to discuss and having more ideas about reading strategies. As students offered their own contributions to

managing the group and offering strategies and perspectives, they gained authority as readers. The article contains excerpts of conversations wherein the discussion expands students ideas of other races, which suggests the presence of using critical pedagogy to see the world and to recreate one's vision of the world through group reading.

In conclusion, based on the findings of studies in the last three sections of Chapter 2, a body of research has made a case that literature instruction may be at its best for diverse learners when it includes a focus on aesthetic reader-response and/ or critical pedagogy. The studies in the Group Discussion section suggest that exploratory, student-directed group discussion provides one structure for aesthetic reader-response and critical pedagogy. Remarkably, the section suggests that group discussion works well for many demographics, and is particularly aided by diversity, such as is present in an inclusion or inner-city classroom. Teachers aided quality instruction by maintaining clear expectations and routines, by reducing teacher-talk, by focusing comments on discussion skills, by focusing comments on extending and eliciting the reasoning of student comments, by providing time for students to reflect on how to improve their discussion skills, by allowing students to use their journals as a reference, and by asserting that there is no predetermined "right" interpretation—all perspectives are important. Students also progressed from literal comprehension to deeper interpretation. They used the discussion as a way to share reading strategies and problem-solve any breakdowns in meaning making. In turn, students made more elaborative comments, integrated alternative views, made textual

reference, and tested hypotheses. Such skills involve higher-order cognition, and seem to provide a venue for comprehension skills. Further sections of this chapter address how the inclusion of open-ended discussion, along with other structures affects achievement.

Motivation & Engagement with Literature

Many of the studies in the previous sections suggested that aesthetic reader-response, critical pedagogy, and discussion structures increase students' internal motivation to read, and that their behavior shows active engagement with the literature for their own purposes. In Parsons (2006) study, she suggested that high achieving, avid readers, are readers that are very motivated and actively transact with the text. This presents a hypothesis that engagement and motivation may positively affect reading achievement. Several meta-analyses suggest that engagement has a positive effect on achievement (Au & Raphael, 2000; Cummins, 2007; Stahl, 1999; Taylor et. al, 2003; Willis & Harris, 2000). The studies in this section do not answer that question, however they do describe what kinds of literacy behaviors become automatic when students are highly motivated and engaged. These studies describe the structures that prompt such engagement. The studies in this section have large similarities to the findings of the studies in the three previous sections, creating an even stronger case for the effectiveness of aesthetic reader-response and critical pedagogy.

As a whole, these studies represent students of many ages and backgrounds and come to similar findings; a brief summary of each study follows.

In the first study, Davis (2010) studied 19 racially diverse second-grade students' perceived value of reading, time spent engaging in literacy activities, and self-concepts as readers in a Northwestern private school and observed that literature-based instruction that included choice and differentiation resulted in increased reading, discussions, and motivation to improving their own reading processes. Next, Turner (1995) conducted a comparative analysis of how reading tasks affected MSES first graders' motivation in literature-based and skills-based curriculums and found that the literature-based classrooms used student-centered and open-ended tasks, and students had more use of strategies, cueing systems, and greater persistence with challenging texts. In the third study, Sulzby, Branz & Buhle (1993) conducted an ethnography of the sequence of emergent literacy in context of repeated readings in a literature-based classroom for African American K-1 students from high poverty backgrounds; they found that repeated readings resulted in children choosing to read more often, having longer attention spans, and displaying increased confidence to experiment with using the cueing systems. In the fourth study, Kantor, Miller & Fernie (1992) sought to observe which kinds of literacy were developed for ethnically diverse MSES and USES pre-school children in a socio-cultural linguistic environment and found that writing in order to meet daily academic, personal, and social needs was the primary mode of literacy development. In the fifth study, McCarthy (2001) studied the role of literature-based culturally relevant curriculum in the identity construction of 12 diverse, LSES fifth-grade students with mixed reading levels and English language

abilities and found that students' identity as readers was positively influenced by self-assessment conferences, open forums for writing, and good communication between the teacher and a dedicated care-taker. Finally, Dahl (1993) made an ethnographic study following LSES African American and Caucasian first-grade students' experiences of literature-based classrooms and found that they exhibited self-monitoring for meaning-making, self-assessment of literacy skills, and plans for creative response. Each of the following critiques contains detailed descriptions of the tasks and practices of the teachers.

Davis (2010) studied 19 racially diverse second-grade students' perceived value of reading, time spent engaging in literacy activities, and self-concepts as readers in a Northwestern private school and observed that literature-based instruction that included choice and differentiation resulted in increased reading, discussion and student involvement in improving their own reading process. Students were reading at a variety of grade levels and all spoke English as a first language. Nine students were Caucasian, six were African American, three were Asian and one was Latino. Davis served as teacher-researcher. She conducted 10 days of literature-based instruction and then used focus groups with students to code and discuss their perceptions of the instruction. She repeated the process for 10 days with skills-based instruction.

Davis (2010) found that choice was the defining difference between student perceptions of the two practices. In the literature-based setting students had choice in small-group collaboration, they self-evaluated their ability to make meaning and their skills development, and they worked with the teacher to

decide on the next activities for their learning. They read more, had more literary talk, and expressed a positive attitude towards the choices offered with differentiated instruction. Students became more competitive as choice and collaborative activities were reduced in the skills-based instruction. Lack of differentiation left some students frustrated and others bored. Students commented on the lack of opportunity to enjoy reading.

Davis's (2010) study has many limitations. The transferability is harmed due to a lack of detailed descriptions of the students. The teacher's two-year long relationship with students, lack of description of prior instructional methods, and the ordering of the ten-day studies all leave room for bias. A twenty day study is a very limited picture, as students would not be accustomed yet to the framework. However, Davis takes some action to support the credibility of the study: she uses a research assistant to take field notes of student engagement, and record audio that provides grounds for inter-rater reliability. She gathers data from multiple sources, and the study is based on student perceptions, which is a form of member checking. Dahl & Freppon (1995), Dahl (1993) and Kantor, Miller & Fernie (1992) also found that choice, and student goal-setting was an essential element of productive literature-based instruction; this makes the findings dependable.

Although the study does not mention aesthetic reader-response or critical pedagogy, Davis (2010) suggests that the element of choice improves individual motivation and the community of the classroom. It created a ground for

collaborative projects and mixed ability literature circles. Such a setting could easily integrate aesthetic reader-response and critical pedagogy processes.

Turner (1995) conducted a comparative analysis of how reading tasks affected MSES first graders' motivation in literature-based and skills-based curriculums and found that the literature-based classrooms used student-centered and open-ended tasks, and students had more use of strategies, cueing systems, and greater persistence with challenging texts. The subjects were 84 first-grade children from two Michigan school districts. Both districts were suburban and the students predominantly came from MSES Caucasian backgrounds. One district used a basal, skills-based curriculum, while the second district used a whole-language approach. Observers used a structured coding system to note students' motivation-related responses to instruction, including: reading strategy use, learning strategy use, and persistence. They then interviewed the students to discern their ideas about the purpose of literacy tasks and their metacognition about their own learning. Turner used a log linear chi-square analysis to establish a model of instructional practices that best predicted student motivation.

Turner (1995) found that students in literature-based classrooms were 1.35 times as likely ($p < .01$) to use learning strategies (rehearsal, elaboration, organization) than expected; in basal programs they were less likely than expected to use learning strategies. When the teacher used open tasks students were 1.59 times more likely ($p < .01$) than expected to use be aware of using cueing systems strategies and 1.19 times more likely than expected ($p < .05$) to

show persistence with challenging tasks. In contrast, Turner found closed tasks to be far less likely than expected to correlate to persistence and use of reading strategies.

Turner (1995) used motivational theory to define open and closed tasks. Open tasks use goals and processes that the child chooses, and they require higher-order thinking. Open tasks may include teaching multiple strategies for self-improvement in the face of challenge, student autonomy to experiment with their own ideas, pursuit of areas of interest, and the option of social collaboration as a resource. They use higher-order thinking because the child evaluates their options to make a choice, or revise a choice. Closed tasks have teacher designated goals and processes, and primarily require recognition skills. In closed tasks, the teacher makes the choices for the students (p. 416 – 418).

Turner's (1995) study has strong internal and external validity. Although Turner does not name how she gained access to the classrooms, she did randomize selection of the focus students based on students who had average scores on the state achievement tests. The fact that teachers had similar amounts of experience with their form of instruction, and that the communities were very similar demographically helps control for other selection differences. There were a total of 42 focus students, making a large number of samples to draw from. Inter-rater reliability was 89% - 99% for the coding procedures. The regression analyses were primarily held to a high p-value. All of these details establish strong internal validity. In critique of the study, Turner uses narrative descriptions of the instruction rather than supplying the coding system used. Use

of a coding system with open and closed task criteria, and analysis of significance, would have made Turner's assertion that open tasks are used in literature-based instruction stronger. However, the narrative examples she does supply seem to illustrate the definitions, even if they do not use the same quantitative methods. The external validity is very good for the specified demographic, and there are ample descriptions of teaching practice to define the kind of instruction that teachers used.

Turner's (1995) study suggests that the element of choice, and a student-centered approach is more common in literature-base instruction, and contributes to increased authority over strategy use and persistence. These are aspects of positive motivation for literacy. The fact that students pursue their own interests through literacy activities, and can use the social context in their exploration, suggests a framework that coincides with both critical pedagogy and aesthetic reader-response. The structure suggests that students are using literature for their own reasons and responding to literature collaboratively. Other studies may suggest additional structures that would enhance student engagement in aesthetic reader-response or empowerment to change societal oppression systems.

Sulzby, Branz & Buhle (1993) conducted an ethnography of the sequence of emergent literacy in context of repeated readings in a literature-based classroom for African American K-1 students from high poverty backgrounds; they found that repeated readings resulted in children choosing to read more often, having longer attention spans, and displaying increased confidence to

experiment with using the cueing systems. The subjects were from the inner city of Detroit, 90% received free or reduced lunch, 40% had been held back one year in kindergarten, and most students spoke a dialect other than Standard English.

Sulzby, Branz & Buhle (1993) followed the students for two years making observations for one week at a time, three times per year. The teacher was asked to do five repeated readings of each storybook. Researchers then audio-recorded, transcribed, and individually coded the children's emergent literacy during the child's re-reading and writing about the books used in the repeated readings. They assessed letter-sound knowledge once per-year.

Sulzby, Branz & Buhle (1993) found that repeated-readings coincided with children developing a longer attention span and choosing to read more often. Students engaged the most when classrooms had libraries, independent reading time, group discussion, and student-teacher collaborative assessment of growth. Students progressed through common developmental stages of literacy, incorporating attention to print, intonation, and using the structure and syntax of written language. When the teacher encouraged the children to read and write emergently using a familiar text, they were able to overcome initial self-perceptions that they were not readers or writers. Student graphophonics developed through explicit modeling of sound-symbol relationships using literature, and benefited from application when the children used invented spelling to re-write the stories.

Sulzby, Branz & Buhle (1993) triangulated data between two observers per classroom, individual assessments, and transcriptions, but they do not provide detailed descriptions of the coding procedure, inter-rater reliability, member-checking or an audit. They do not provide illustrative examples of their findings. All of this leads to room for bias and decreased credibility and confirmability. The observations of the classroom practices are consistent with the definition of literature-based instruction in Chapter 1 of this paper. They also concur with Morrow's (1988) findings about repeated readings, which supports the dependability of the study. The study is highly transferable due to all teachers being very new at implementation, and detailed description of the demographics of the subjects.

Despite the fact that Sulzby, Branz & Buhle (1993) should have done more to control for their own biases, the study challenges the view that children who speak an African American dialect, and come from backgrounds with little experience of literature, benefit most from basic skills instruction before they can engage complexly with literature. The students did receive explicit instruction in skills, but it was contextualized by developing their familiarity with books through repeated readings, choice, and discussion of literature. Students were also encouraged to read and write emergently, incorporating new skills as they went, and self-assessing, rather than working from a passive right/wrong framework. The researchers suggest that this built student motivation and confidence in their ability to be readers and writers. Because this framework included response through discussion and writing, it is one that would support and may have

included aesthetic reader-response and critical literacy, even if the researchers did not mention these features in their report on writing and discussion.

Kantor, Miller & Fernie (1992) sought to observe which kinds of literacy were developed for ethnically diverse MSES and USES pre-school children in a socio-cultural linguistic environment and found that writing in order to meet daily academic, personal, and social needs was the primary mode of literacy development. The subjects were situated in a university lab setting in the Midwest and were from families that prized literacy. The purpose of the study was to create an ethnography that would develop of the social-constructivist approach to literacy.

Kantor, Miller & Fernie (1992) gathered data from multiple sources: videos, transcriptions, field notes, interviews, and retrospective journals of teachers. Researches did a grand tour of the data to develop and analyze where literacy activities were occurring, how the students used the activities, and how the teachers implemented the activities.

Kantor, Miller & Fernie (1992) found that writing was the primary mode of literacy development. The teacher dictated the children's written work to make community reference lists of the children's future plans, democratically post rules, work collaboratively to publish and plan projects, and organize and label student work. The literacy activities focused on learning to function in the culture of school, and to use reading and writing as a way to meet daily personal, social, and academic needs.

Given that Kantor, Miller & Fernie's (1992) study took place in a lab setting, included participant observers, and did not supply detailed information about the subjects, the transferability is marginal. The study does not have transferability for more economically diverse public settings. Credibility and confirmability is reduced because there was no inter-rater, and no outside audit. On the other hand, the coding procedures are detailed and illustrated with ample examples. The data has several points of triangulation, and different researchers than the participant observers coded the data, which may reduce bias. Interviews serve to member-check the findings. Each of these aspects supports the credibility of the study.

Despite the fact that Kantor, Miller & Fernie's (1992) study is not particularly transferable it makes an interesting assertion that writing to organize and revise school life is a way for children to begin engaging with literacy. Dahl's (1993) study came to similar findings about students' use of integrated writing as a way to apply reading to their lives and ideas. The theory of critical pedagogy suggests that literacy is a way to make and remake one's view of the world, and consider one's own role. In that sense, although Kantor, Miller & Fernie do not mention literature, ways of thinking consistent with critical pedagogy are still present, and the researchers observed that these ways of thinking were a useful vehicle for young children's engagement with literacy.

McCarthy (2001) studied the role of literature-based culturally relevant curriculum in the identity construction of 12 diverse LSES fifth-grade students, with mixed reading levels and English language abilities, and found that students'

identity as readers was positively influenced by self-assessment conferences, open forums for writing, and good communication between the teacher and a dedicated care-taker. The students were from a school in Texas that served inner-city, suburban and rural students. All of the 12 students received free or reduced lunch. Seven were Latino/ Latina, two were African American, one was Caucasian, and one student was Latina and Caucasian. Based on teacher assessment, students were classified in the following way: one reading above grade level, four reading at grade level, four reading just below grade level, and three reading far below grade level. There were two ESL students.

McCarthy (2001) used qualitative narrative methods to transcribe interviews with the students, the teacher, parents, and peers regarding the curriculum, habits, and perceptions of students as readers. Additionally, McCarthy collected student journals and literary letters written to peers or the teacher. McCarthy analyzed the data with the constant comparative method (Stauss & Corbin, 1990).

McCarthy (2001) found that students who chose to read and write as part of their identity had literacy encouragement from at least one caretaker, the caretaker and teacher had strong communication, and all sources viewed the student as a "strong reader or writer." When the school program found students to be "average readers," they found that the students had less support at home and had inconsistent reputations as readers. And when the school program found readers to be far below grade level, the students emphasized different aspects of their identities as more important. The school wide reading

assessments greatly influenced self and peer views of students as readers. Students tended to self-assess their writing because the school did not assess a level of writing ability for each student. The students used literary letters and group discussions more actively to develop identity and respond to literature than the prompted literature response journals.

Aside from two minor reservations, McCarthy's (2001) study meets all criteria of design for quantitative studies. McCarthy gives illustrative examples of the findings, analysis procedures, and selection process; he supplies information about each individual student's development and background, and he member-checks and triangulates all the informants' perspectives. Credibility, confirmability and transferability are all strong. McCarthy selects students based on their background, parental permission, and school assessment of reading levels, but does not say if the study excluded some students who also met criteria, which may be a credibility flaw. There was no outside audit, however, the report provides ample examples from multiple sources make up for this.

The McCarthy (2001) study is interesting in that it suggests that conversational written response and discussion of literature in mixed ability groups provides a ground for students to articulate their use of reading in their own lives. The reputation of students, based on fixed classification of ability in the school assessments, negatively or positively impacts a student's status and own view of him or herself as a reader. Dahl & Freppon (1995), Dahl (1993), Davis (2010) and Sulzby, Branz & Buhle (1993) all found that developmental

models of collaborative teacher-student assessment positively influence engagement with literacy. This data suggests that students may have been more open to seeing literature as part of their identity, if they had ongoing control over assessing their own needs as readers, rather than a fixed reputation. Regardless, literary letters and discussion involving culturally relevant texts provided useful grounding for aesthetic reader-response.

Dahl (1993) did an ethnographic study following LSES African American and Caucasian first-grade students' experiences of literature-based classrooms and found that they exhibited self-monitoring for meaning making, self-assessment of literacy skills, and plans for creative response. The subjects were first-graders from two Midwestern cities. Most children qualified for free and reduced lunch; out of the twelve children six were Appalachian Caucasians and six were African-American; five were female and seven were male. In this two-year-long study, Dahl followed one focus child per month audio-recording their spontaneous utterances during activities. She collected all student work to form additional triangulation. Dahl derived categories using constant comparative method (Glaser, 1969).

Dahl (1993) found that learners showed a sense of self-monitoring and self-assessment. The students made hypotheses about print details, sound-symbol relationships, language patterns, and strategies for reading that were focused on making meaning from literature. They commented on their plans to produce writing that was inspired by their reading.

Dahl (1993) arranged no outside audit and reports no inter-rater reliability for the categorization, reducing the credibility and confirmability of the study. More details about the home context of literacy and the particular needs of students would be helpful for transferability. However, there are many strengths to the study. Dahl describes how focusing on metacognition, interaction, and the role of talk influenced the categories she derived and gives clear examples of her findings. The study includes three points of triangulation. The findings were member-checked. All of these details create strong credibility, even if there was no inter-rater. The study also has dependability because Morrow's (1988) study also finds that diverse pre-literate LSES children approach books with complex comments and questions if given the opportunity.

Many aspects of the instruction that Morrow (1988) documents are noteworthy to this paper's exploration of aesthetic reader-response and critical literacy. Children had access to an extensive library organized by a wide variety of print forms and genre. The teachers introduced literature through themes in a two-hour reading block, followed by continued writing about the theme in content areas. The teacher used conferences to determine which skills students needed to work on next, and provided instruction in context of theme related materials, followed by revisions of written work. Discussions focused on intertextual relationships, opinions comparing and contrasting different authors, and problem-solving their own writing. The teachers encouraged collaborative projects and student choice. All of these structures allow opportunities for children to use reading to articulate their own ideas and identify what they need to do improve

their efforts. Such habits of discussion, creative response, and self-assessment are inherent in aesthetic reader-response and critical pedagogy processes.

When literature-based instruction included both discussion and written response, Morrow observed the students making comments about their own identity as readers and writers, including their internal motivation to acquire basic skills.

Morrow (1988) observed diverse LSES children naturally pursuing skills as they self-directed their own ability to make meaning and create literature in response to literature. The study is significant because it suggests that, for this population of students, basic skills did not need precede complex reading processes. The children in this study decided to learn basic skills as they made authentic connections with what they read.

In conclusion, the Motivation & Engagement section revealed that structures that contribute to student motivation and active engagement are the same structures used in aesthetic reader-response and critical pedagogy. Some studies were more credible or valid than others, but all studies had similar findings, which contributes to dependability. Similar to the findings in the Group Discussion section, this collection of studies found that choice, open-ended tasks, self-assessment, and goal setting conferences seemed to positively affect motivation and engagement with literacy. Teachers also used repeated readings; themes; strategy instruction; integration of literature, writing, and discussion; extensive libraries; independent reading time; culturally-relevant texts; and strong home to school communication. With these structures, researches observed that students read more often, were motivated to improve

skills, developed a confident reader-identity, and were more persistent in the face of challenge. Contrary to the belief that basic-skills precede higher-level engagement with literature, these students were motivated to improve skills in order to make-meaning of texts that interested them. This further suggests that structures that coincide with aesthetic reader-response and critical pedagogy contribute to student ownership of their own improvement of their reading process, including basic skills. Future studies will need to empirically test this hypothesis.

Best Practices in Literacy Instruction

A group of studies qualitatively and quantitatively approached literacy instruction by identifying effective teachers and schools, and then noting which structures and processes they used. This group of studies is useful to the central questions of this paper because elements of the findings confirm findings in the previous four sections of this chapter. The studies add additional practices that may serve to qualify effective facilitation of literature-based instruction and the context in which aesthetic reader-response and critical pedagogy occur. The studies also take a large view of the classroom management, community, and policy contexts that affect instruction, and suggest that a limitation of much research on literacy is that it often does not take into account larger contexts.

A brief summary of the studies included in this section follows. In the first study, Taylor, Pearson, Clarke & Walpole (2000) studied the characteristics of effective schools and K-3 reading teachers in mid to high poverty schools in rural, suburban and urban settings across the U.S. and found that strong

communication with families, small group work, independent reading, strategy instruction, high-order questions, and the use of response journals were characteristics of the most effective schools. In the second study, Morrow, Tracey, Woo & Pressley (1999) sought out experienced first-grade teachers from three MSES to LSES New Jersey school districts that administrators considered exemplary and found that common practices were: collaboration; differentiation; multiple venues for reading and extending reading experiences; multiple texts on a chosen theme; journals; charts and tasks to integrate strategies and skills; and self-evaluative conferences with students. In the third study, Wharton-McDonald, Pressley & Hampston (1998) studied the characteristics of high-quality literacy instruction for economically and culturally diverse first-grade students in rural, suburban, and urban settings and found that outstanding teachers were distinguished by instructional balance as evidenced by differentiated skills instruction integrated coherently into literature-based experiences with high-quality literature. In the fourth study, Mosenthal, Lipson, Torncello, Russ & Mekkelsen (2004) examined the contexts and practices of schools whose students have good scores on state reading tests and found that schools that provided ample time to read, write, and discuss in response to literature were most effective for every demographic and setting. Finally, Gersten (1996) studied language arts learning and literacy engagement for LSES, urban, culturally diverse, language minority students from California, grades 3-6, in order to discern best practices and found that characteristics of best practice included: strategy instruction that focused on meaning-making; extended

discussions; integration of native languages and cultural assets; and higher-order questioning.

Taylor, Pearson, Clarke & Walpole (2000) studied the characteristics of effective schools and K-3 reading teachers in mid to high poverty schools in rural, suburban and urban settings across the U.S. and found that strong communication with families, small group work, independent reading, strategy instruction, high-order questions, and the use of response journals were characteristics of the most effective schools. The researchers tried to include a broad range of subject demographics. The subjects were 14 schools in Virginia, Minnesota, Colorado and California, and were from a proportional mix of rural, small town, suburban and inner-city settings, with a varied poverty rate of 28% to 92%. 6 of the schools were bilingual schools. Schools used national reading programs, reforms integrated into a homegrown program, a homegrown program alone, or no designated school-wide reading program.

Taylor, Pearson, Clarke & Walpole (2000) used quantitative correlation analyses and significance tests. They used three reading measures (fluency, comprehension and word recognition) and the districts' grade 3 reading tests as pre and post tests to measure the growth in student achievement of two teacher-selected students per teacher. They observed and video recorded each teacher five times, collected the teachers' logs of daily literacy activities, conducted a survey and used an interview to obtain data on classroom practices and school environment.

Taylor, Pearson, Clarke & Walpole (2000) found that teacher communication with families occurred significantly more often (73%) in the most effective schools than schools they designated as average and low achieving based on student growth on tests, $F(2, 65) = 5.25, p < .01$. An ANCOVA revealed that effective teachers had significantly more time spent in small groups, $F(2, 60) = 9.63, p < .01$, and more time spent in independent reading, $F(2, 60) = 4.24, p < .05$. A chi-square analysis found that the most effective schools spend 53% of the time in strategy instruction versus 13%, chi-square = 5.4, $p < .05$, and 17%, chi-square = 5.0, $p < .05$, in the schools they rated as low and moderately achieving respectively. Researchers did not find differences in the instructional philosophy of the teacher to be significant. Finally, some correlations are interesting, although the researchers did no significance statistics. The study found that teachers in the most effective schools spent 31% of the time asking higher-level questions (versus 0% in the schools with lowest scores), and effective schools spent 48% of the time using written response journals (versus 17%). However, only 16% of teachers in the whole study asked higher-level questions that required students to relate text to their own perspectives and experiences. The researchers also found that school status did not always correlate with teacher effectiveness, and that all of the most effective schools used homegrown reading programs instead of national programs.

Although the statistical analyses of the correlations are strong, there are some problems with the selection and the initial measures of student achievement. The researchers do not name why or how they selected these

particular schools, and the fact that teachers chose the focus students leaves room for biases to affect the results. Teachers may have given special attention to the development of the focus students. The researchers used district tests to establish growth of students and this is not really reliable given that different districts use different tests. In addition they say they adjusted test results for poverty, and then the most accomplished schools often had high poverty rates—this suggests that the adjustment may not have been appropriate, or that students from high poverty grew the most because they entered with the least experience with literacy. The researchers do not address these details, and so their ranking of most effective to least effective teachers may or may not be valid. None-the-less, the statistical analysis connecting practices with high growth as determined by the district are very strong. Many of the findings have high confidence levels and large f-values, and the correlations point to directions for further research or interesting comparison to other studies despite the limitations. The inter-rater reliability for the observational coding was also strong (within the 90% to 100% agreement range). The external validity is wide, but the researchers do not supply details about differences between demographics and this limits specific generalizability.

The findings of this study suggest that strategy instruction is a useful structure to include, and that writing and discussion were useful for the development of literacy and achievement on basic skills. Finally, the fact that few teachers used aesthetic response (as suggested by questions related to student perspectives and experiences, and response journals), despite this being a trait

of the most effective schools, and the fact that the researchers did not study the use of critical pedagogy, points to lack of understanding of these distinctions in literature-based instruction in the field of educational practice.

Morrow, Tracey, Woo & Pressley (1999) sought out experienced first-grade teachers from three MSES to LSES New Jersey school districts that administrators considered exemplary and found that common practices were: collaboration; differentiation; multiple venues for reading and extending reading experiences; multiple texts on a chosen theme; journals; charts and tasks to integrate strategies and skills; and self-evaluative conferences with students. They studied the teachers' characteristics, practices, and philosophies of reading instruction in order to gain insight into best practices for literacy instruction. The teachers had 9-25 years of experience and all had master's degrees. The teachers came from three New Jersey school districts with middle to low income families. 50% of their students were Caucasian, 20% African American, 10% Hispanic, 10% Asian American, and 10% other ethnicities. Morrow et al. asked administrators from the three school districts to select six teachers who had most students emerge as readers and writers. They triangulated consensus with achievement records, test scores, ability assessments of students, positive comments by families and staff, and the known ability of the teacher to articulate his or her philosophy of teaching. Researchers then observed each classroom eight times throughout the school year taking detailed field notes. They developed categories on an ongoing basis, interviewed the teacher on their

philosophy at the end of the study, and then synthesized the findings based on reoccurring trends.

Morrow, Tracey, Woo & Pressley (1999) found trends in the environment of the school and classroom. Exemplary teachers benefited from collaboration with administrators, and parents, but also by having the freedom to choose how they would instruct. Classroom management included regular agreements and goal setting with students as well as clear expectations. All of these teachers worked to democratically establish rules, and spent a considerable amount of time teaching children classroom routines so that they could work independently while the teacher met with small groups and individuals. The teachers used read-aloud time, partner reading, guided reading, and independent reading throughout the week. They used a mix of teacher selected, student selected, and leveled text work as needed. Teachers used the morning message, small flexible groups, and centers to focus on strategies and skills as needed for different groups of students. They had regular conferences with students to read over the child's activity reflections and self-assessment, and to share ideas and write notes to parents. Effective teachers used a variety of charts to focus students on word level, syntactical, or textual features of the literature they were engaged in. Teachers based literature study on themes that they integrated with all subject areas. Students wrote regular journals of events and reading experiences, dictated stories, and learned skills and structures by revising writing in order to publish it in the classroom library. In addition, students had opportunities to re-enact, re-read, discuss, and re-write the literature they read.

Although the study is very detailed in the synthesis of findings and makes many specific suggestions for best practice, Morrow, Tracey, Woo & Pressley (1999) do not give enough information on the method to be credible. They do not say how they chose the administrators who selected the teachers, and they mention that the districts had been undergoing extensive staff development that they also do not name. These facts leave considerable room for a biased sample, as the districts and administrators may favor a social-constructivist approach. Although they mention triangulation of data and analyses by multiple researchers, they do not describe the procedures or give examples to illustrate the categories they arrived at. The researchers of this study do not mention using a rotation of observers per classroom, which leaves more room for biases to develop as one observer becomes familiar with one classroom. Additionally, the confirmability is harmed because they had no outside audit. This study is credible in that other stronger studies come to similar findings.

The descriptions of materials, classrooms, and school environments, and routines, rules and practices are very valuable in context of the other studies in this paper. Since many studies find that literature-based instruction is beneficial for achievement and other aspects of literacy, the fact that the practices in this study match the definition of literature-based instruction reduce the need for this study to be internally credible. It is more useful to see the study as a survey of practices found in classrooms that match a definition of literature-based instruction. Other studies may more aptly study the effectiveness of the specific features presented in this study. Clearly, aesthetic reader-response was at play

through the use of experience journals and multiple ways to relive the story. Since students could change the story, they may have been using critical pedagogy to adjust the culture or power relationships the author used.

Wharton-McDonald, Pressley & Hampston (1998) studied the characteristics of high-quality literacy instruction for economically and culturally diverse first-grade students in rural, suburban, and urban settings and found that outstanding teachers were distinguished by instructional balance as evidenced by differentiated skills instruction integrated coherently into literature-based experiences with high-quality literature. The study was qualitative and had the purpose of developing the theory of a balanced approach to literacy, meaning that skills are taught in context of meaningful literature experiences. The subjects were from suburban and rural school districts. One district served low to middle income families, one served upper-middle-class families, and one served rural families with LSES; these three districts primarily served Caucasian students. A fourth suburban school district served ethnically and economically diverse students. The classes served some students who received special education services. The districts had adopted different approaches to literacy, and language arts coordinators selected teachers that they considered very effective, moderately effective, and least effective. Wharton-McDonald, Pressley & Hampston observed each classroom, compared notes, interviewed the teachers and assessed student amount of reading, writing coherence and engagement. They selected three teachers based on the assessments and report the findings for these teachers.

Warton-McDonald, Pressley & Hampston (1998) found that outstanding teachers were distinguished by instructional balance as evidenced by teaching skills integrated coherently into literature-based experiences with high quality literature. They had high instructional density, meaning that they had multiple goals in each lesson, integrated content areas, and stopped for mini-lessons as needed. They scaffolded students' self-regulation by asking metacognitive questions about strategy use. They employed thorough integration of reading and writing activities. They had predictable routines and expectations that incorporated student input with management and lesson goals. The two outstanding teachers in the rural district used a balance approach to literacy despite district assertion of having no whole language and using only basal series instruction. They designed their own lessons and provided additional avenues for reading experiences.

Wharton-McDonald, Pressley & Hampston's (1998) study does not sufficiently account bias, but does take other measures for credibility. The selection of the study leaves some room for bias as the language arts coordinators selected the teachers, and the researchers do not say how they gained access to the language arts coordinators. They do make an effort to control for these biases by doing their own assessment of the effectiveness of each teacher, but if everyone shared the same bias then this move does not do much to control for bias. They give many examples of practices, paired with member-checking interviews, which helps the credibility of the study. There is no confirmability due to no outside audit. The study is transferable for Caucasian

suburban and rural students with diverse levels of income. Unfortunately, they do not describe the circumstances that surround the marginal effectiveness in the more ethnically diverse and the upper-class school districts, so readers can not discern transferability for those demographics.

Studies with stronger credibility came to similar findings as Wharton-McDonald, Pressley & Hampston (1998), making the dependability of the study strong overall. The study reasserts the idea that students can develop strategies and skills at the same time as meaningful experiences of literature. Aside from that it does not go deeply enough into the literature instruction to make inferences about aesthetic reader-response or critical pedagogy. However it suggests that quality literature-based instruction should incorporate strategy work. The next section, Strategy Instruction, will address this idea.

Mosenthal, Lipson, Torncello, Russ & Mekkelsen (2004) examined the contexts and practices of schools whose students have good scores on state reading tests and found that schools that provided ample time to read, write, and discuss in response to literature were most effective for every demographic and setting. Particularly, they sought to see if effective practices differed for schools with different socio-economic and cultural demographics. The study focused on two successful schools, and one school deemed less successful, in each of three categories of schools in Vermont. The first group was “country schools” that served primarily Caucasian students from LSES backgrounds and little history of education in the families. The second group was “main street schools” serving mostly Caucasian students from MSES backgrounds. The final cluster, the

“uptown schools,” served MSES to upper income families, and had schools with the most resources, classroom aides, and teachers living in the neighborhoods they taught in. Vermont’s standards were based on literature-based instructional practices.

In order to ensure representative selection, Mosenthal, Lipson, Torncello, Russ & Mekkelsen (2004) used quantitative analysis for two years of test scores and the associated demographics. Then they conducted a qualitative case study, observing classrooms, transcribing interviews, and conducting a cross-case analysis for critical attributes followed by comparison based on demographics.

Mosenthal, Lipson, Torncello, Russ & Mekkelsen (2004) found that schools that provided ample time to read and discuss literature were most successful. Teachers encouraged and modeled metacognitive approaches to reading, and shared a long-term common vision with the school. The successful schools from the rural LSES community integrated Reading Recovery into literature-based contexts, provided large libraries, and integrated response journals and discussion into SSR time. The same was true of the main street schools. The less successful MSES, LSES and USES schools all focused on reading at the word level in the early grades and used 20- 30 minutes of SSR with no discussion or written response. They found no differences in scores between successful schools of different SES communities.

The quantitative method Mosenthal, Lipson, Torncello, Russ & Mekkelsen (2004) used to select the participants controls for bias. Additionally, researchers

triangulate test scores, interviews, and classroom observations to derive attributes. Although they did not conduct the evaluation process separately and then establish inter-rater reliability, they did have ongoing discussion between six researchers, using commonly established qualitative methods, which may help to control for bias. By these terms, the study has strong credibility. It is dependable in that other studies distinguish quality literature-based instruction as including a focus on meaning-making, metacognitive strategies, and elements of reader-response and discussion and not just the presence of trade books. They had no outside audit, which harms confirmability. The study has good transferability for multiple Caucasian demographics in that it describes the schools, communities, and the resources of the families in detail.

Mosenthal, Lipson, Torncello, Russ & Mekkelsen's (2004) study is useful in defining some features of successful literature-based instruction. There is no mention of critical pedagogy, but the structures that allow for aesthetic reader-response were present in the successful schools. This study adds to the ongoing trend of the combination of strategy instruction, discussion, and written response to literature as effective teaching practices.

Gersten (1996) studied language arts learning and literacy engagement for LSES, urban, culturally diverse, language minority students from California, grades 3-6, in order to discern best practices and found that characteristics of best practice included: strategy instruction that focused on meaning-making; extended discussions; integration of native languages and cultural assets; and higher-order questioning. The subjects were from three large urban school

districts in California, in which 60% to 85% of the students spoke a native language other than English, and were in their first or second year of speaking English. 91% to 96% of the students in each school received free or reduced lunch. The study included schools with Spanish speakers, Southeast Asian languages, and some students who spoke Ethiopian and Somalian languages. The English speaking students were primarily African American. Additionally they included two Spanish-English bilingual classrooms in El Paso, Texas, that had teachers who had been using literature-based instruction for over 10 years. All but one classroom used trade book literature.

Gersten's team of researchers (1996) observed each classroom for 200 hours over 2 years. They used pre-established categories to focus on one focus student per year who the teacher deemed as a struggling reader. They triangulated interviews with administrators and the teachers. The research team created final themes of the findings through several sessions of debate. Outside consultants evaluated the findings twice per year.

Gersten (1996) found several themes for promoting literacy. Teachers focused on think-aloud strategies and made extensive use of graphic organizers to build background knowledge, chart vocabulary, and predict and confirm meaning making. Teachers provided feedback based on meaning-making, reasoning, and strategy use, rather than grammar, syntax, and pronunciation. They allowed students to use their native language in their early explorations of the literature. The teachers fostered active involvement and extend literature discussions. They made use of higher-order questions and they worked to use

the assets of the cultures of the students in the choice and approach to the literature.

Gersten (1996) indicates that he had investigated direct instruction for 12 years and that this was the first study he did looking into literature-based instruction, this fact makes the study potentially less biased than studies conducted by life-long advocates of literature-based instruction. The extensive observations and regular auditing and consultation help to confirm the findings and establish credibility. They do not report which findings applied in which schools and this clouds the transferability of the study. If the themes were consistent in all schools then the transferability is strong for students from many cultures and who live in high poverty settings. The study names a large body of research that affirms the findings.

Gersten's (1996) study is interesting because it suggests that language development is driven by meaning-making, and that contrary to some opinions, basic-skills need not come before a high level of engagement with literature. It does not specifically name the use of aesthetic reader-response. It also does not name the use of critical pedagogy, for which this could have been a ripe environment. But the study does mention using the students' cultures and experiences as an asset, aligning materials to the students' concerns, and involving students in written and verbal response groups. All of these aspects in addition to the structures of literature-based instruction suggest that environments that allow for aesthetic response and critical pedagogy may be very beneficial for LEP and ELL students. These aspects are further supported

by the use of strategy focus, graphic organizers, and incorporation of home language.

In conclusion, many aspects of the “best practices” studies confirm the studies in earlier sections of this chapter. All of these studies recommended using meaningful contexts, rather than extracted skills instruction. The studies found that discussions and journals with a focus on the perspectives and experiences of students were effective structures. Students also use theatre, repeated readings, and re-writing books to re-experience texts. Like the studies in the Engagement and Motivation section these studies recommend using conferences with students to focus on how a student can gain strategies and skills to make meaning from texts. The studies also encourage the use of the home language and cultural assets of the students. They assert that while school to home communication, and a shared vision within the school are helpful, they are not essential to good teaching. Additionally, they suggest that democratic and explicit rules and routines greatly facilitate instruction. In addition, these studies add the idea that students need multiple venues to interact with books: large group, partner reading, flexible small groups, and independent reading. Finally, the studies asserted that strategy instruction within a meaning-making focus was a key component of effective teaching. The next section explores this idea further.

Strategy Instruction

The last section contributed the idea that the incorporation of strategy instruction into literature-based meaning-making contexts may make or break the

effectiveness of the instruction. One controversy related to literature-based instruction is the idea that it does not include skills, and that this is especially insufficient for students who enter school with little knowledge of books. Given that literature-based instruction is interpreted in many ways, some literature instruction may in fact avoid integration of explicit attention to skills. However, most advocates assert that the definition includes skills in context of meaningful experiences with literature. But skills are different than strategies. Strategy instruction is based on the ideas that four cueing systems must be simultaneously at work for effective meaning making: graphophonics, syntax, semantics, and pragmatics. Strategy instruction also includes the recursive reading process; using prior experiences, predicting, correcting or confirming, and then evaluating the validity of the experience or ideas that the text presents against prior knowledge. This provides the opportunity for the reader to assimilate or accommodate a new view of the world (Gilles, 1988). Therefore, theoretically, strategy instruction is a way to understand how literature-based instruction incorporates skills work into a meaning-making driven reading process. The studies in this section further define strategy instruction, describe how teachers can scaffold strategy instruction, and then evaluate the resulting benefits for students.

A summary of each study in this section follows. In the first study, Taylor, Pearson, Peterson & Rodriguez (2003) did a correlative study on how active higher-order meaning-making strategies affect ethnically diverse, LSES, and ESL first grade student's growth in comprehension and fluency; they found that

higher-order thinking correlates with a 35% positive difference in fluency and comprehension in contrast to teachers who focus mostly on skills and basic comprehension. In the second study, Duffy (1993) sought to shed light on why some rural second grade students who were exposed to strategy instruction in a literature-based context were not developing metacognition or personal uses for reading and found that this may have resulted from strategy instruction that was focused on naming the steps of a strategy rather than a central focus on creating a meaningful interpretation of the text. In the third study, Block (1993) used a quasi-experimental study to examine the effects of strategy instruction in a literature-based classroom; the students were in grades 2-6 and from three schools with very different demographics and she found that students in the experimental groups used more strategies to problem solve, were able to explain context cues more often, scored higher on two standardized tests, and developed social skills for literature discussion that control groups did not display. Finally, Bauman & Ivey (1997) wanted to know what ethnically diverse second graders learn about reading, writing, and literature in a yearlong program of strategy instruction integrated with a literature-based environment; the students had developmental delays and behavioral disorders and came from an LSES community with high mobility rates; they found that the program resulted in increased engagement, reading levels, articulation of context and syntax cues. As a whole this section show several cases where strategy instruction focused on meaningful literature resulted in a high degree of confidence and ownership of the reading process.

Taylor, Pearson, Peterson & Rodriguez (2003) did a correlative study on how active higher-order meaning-making strategies affects ethnically diverse, LSES, and ESL first grade student's growth in comprehension and fluency; they found that higher-order thinking correlates with a 35% positive difference in fluency and comprehension in contrast to teachers who focus mostly on skills and basic comprehension. The researchers defined higher-order questioning as instruction that positions meaning making as the purpose of reading. Students make connections to prior-knowledge, examine theme, and interpret the motivations and actions of characters. Strategy instruction includes supporting metacognition about when to use the four cueing systems to solve problems in meaning making. The subjects were 135 first grade students in nine schools with 70%-95% of students receiving free or reduced lunch. Schools were in the Midwest, Southeast, Northeast, Southwest, from rural and urban settings. From 2% to 68% in each school were ESL students, and 67% to 91% were non-white, however the study does not identify these students further. The teachers had 0 to 35 years experience and 40% had master's degrees.

Students were part of a one-year assessment of practices in the CIERCA School Change Project; they presented findings with HLM analysis of variance reporting mean statistics and correlations of specific factors of instruction to growth. Taylor, Pearson, Peterson & Rodriguez (2003) conducted pre- and post-tests. They triangulated data on instructional practices with interviews, observational coding, and multiple raters. For first grade fluency ($N = 135$), they found 35% of the 35% variance between teachers correlated with higher-order

thinking. For first grade comprehension ($N = 107$), 27% of the 35% variance between teachers correlated with higher-order questioning.

Taylor, Pearson, Peterson & Rodriguez's (2003) study has some significant problems in the internal design and identification of the participants. The pre and post tests were different and researchers give no reliability statistics between tests, this means that the findings could be due to the change of test rather than the instruction. Researchers report that teachers used trade books some of the time, but do not study how this correlated with higher-order thinking. The inter-rater reliability was 90% on the coding of instructional practice, which is good, but given the other flaws the validity of the internal design is poor. Because they do not give specific information about the demographics of students in the different instructional settings it is impossible to know which demographics this study applies to, which greatly harms the external validity of the study.

Duffy (1993) sought to shed light on why some rural Caucasian second-grade students who were exposed to strategy instruction in a literature-based context were not developing metacognition or personal uses for reading and found that this may have resulted from strategy instruction that was focused on naming the steps of a strategy rather than a central focus on creating a meaningful interpretation of the text. The four focus students were from a rural school receiving Title I funding. Based on Duffy's use of Caldwell's (1990) *Qualitative Reading Inventory* he states that none of the students had

measurable reading levels prior to the study. The teacher had 17 years of experience, but was only in her first year of using literature-based instruction.

For one year, twice monthly, Duffy (1993) audio-recorded and transcribed conversations with the teacher before and after the lesson, and conversations with students about what the lesson was about, when the student planned to use what was learned, and how to do what they learned. He also observed the implementation on the days of the interviews. In the analysis he compared the expectations and thinking required during instruction to what students claimed they were supposed to learn.

Overall, Duffy (1993) found that students made more comments about the isolated strategies they were learning, but they had misconceptions about when and how to use the strategies, and could not comment on what it meant for their growth as readers. The teacher emphasized naming the strategy, steps for using the strategy, entertaining lessons, and literal questions. The teacher did not model flexible, metacognitive use of strategies, and she did not encourage student discussions with multiple interpretations. The teacher failed to name a connection between the literature and students' interests.

Aside from a lack of an outside audit, this study is very strong for all other criteria of design. Duffy (1993) selected the participants through his involvement in the literature-based strategy instructional training program for teachers in the district, and a five-year study. He selected the teachers that were having different outcomes than the targeted outcomes, and only one of the teachers was actually trying to implement literature-based instruction, the teacher featured in

this study. He triangulated data through member checking and has detailed descriptions and examples of the findings. Both of these aspects support credibility. The descriptions of the students as rural, Caucasian, and designated as reading far below expected levels created strong transferability. The definition of literature-based instruction matches the definition in Chapter 1 of this paper.

Duffy's (1993) study suggests that literature-based strategy instruction does not help metacognition and purposeful reading if the teacher does not cultivate structures to place meaning making as the driving process. Both aesthetic reader-response and critical pedagogy put meaning, social diversity of values, and use in real life at the center of instruction, therefore, the facilitation of these processes along with strategies may be a more helpful form of literature-based instruction.

Block (1993) used a quasi-experimental study to examine the effects of strategy instruction in a literature-based classroom; the students were in grades 2-6 and from three schools with very different demographics and she found that students in the experimental groups used more strategies to problem solve, were able to explain context cues more often, scored higher on two standardized tests, and developed social skills for literature discussion that control groups did not display. Subjects were 352 elementary students were located in the Southwest and were in one private school and two public schools with very different demographics. School 1 had 68% Mexican American, 16% African American, and 14% Caucasian, and 2% Asian American students. School 2 was comprised of 52% Caucasian, 30% African American, 12% Mexican American, and 6%

Asian American students. School 3 had 61% Caucasian, 22% African American, and 17% Mexican American students. Each school had experimental students in eight experimental classrooms, and eight control classrooms.

Research assistants for Block's (1993) study gave strategy mini-lessons, with graphic organizers, followed by application in literature and content areas of students' choice. They had a pretest for self-esteem/ social competence, the Harter self-perception profile (1985), and a post-test. During the study the researchers asked students how they were applying strategies in other contexts. After the study researchers asked students what they learned. Students responded in writing. On the last day of the study (32 weeks) the lesson was recorded and scored by outside raters for strategy use and themes. They were given the Iowa Test of Basic Skills, a problem-solving test, and a reasoning test post-study.

Block (1993) found that students in the experimental groups had significantly better scores on most measures, and were equal on two measures. On the Iowa Test of Basic Skills experimental groups scored significantly higher, $F(2, 252) = 12.65, p < .001$. On number of strategies used experimental groups used eight strategies versus two in the control group with significance, $F(2, 252) = 258.52, p < .0001$. For problem solving ($t = 11.15, p < .0001$). For providing justification for judgments Block compared scores to expected scores based on the California State Department of Education Statewide Assessment Test (1989); for the experimental group, Chi-squared was equal to 361.98, $df = 4, p < .0001$. For reflective thinking (using context cues to identify the narrator) experimental

students had 84% to 88% on the pre-test and 96% on the post-test, while the control group did not improve. Finally the researchers noted that experimental groups showed some unique characteristic traits that control students did not display: group discussion skills, peer questioning, and providing evidence.

Block's (1993) selection of the study is strong due to randomization and including a large sample of students, which most likely controls for any differences pre-study. Block does not report any test reliability or inter-rater reliability. She provides standard deviations on some measures but not on others. In the ones she does provide the standard deviations for the experimental group were larger than for the control group, which might suggest less consistent outcomes for the experimental group. However the p-values for the study are set at a very high level of confidence. The statistics may not mean much if the tests and rating are not reliable, therefore the internal validity may be compromised. The external validity is weak because it is not naturalistic for a research team to come in and perform instruction and because Block does not provide specific outcomes for different demographics, rather she gives an overall mean. None-the-less these were the outcomes given highly diverse groups of students.

Block (1993) suggested that the significant differences in scores were due to increased opportunities to use strategies in student-lead social meaning-making conversations about literature. Teachers scaffold strategy use and self-evaluation through the use of graphic organizers. Strategy instruction paired with student control of the conversation may have helped the student reflect, reason,

and problem-solve. The structure of discussion focused on social meaning making may or may not have included aesthetic reader-response and critical literacy, but it does support those practices.

Bauman & Ivey (1997) wanted to know what ethnically diverse second graders learn about reading, writing, and literature in a yearlong program of strategy instruction integrated with a literature-based environment; the students had developmental delays and behavioral disorders and came from an LSES community with high mobility rates; they found that the program resulted in increased engagement, reading levels, articulation of context and syntax cues. Students were from Athens, Georgia and lived in a high-crime neighborhood with high mobility. Out of the 19 students, 13 remained in the program for seven months; these 13 were the focus of the study. In the focus group, 75% received free or reduced lunch, eight were classified as having developmental delays in language, and two were classified as having behavioral disorders. Seven were female and six were male. Nine were African American and four were Caucasian. The teacher was a university researcher and professor on leave.

Bauman & Ivey (1997) collected personal journals by the teacher and a participating observer, student interviews, care giver interviews, videos of classroom literacy activities, artifacts of student work, assessments of student literacy learning (reading level, view of literature, engagement, word identification, fluency, comprehension, and composition, IRI analysis), and the teacher's daily plan book. Bauman & Ivey (1997) derived categories based on the data, compared them to their own theoretical positioning, conducted a

negative-case analysis to identify and reconcile discrepancies between individual students and the categories, and had an outside audit.

Bauman & Ivey (1997) found students had high engagement with books, as evidenced by children's journals about reading, diverse choices of books, knowledge of many kinds of books and authors, and criticism of books. Some students did not have growth on the IRI, but did grow in other ways—moving from no reading into enjoying mid-first grade level books. Students who struggled on the tests did well when they had an opportunity to re-read favorite books multiple times, and write their own versions or re-illustrate those texts. However, there was overall growth in reading level, as evidenced by the IRI. The 13 students developed skills in word identification and fluency, as evidenced by automatic word recognition and metacognitive answers to how they solve reading problems using reading cues. They had strong comprehension as evidenced by personal and informational meaning-making choices for books and criticism of books. They examined the purpose of various structures, how the pictures added more to the story, and how their thinking process evolved. For composition they went from listing facts at the beginning of the year to constructing a consistent sense of audience. They related elaborated descriptions of personal experiences and their interpretation of the meaning of those experiences as expressed through the style of the writing.

Bauman & Ivey's (1997) study is strong in every way with only one reservation concerning transferability. Although the detailed accounts of each student support specific transferability for students with specific mix of learning

strengths and needs, the fact that the teacher was a university professor, with an assistant teacher, a small class size, and a community of devoted parents is a rather privileged context. None-the-less these limitations suggest the many contextual factors that need to be in place for ideal implementation. The study meets the other criteria of design. The study is dependable because it builds on prior research, such as Morrow (1992) and Block (1993), and the researchers use established research techniques. Every aspect of the process is methodically detailed, containing multiple points for reanalysis, extensive examples of findings, and inclusion of an outside audit; this greatly controls for bias.

Since Bauman & Ivey's (1997) study suggests a remarkable degree of success, it is worth mentioning some of the key structures of the instruction. There was a library of over 500 trade and student published books, including culturally relevant texts and multiple genres. Strategy lessons focused on challenges that arose within specific texts. The teacher used differentiated groups for part of the day that included guided reading, repeated reading, open-ended choice half the time, and independent and group work. The teacher read aloud from multiple genres, and students read aloud using the author's chair. Texts were embedded in themes that took over the environment of the classroom and integrated writing in the content areas. The teacher used brief one-on-one conferences, and integrated skills work into the writing process, which began with pictures and journals and extended through revisions to publication. Students had many opportunities to do hands-on and memorial activities to extend their reading experience.

Bauman & Ivey (1997) identify a context that clearly integrated both aesthetic reader-response and critical pedagogy, due to the application to social context, personal development of values, and group process. The study suggests that when the teachers included strategy and skills mini-lessons, culturally relevant texts, aesthetic reader-response, and critical pedagogy, implementation of literature-based instruction was effective for strategy, skills, ownership, and meaning making. It contradicts the perspective that basic skills need to be taught in isolation before students can have higher-level engagement with literature.

In conclusion, the studies in the Strategy Instruction section suggest that strategy instruction is beneficial for skills such as fluency, word-recognition, and comprehension. However, it is ineffective if it is not integrated into a meaningful-context, wherein students are in charge of their own development. One structure that facilitated strategy instruction was group discussion of strategies to address instances where meaning broke down. Parsons (2006) and Kersten (2006) suggested that literature with discrepant events contributed to aesthetic reader-response and critical pedagogy; these studies add that discrepant events (or events where meaning-making is challenging or unsettling) also serve as a useful ground for strategy discussion between peers. This provides more support for the use of quality literature as a helpful context for strategy and skills work. Additionally, teachers supported the effectiveness of strategy instruction with opportunities for students to self-evaluate, re-experience or revise texts, receive attention in small flexible groups, and to use thematic study.

Achievement & Additional Literacy Skills

The last section suggested that literature-based instruction, with cultural relevance and strategy instruction, incorporates skills work and has a positive influence on basic skills achievement scores. Because one view on the effectiveness of instruction only takes into account test scores, it is important to examine studies that relate to how literature-based instruction affects basic skills as opposed to isolated direct instruction of the skills measured by the tests. Unfortunately, one limitation of many of these studies is that literature-based instruction is not defined, or it is defined in ways that are inconsistent with advocates' definitions. Rather literature-based instruction is treated as a packaged program that can be defined by its name only, which is inaccurate because literature-based instruction is a philosophy that can be implemented in many ways, not a program. None-the-less, it is useful to see how literature-based structures stack up to skills-based instruction as a way to bridge conversations across the controversy.

Half of these studies focus on skills such as comprehension, graphophonics, fluency, vocabulary, and syntax (the basic skills prized by the NRP meta-analysis). The other studies address concepts of print, written narrative knowledge, strategy use, and other benefits. The studies focus on diverse demographics. All studies are quantitative, but may include an additional qualitative study.

This paragraph includes a summary of each study in this section. Kamps, Wills, Greenwood, Thorne, Lazo, Crockett, Akers & Swaggart (2003) conducted

a quantitative observational study comparing the effects of three school-wide reading programs on the literacy skills of LSES, urban, and ethnically diverse students with academic and behavioral risks in first through third grade; they found that students in programs that used direct instruction of graphophonics with no literary context had the best scores on the DIBELS test. In the second study, Klesius, Griffith & Zielonka (1991) compared literature-based instruction to skills-based instruction for their effectiveness at closing the gap in reading and writing abilities of incoming semi-rural first grade students with varying SES, as measured by development of alphabetic knowledge; they found that the form of instruction had no significant impact on the achievement gap. In the third study, Juel & Minden-Cupp (2000) tested which incoming groups of African American and Caucasian first grade students, with LSES and with different levels of scores on literacy tests, benefit most from skills based versus literature-based instruction; they found that low ability groups benefited most by half a year of intensive skills instruction followed by half a year of literature-based instruction, while other groups benefited most from literature-based instruction with reduced direct skills work. In the fourth study, Morrow (1992) explored the effects of literature-based instruction on culturally diverse second grade students and confirmed that they had equal scores as skills-based treatment on the measure of the California Test of Basic Skills; she also suggested that there are additional literacy benefits for students in literature-based classrooms. Fifth, Dahl & Freppon (1995) conducted two related studies questioning the similarities and differences of inner-city K-1 children's experiences in skills-based versus

literature-based classrooms and found that growth was equal in both instructional settings; literature-based groups showed better growth in writing and additional qualitative benefits such as persistence, use of more cues, use of inquiry and critique, self-assessment of progress, and use of writing to respond to literature. In the sixth study, Purcell-Gates, McIntyre & Freppon (1995) focused just on the differences between written story-book language in skills-based versus literature-based classrooms in diverse LSES settings to see whether they would catch up to diverse well-read-to students; they found that children in literature-based classrooms were able to acquire scores that indicated a breadth of lexical and syntactic story-book language similar to the entry scores of children from well-read-to homes. Finally, Sacks & Mergendoller (1997) conducted a study comparing test scores from literature-based classrooms with embedded skills instruction to test scores in classrooms emphasizing graphophonics instruction with much less literature context and found that students who had low scores initially improved more in the literature-based classrooms, and that the engagement of all ability groups was equal in literature based classrooms whereas skills-based classrooms displayed a gap in engagement. These studies suggest there is still debate in regards to the test results of basic skills.

Kamps, Wills, Greenwood, Thorne, Lazo, Crockett, Akers & Swaggart (2003) conducted a quantitative observational study comparing the effects of three school-wide reading programs on the literacy skills of urban, and ethnically diverse students, with LSED and with academic and behavioral risks, in first through third grade; they found that students in programs that used direct

instruction of graphophonics with no literary context had the best scores on the DIBELS test. The programs were two skills-based programs and one literature-based program. They studied 383 students in a Midwestern urban area and describe the entire group as: 40% African American, 34% European American; 8% Hispanic, 7% Somolian or Sudanese, 7% Asian, and 5% unknown. They were from 5 schools with LSES demographic, but specifics are not named. Screening procedures revealed that 10% had behavioral problems, 36% exhibited academic delays, 16% had both behavioral risks and academic risks, and 38% had no risk factors.

They conducted a total of 16 DIBELS (which tests letter-naming, nonsense-word fluency, and oral reading) on two cohorts of students, over the course of the two to three year-long study. The three treatments were *Reading Mastery*, *Success For All*, and three schools with school-wide literature-based programs.

Kamps, Wills, Greenwood, Thorne, Lazo, Crockett, Akers & Swaggart (2003) found that all three curriculum significantly influenced growth in all three literacy categories. They found that programs that used direct instruction of graphophonics (without a literature context) had the best mean scores on the DIBELS for all designated risk groups, with students in the *Reading Mastery* program receiving scores 13.13% to 17.41% higher than literature-based groups at $p = 0.0001$. A HLM analysis revealed that the differences of curriculum were significant (intercept $t(377) = -1.922$, $p = 0.05$; slope $t(377) = 3.607$, $p = 0.001$;

and acceleration, $t(377) = 3.800$, $p = 0.0001$. They did not find the interaction between curriculum and risk factors to be significant.

Although the statistical values of Kamps, Wills, Greenwood, Thorne, Lazo, Crockett, Akers & Swaggart's (2003) study have high confidence levels, there are problems regarding differential selection of the subjects; these problems diminish the internal validity. Because statistics on the backgrounds of students are provided for all students in the study, rather than for each classroom, there is no way to tell if there were great differences in the demographics between classrooms/ treatments, and it harms the validity of any comparison. The one school using *Reading Mastery* as a stand-alone program was a charter school, while all other schools in the study were public schools. Charter schools are known to include students whose parents are more motivated to be involved in their child's education because inclusion in the school is based on making the effort to enter a lottery; parental involvement positively impacts achievement (Ravitch, 2010). There were no statistical tests at the beginning of the study to ensure that all treatment groups had the same incoming scores on the DIBELS. This combined with a lack of transparency about between school differences in demographics is significant enough to reject the validity of the study.

Additionally, because there is little to no description of the practices of any of the classrooms the external validity of Kamps, Wills, Greenwood, Thorne, Lazo, Crockett, Akers & Swaggart's (2003) study is compromised. Since *Reading Mastery* and *Success for All* use scripted lessons that are the same in every school one can look up the program and discover the method of

instruction. But literature-based programs are not uniform and they vary greatly. The study says that they used literature-based basals. This fact alone contradicts what many educators would define as literature-based instruction, which often uses trade books, and student-selected reading. Additionally, they do not mention if teachers gave the students embedded support in graphophonics and reading strategies, which many educators include in a definition of literature-based instruction, but some do not. In studies on literature-based instruction it is imperative to describe the instruction teachers in the study practice in order to be able to draw conclusions from comparisons, and in order to generalize the results.

Overall, Kamps, Wills, Greenwood, Thorne, Lazo, Crockett, Akers & Swaggart's (2003) study suggests that diverse urban students with varying academic and behavioral skills benefit from direct instruction in graphophonics. Because literature-based instruction is not defined, and selection may have been differential between groups the comparative results are not noteworthy. However, the study could be supported by other studies.

Klesius, Griffith & Zielonka (1991) compared literature-based instruction to skills-based instruction for their effectiveness at closing the gap in reading and writing abilities of incoming semi-rural first grade students with varying SES, as measured by development of alphabetic knowledge; they found that the form of instruction had no significant impact on the achievement gap. The subjects were 112 first-grade students from two schools in a semi-rural area of Florida. In the school using skills-based instruction 22% of students were eligible for free and

reduced lunch, in the school using literature-based instruction 40% of students were eligible for free and reduced lunch. In their year-long quantitative observational study they used ANOVAs to compare the growth in alphabetic knowledge in three skills-based classrooms to three literature-based classrooms by conducting pre-tests and post-tests with the Gough-Kastler-Roper Phonemic Awareness Test (PAT) (1984), a Features Spelling Test (FST), and an activity involving writing a story about a picture. The skills-base classrooms had long been using skills-based instruction; the literature-base classes were in their first year of literature-based instruction.

Klesius, Griffith & Zielonka (1991) found that it makes no difference if phonics only is used or if literature-based instruction is used; all students improve on testing by the same degree. For PAT, $F(5, 5) = 4.114$, $p = .112$; for FST, $F(5, 5) = 2.348$, $p = .200$; and for writing, $F(5, 5) = .945$, $p = .403$. The achievement gap between high and low scoring students remains with either instructional method, except for writing. In the literature-based groups, for PAT, $F(2, 3) = 34.10$, $p = .0000$; for FST $F(2, 3) = 22.21$, $p = .0000$; and for writing, $F(2, 3) = 3.66$, $p = .0622$. In the skills-base groups, for PAT, $F(2, 3) = 42.35$, $p = .0000$; for FST $F(2, 3) = 20.86$, $p = .0000$; and for writing $F(2, 3) = 5.49$, $p = .0229$. This suggests that literature-based instruction does no harm overall for students in its first year of implementation, and that it may marginally close the gap in scores for writing as compared to skills-based instruction.

The design of the Klesius, Griffith & Zielonka (1991) study is fairly strong given that they control for differences in selection by measuring growth in scores

rather than final scores, but they do not supply reliability scores for the two researcher-created tests, and the PAT test only had $r > .7$, when a score over 80% is preferable. The statistical report on closing the achievement gap is strange due to the low p-values reported (.0000), when some of the f-values do not exceed critical values at the .01 confidence level, 30.82; at the .05 level they are all (except literature-based writing) very significant.

The external validity of this study is strong for semi-rural populations with some SES diversity. Klesius, Griffith & Zielonka (1991) do define literature-based instruction in the same terms as this paper, and conducted informal observations to ensure the practices were upheld; this makes the study generalizable for accurate literature-based practices in the first year of implementation.

The meaning of the Klesius, Griffith & Zielonka (1991) study for this paper's central question is that literature-based instruction, when it implicitly includes skills work, may be equally effective for teaching skills, and equally ineffective at closing the achievement gap on alphabetic principle. Again, there is no mention of aesthetic reader-response or critical literacy within the literature-based instruction. The format of interactions and real-life uses for reading is present, although the researchers do not mention group discussions of different interpretations of meaning, and strategy instruction.

Juel & Minden-Cupp (2000) tested which incoming groups of African American and Caucasian first grade students, with LSES and with different levels of scores on literacy tests, benefit most from skills-based versus literature-based

instruction; they found that low ability groups benefited most by half a year of intensive skills instruction followed by half a year of literature-based instruction, while other groups benefited most from literature-based instruction with reduced direct skills work. The subjects of the study were from a city in the Southwest and attended two schools with 70% free and reduced lunch, and had classrooms of 18 or fewer students wherein students were grouped into low, medium and high ability groups within the class. At each school 60% of students were African American, 36% were Caucasian, and 4% were from other ethnic groups. They conducted a year-long quantitative comparative observational study. For pre-tests and post-tests they triangulated the BBELS (which tests word recognition strategies, oral reading, and comprehension) the WRAT (which tests word recognition strategies) and a story-reading test (which tests for word-recognition strategies with a think-aloud procedure). By observing classrooms they developed codes for the instruction in each group. They included four groups of students in the study, each receiving a different mix of skills and literature-based instruction.

Overall, Juel & Minden-Cupp (2000) found that students who had low initial scores had the highest scores on post-test when they received half a year of intensive skills-base instruction followed by half a year of integrated literature-based instruction. Using an ANCOVA on the BBELS post-test they found significant difference among the four classrooms, $F(3, 50) = 6.60, p < .001$. They followed this by pairwise tests using Bonferroni adjustment and found that the mean scores of the classes on the BBELS post-test were highest in the

classroom that differentiated instruction with a mix of skills and literature-based methods based on designated ability level. By comparing expected scores to observed scores on BBELS post-tests, they found that students who had low scores on the pre-tests met end-of-year expectations when they received half a year of intensive skills-based instruction, followed by half a year of integrated literature-based instruction (16% skills emphasis), $M = 91.3\%$, $SD = 5.8$ on the end of first grade level, $M = 94\%$, $SD = 3.5$ on the primer test. They fell far below post-test expectations when they received literature-based instruction that included almost no skills instruction, with no scores on end of first grade test due to frustration, and on the primer $M = 80.7\%$, $SD = 17.1$. Students who scored medium and high on the pretest benefited most from literature-based instruction integrating only 6% emphasis on skills, exceeding expected scores by one standard deviation.

The internal design of the Juel & Minden-Cupp (2000) study was strong, however, omission of certain key details in the statistical report make the findings hard to evaluate, and therefore it has questionable validity. The use of pre-tests and post-tests with comparison treatments is a strong design, and ensured that subjects in every classroom had the same mean incoming scores on literacy tests. They selected the classrooms through identifying similar demographics, similar literacy structures of the schools, and administrators' suggestion of effective teachers; all of this helps control for selection differences and researcher bias. The use of three similar tests on the pre-tests suggests a way to determine reliability of the tests, but they do not report those statistics, and

results are only given for the BBELS. However, they do say that BBELS has otherwise been determined to have test re-test reliability of .78 to .91, which is good. Observers were all reading specialists and had strong inter-rater reliability, $r = .94$. Likewise, two reading specialists independently coded the observation narratives and had 97% inter-rater reliability. As mentioned previously, all of this suggests a very careful and strong internal design.

However, Juel & Minden-Cupp (2000) omit numerical details such as subject attrition, the number of students in each ability group within the classroom, and whether ability groups were flexible or fixed throughout the year, and this contributes to clouded statistical reporting. They do not report the pairwise statistics of the follow-up to the ANCOVA that was used to rank the mean effectiveness of each treatment. Without knowing the sample number, means, standard deviations and the selected p-value reader's cannot evaluate the confidence level of the most important findings. The standard deviation of the primer test for the designated low group receiving literature-base instruction with only 6% skills emphasis, was quite large, without a mode or range it is hard to say if the choice not to give the end-of-first grade test to at least some of those students was really valid.

In short, because Juel & Minden-Cupp (2000) had a strong design, and they set their p-value at a high level (99.9%) for the ANCOVA their study has initial credibility and one can be inclined to give them the benefit of the doubt on the incomplete parts of the statistical reporting. If other studies come to similar findings, with clear statistical reporting then it will lend the findings of this study

credibility. The external validity of this study is strong for classrooms that use ability grouping and have high poverty rates and high numbers of African American students. Also teachers were accustomed to their style of instruction and researchers did not coach them, which presents circumstances that eliminate the reduced generalizability that a switch in instruction presents.

The significance of Juel & Minden-Cupp's (2000) study is that it suggests that literature-based instruction is most beneficial for early readers once they have developed strong graphophonics and word recognition strategies. If the instruction does not integrate a skills emphasis then it may influence poorer scores on word recognition as a result. The definition of literature-based instruction Juel & Midden-Cupp used does not match the standard definition among literature-based instruction advocates, because skills were not by definition included in interactions with literature, thus contributing to the inappropriate dichotomy that is prominent in research. The limitation of this study is that it is not clear how a year-long program with explicit graphophonic and strategy instruction embedded in literature study would have benefited students. In terms of the central question of this paper, regarding the effectiveness of aesthetic reading and/ or critical pedagogy for diverse learners, these approaches may or may not have been used in the discussion groups that were used in the literature-based instruction, the study does not go into enough detail to know. Certainly, the structure was such that aesthetic reader-response or critical pedagogy may have been used, and that structure was found to be

beneficial for all students' word recognition scores, as long as it included strong emphasis on graphophonic skills as needed.

Morrow (1992) explored the effects of literature-based instruction on culturally diverse second grade students and confirmed that they had equal scores as skills-based treatment on the measure of the California Test of Basic Skills; she also suggested that there are additional literacy benefits for students in literature-based classrooms. The subjects were 166 second-graders from nine classrooms in two schools; 72 were African American, 62 were Caucasian, 23 were Asian-Pacific Islander, and 9 were Cuban or Puerto-Rican, with equal distribution per classroom; 24% were eligible for free or reduced lunch. In this eight-month quasi-experimental study researchers trained teachers in the two experimental groups to implement literature-based instruction; the control group continued with skills-based instruction. Randomly selected focus students were given pre and post-tests, analyzed by ANCOVAs.

Morrow (1992) found no significant differences on the standardized test, $F(2, 5) = .02, p < .10$. The literature-based classrooms scored significantly higher on all additional tests. Comprehension: story retelling, $F(2, 5) = 12.11, p < .01$; story rewriting, $F(2, 5) = 10.75, p < .006$; Probed Recall Comprehension, $F(2, 5) = 9.24, p < .03$. Use of story conventions in original creations of stories orally: $F(2, 5) = 9.87, p < .02$, and in writing, $F(2, 5) = 9.70, p < .01$. Vocabulary: in written original story, $F(2, 5) = 5.21, p < .05$, and in oral original story, $F(2, 5) = 6.46, p < .01$. Complexity of Syntax: in the written original story, $F(2, 5) = 9.79, p < .05$, and in the oral original story, $F(2, 5) = 4.01, p < .05$. Additionally, they

found no differences in mean scores or standard deviations between African American and Caucasian students. Other collections of data found that children in literature-based classrooms reported more reading after school on two single days, $F(2, 5) = 4.14$, $p < .01$, and that they could identify 28% to 60% more authors, illustrators and book titles.

The internal design of the Morrow (1992) study and the significant f-values are questionable. A design that uses pre-tests and post-tests with comparison groups, and random selection of focus students, controls for many variables, and Morrow adjusted post-test scores for pre-test differences. All the tests had reliability of $>.90$, which is very good. Morrow does not say how access was gained to the teachers or if there were any mortality; mortality would be a big issue given that there are only nine focus students per group. These issues leave room for bias and may harm internal validity. Morrow reports f-values for the .01 and .05 levels of confidence that she identifies as significant, however they sometimes do not exceed the critical f-value at the p-value. At .01 the critical value is 13.27, and at .05 it is 5.79; therefore almost none of the statistics stand up to the .01 confidence level, but most stand up to the .05 confidence level. The fact that she reports such significance harms the validity of the study. Even so, the trends she identifies are interesting.

Morrow (1992) compromised the external validity of her study by providing ongoing coaching to teachers throughout the study. Such coaching is often not present, but it does ensure that the instruction given matches the definition of literature-based instruction in Chapter 1 of this paper. The student

sample was diverse, and no differences were found between ethnicities for the effectiveness, so the study is generalizable for several cultural backgrounds.

Although the design and significance of the statistics in the Morrow (1992) study left room for questions, the study does point out an interesting direction for research: if literature-based instruction matches skills-based on discrete skills, does it offer any additional benefits for other areas of literacy? The description of the literature-based instruction included culturally relevant texts, and opportunities for shared discussion, retellings, reenactments and rewritings of stories. These activities create a strong framework for the presence of aesthetic reader-response and critical pedagogy, even if they are not directly mentioned.

Dahl & Freppon (1995) conducted two related studies questioning the similarities and differences of inner-city K-1 children's experiences in skills-based versus literature-based classrooms and found that growth was equal in both instructional settings; literature-based groups showed better growth in writing and additional qualitative benefits such as persistence, use of more cues, use of inquiry and critique, self-assessment of progress, and use of writing to respond to literature. Their study focused on the categories of persistence, concepts of print, story structure, written narrative register, concepts of writing and sense-making strategies, alphabetic knowledge, and syntactical knowledge. With such a large scope of literacy abilities, this study goes well beyond the basic literacy skills focused on in the NRP analysis, and it contributes more breadth to this paper's analysis of the kinds of literacy outcomes of literature-based instruction.

The subjects of Dahl & Freppon's (1995) study were 48 (12 focus) K-1 students from two Midwestern cities; most students received public assistance, and attended schools with high mobility rates. Half the students were Appalachian Caucasians and half were African American. Dahl & Freppon conducted a quantitative and a qualitative study, following the subjects for two years. The comparative quantitative study followed a pre and post-test model. They used repeated 1 X ANOVAs comparing each focus child, followed by a 2 X ANOVA to compare outcomes for each instructional group. Dahl & Freppon used the qualitative study to develop patterns through a twice weekly observational process, focusing on a different child each session. They audio-recorded student utterances during instruction, probed for metacognition, conducted miscue analysis, and collected artifacts of student work in both settings.

In the quantitative study, Dahl & Freppon found no significant difference in growth between instructional groups for concepts of print, story structure, alphabetic principle, and concepts of writing. Literature-based classrooms had significantly higher post-scores on written narrative register ($F(1, 2) = 27.95$, $p < .05$).

In the qualitative study, Dahl & Freppon found differences in the cueing systems student's used in each setting, and differences in patterns of learner sense-making in each setting. Miscue analysis revealed that the skills-based students relied on graphophonic cues, while students in the literature-based classroom used semantics, syntax and graphophonics, increasing their ability to

read independently. Skills-based students passively showed enjoyment of read-aloud time, answering teacher-led questions with facts from the story. Literature-based students responded to literature by making predictions, questioning, stating opinions, discussing intertextual relationships, critiquing the story, making suggestions based on real life experience, acting out interpretations, and rewriting stories to change language or plot. Literature-based students had more persistence with independent reading, often coping with difficulty by choosing to make use of discussions with peers and collaborative reading. Skills-based students completed assigned work but had much shorter persistence with independent reading, abandoning books after a page or two, or relying the teacher, or copying to get the correct answer for worksheets. Literature-based students frequently made comments about their progress, difficulties, and their plans for what to do next. Skills-based students completed assigned work with no comments about their own learning process, and more comments about accuracy. Literature-based students used writing to explore texts and mimic environmental print; they worked on the sentence, paragraph and story levels of writing. Skills-based students used writing primarily on worksheets and were only able to write on the sentence level on a mid-year writing event.

Concerns about selection threaten the validity of the internal design of the Dahl & Freppon (1995) studies, however, illustrative examples and an outside audit give the qualitative study credibility. The quantitative model of pre and post tests analyzing growth controls for initial differences in the sample. The controlled for selection bias in three ways: demographic matches, consistency

with instructional approach as evidence by the Theoretical Orientations to Reading Profile used in other studies, and random selection of focal students. Bias may have entered because Dahl & Freppon do not name the initial access to the sites, or the experience levels of the teachers. It is also suspect that only seven students were included in the skills-based group, while 12 were in the literature-based group; this could greatly skew both studies results. The critical value of the ANOVAs was set at the .05 level, which has less confidence than the .01 level in determining significance. No inter-rater reliability or rotation of observation sites is present, and Dahl & Freppon supply no reliability statistics for the measures. All of these factors greatly reduce the validity of the quantitative analysis.

In the qualitative study Dahl & Freppon (1995) describe the data collection process, and how categories were derived by individual researchers, and debated as a team, and confirmed by two outside audits. They provide selections from the transcriptions to illustrate the differences in instructional settings. These moves make the qualitative study both credible and confirmable.

The external validity and the transferability of Dahl & Freppon (1995) are for low-income Appalachian and African American students. The school and community context could be stronger for the qualitative study in order to know what other support networks may have influenced the culture of the classrooms.

Despite the issue with the design of the Dahl & Freppon (1995) quantitative study, it does concur with Morrow (1992) and Klesuis, Griffith & Zeilonka (1991), which strengthens the ongoing trend in the research, suggesting

that there are no differences in skills like alphabetic knowledge and syntax usage. It does contradict Juel & Minden-Cupp (2000) and Sacks & Mergendoller (1997) who found differences in achievement in favor of integrated literature-based instruction. The qualitative study provides insight as to which classroom structures support active response, and use of literature for real purposes. The teacher emphasized use of all cueing systems, centered book talk on meaning-making, and provided ample opportunities to collaboratively or independently extend their interaction through theatre, writing, play, and discussion. Additionally, the teachers put emphasis on creating independent readers more so than accurate oral reading. Such an environment had readers who were persistent problem-solvers and planners of their own development. Taken along with Morrow (1992) this suggests that literature-based instruction may help children be inspired to integrate reading into their whole lives, both in and out of school. The study certainly confirms the recommended structures in all previous sections of this chapter.

In the same year Purcell-Gates, McIntyre & Freppon (1995) focused just on the differences between written story-book language in skills-based versus literature-based classrooms in diverse LSES settings to see whether they would catch up to diverse well-read-to students; they found that children in literature-based classrooms were able to acquire scores that indicated a breadth of lexical and syntactic story-book language similar to the entry scores of children from well-read-to homes. The subjects of this two-year quasi-experimental study were K-1 students from three groups. The base-line group of subjects were from an

urban area in northern California, they were racially and ethnically diverse, from MSES to LSES backgrounds, spoke English as a first language, were not yet independent readers; they had been read to extensively at least once a day since six months of age. The second group were LSES students from a Midwestern city; 67% were African-American and 33% were Appalachian Caucasians given skills-based instruction. The third group were of LSES from a Midwestern city and were 46% African-American and 54% Appalachian Caucasian and received literature-based instruction. All groups were gender-balanced.

Purcell-Gates, McIntyre & Freppon (1995) administered pre, and post tests of pretend book reading; they were scored based on their intonations. The data from these measures was then compared using weighted ANOVAs to account for the unbalanced design. They used independent t-test to follow up with pairwise contrasts.

The initial test revealed that the well-read-to group scored significantly higher on written registers than the other two groups, and that there were no significant differences between the skills-based group and the literature-based group. Purcell-Gates, McIntyre & Freppon (1995) found that after two years there were significant differences between each of the experimental groups in their relation to the well-read-to group's scores on breadth of written narrative register, with the literature-based group making significantly more gains than the skills-based group, $t = 3.06$ with $p = .007$; neither skills-based nor literature-based group differed significantly from the beginning of kindergarten scores of the well-read-to group. Children in literature-based classrooms were able to

acquire scores that indicated a breadth of lexical and syntactic storybook language similar to the entry scores of children from well-read-to homes.

The internal design of this study is strong. Purcell-Gates, McIntyre & Freppon (1995) should have provided information as to how much the experimental groups had been read to prior to school entry, in order to see if the prior experience with reading or one of the many other differences (such as location, SES, preschool experience etc.) accounted for initial score differences with the well-read-to group. The researchers did not name how they gained access to the subjects, however for the experimental groups, they assessed the Theoretical Orientation to Reading of the teachers, students were randomly selected, and administrators rated all teachers as average. The test did have strong inter-rater reliability, $r = .99$. They hold the statistical measure to a high level of confidence: .007, however the standard deviations are very large for both groups, 12 to 14, out of 18 or 19 samples; it is therefore hard to know if the means were really indicative of the group trends without a median or mode. There were no significant differences between the well-read-to group's entry scores and skills-based or literature-based classes, which shows that although the difference between the experimental groups is significant, in the end, both groups caught up to the entry scores of the well-read-to group.

The Purcell-Gates, McIntyre & Freppon (1995) study is externally valid for the selected urban LSES populations, who may or may not have been read to. The researchers do not name other contextual factors such as the support

systems of the schools, families, and communities. The definition of literature-based instruction matches the Chapter 1 definition of this paper.

Purcell-Gates, McIntyre & Freppon's (1995) study suggests that the literature-based classroom approached the breadth of written narrative knowledge that well-read-to students had due to ongoing interactive engagement with wider variety of texts, at least as much as skills-based classrooms, if not more. There are three studies in this paper so far that suggest children in literature-based classrooms may have greater written language knowledge than those in skills-based classrooms. The instructional format involved high amounts of discussion, predictions, interactive interpretations, and multiple genres of trade books. These formats are foundational to the presence of aesthetic reader-response and critical pedagogy, even if these forms of instruction are not explicitly mentioned.

Sacks & Mergendoller (1997) conducted a study comparing test scores from literature-based classrooms with embedded skills instruction to test scores in classrooms emphasizing graphophonics instruction with much less literature context and found that students who had low scores initially improved more in the literature-based classrooms, and that the engagement of all ability groups was equal in literature based classrooms whereas skills-based classrooms displayed a gap in engagement. They asked two questions. Which instruction affected higher test scores for students who scored above or below the mean? Which instruction had more highly engaged students with the respective above mean and below mean scoring groups? Their subjects were 132 kindergarteners in 11

Marin County, Californian classrooms; 75% were Caucasian, 11% were African American, 6% were Asian-Pacific Islander, 5% were Hispanic and 2% unknown. The SES is not stated, however census statistics report that Marin County is affluent. They conducted a year long, quantitative observational comparison between six skills-emphasis classrooms, and five literature-based classrooms with integrated skills instruction. Students were given the TERA-2 for pre and post-test (which measures concepts of print, alphabetic knowledge, print awareness in different pragmatic contexts, and knowledge of relationships between vocabulary). Researchers observed the classrooms on a rotating schedule and used pre-established codes to mark instruction characteristics and engagement.

Using a 2 X ANOVA, Sacks & Mergendoller (1997) found that students who had pretest scores below the mean and were given literature-based instruction with embedded skills instruction improved their scores by about 10 points on average on the Test of Early Reading Ability 2 (TERA-2), whereas students in all other designated groups and instruction formats only improved by six to eight points, $F(1, 9) = 5.00, p = .05$. The differences between instruction for students who scored above the mean were not significant. The 2 X ANOVA for engagement revealed significant differences between ability groups, with students who scored above the mean on the pre-test having more engagement overall $F(1, 9) = 7.30, p = .02$. There was marginal interaction between instruction, pre-test score grouping, and post-test scores. Students in the two ability groups had almost equal engagement in literature-based classrooms, but

students in skills-based classrooms had a gap in engagement between ability groups, with students who scored below the mean on the pre-test being less engaged, $F(1, 9) = 4.78, p = .06$.

Overall the internal validity of the study is good and the findings may or may not be significant due to the lower confidence levels Sacks & Mergendoller (1997) used. They do not name how they gained access to the teachers, but they selected subjects based on teacher ratings of experience with reading followed by scores on the TERA-2 in order to create two ability groups among 48 target students. Teachers were selected based on a Theoretical Orientation to Reading measure, but had great differences in the number of years of experience teaching. These factors may leave some room for biases to enter into the selection. Sacks & Mergendoller do not provide data on subject attrition. The TERA-2 is a reliable test, with $r = .89$. The method of counterbalancing observations and obtaining a .84 inter-rater reliability is very reliable. If researchers had used a confidence level of .01 instead of .05 their statistics would be more significant. F-values below five may or may not be significant. Sacks & Mergendoller do not provide a range or mode for the means, which also harms a reader's ability to evaluate the findings. It is not a newsworthy study on its own, but additional studies may lend it validity.

What does make Sacks & Mergendoller's (1997) study outstanding is its detailed definition of terms, and detailed coding of actual practices that occurred in the classroom. This greatly supports the external validity of the study. Their definition of literature-based instruction matches the definition in Chapter 1 of this

paper. The generalizability is limited because subjects were not very diverse and mostly likely came from affluent backgrounds.

In terms of the central question of this study, Sacks & Mergendoller's study does contribute to the review of effectiveness of aesthetic reader-response and critical pedagogy. Based on the findings of the coding, students in the literature-based classrooms spent significantly more time interacting with books with others ($F(1, 9) = 8.86, p = .016$) and more time writing ($F(1, 9) = 10.22, p = .01$), using invented spelling ($F(1, 9) = 7.21, p = .025$), and dictating stories ($F(1, 9) = 4.17, p = .07$). All of these activities create avenues for children to do something with what they read or respond to the meaning with others, which are essential aspects of aesthetic reader-response and critical pedagogy. In the case of a test that measures concepts of print and pragmatics, students in such settings scored better and may have been slightly more engaged.

The Achievement & Additional Literacy Skills section of studies presented the most controversy among the findings. One interesting trend suggested that students in literature-based classrooms may have scored equally on standardized tests, but had higher scores on alternative testing. Kamps, Wills, Greenwood, Thorne, Lazo, Crockett, Akers & Swaggart (2003) and Juel & Minden-Cupp (2000) were the only two studies that found literature-based instruction to be less effective on standardized tests, however their definitions of the instruction are inconsistent with the definitions of the other studies in this chapter. All studies in which literature-based instruction included skills and strategies work found that literature-based instruction was at least equally

effective on basic skills, and added more benefits in writing, active engagement, comprehension, persistence, strategies, and voluntary reading. The limitations of this group of studies are that most of them are older studies (most likely due to the greater amount of funding during the era of the “reading wars”), the studies do not define further distinctions in the literature-based instruction, and some of the studies do not have accurate definitions. When one synthesizes these studies with the other studies in this review of research, one can arrive at a hypothesis that there are mixed findings because the literature-based instruction may or may not include the structures that make it most effective.

Summary

Chapter 2 was a review of the research on literature-based instruction with a focus on aesthetic reader-response and critical pedagogy. Chapter 2 contained analyses and critiques of the findings of the studies for credibility and validity. The purpose of the chapter was to explore the benefits of aesthetic reader-response and critical pedagogy for diverse elementary school students.

The research in the Aesthetic Reader-Response section indicated that the process of aesthetic response was a natural way to deepen cognitive processes with reading, generate deeper understanding, take ownership over a meaning-making process, and understand ones self in context of a diverse community. The Critical Pedagogy section suggested that students developed higher-level cognitive processes, articulated their voice in response to systems of power, and took ownership of literacy as an important way to develop identity. The Group Discussion section found that given open-ended student-led formats, students

used the discussion as a way to share reading strategies and problem-solve any break downs in meaning-making; they developed higher-order cognition as they made elaborative comments, integrated alternative views, made textual reference, and tested hypotheses. The Motivation and Engagement section indicated that student engagement was high when they could collaboratively self-assess, make their own choices, and transact with culturally relevant texts. The Best Practices section found that the structures indicated in the previous sections corresponded to effective practice, and added additional suggestions. The Strategy section suggested that literature-based instruction contributes to perseverant and independent readers when it includes metacognitive strategy instruction. Finally, the Achievement section indicated that literature-based instruction might result in equal improvement on standardized tests as skills-based instruction, and offer additional literacy skills.

CHAPTER 3—CONCLUSION

Introduction

Chapter 1 introduced the controversy and historical background of aesthetic reader-response and critical pedagogy. It posed the questions of how these processes further distinguish quality literature-based instruction, and how they benefit students' literacy. Chapter 2 reviewed the research on literature-based instruction, with a focus on aesthetic reader-response, critical pedagogy, group discussion, and strategy instruction, and noted how it affected the motivation and literacy skills of diverse elementary school students. This section concludes the paper by synthesizing the findings of the research. It then discusses implications for instructional practice and suggests directions for further research.

Summary of the Findings

In summary of the review of research in chapter 2, a few themes and questions arise. This section discusses the implications for the controversial questions of how literature-based instruction benefits basic skills, and how useful it is for elementary school readers and diverse demographics of students. Then it summarizes the findings regarding the benefits of aesthetic reader-response and critical pedagogy, and the structures that support these processes.

A problem in the controversial arena of comparative quantitative studies on basic skills achievement is that literature-based instruction is not defined according to its theoretical foundation, as evidenced in Kamps et al (2003) and Juel & Minden-Cupp (2000). Therefore, the question still remains as to whether

literature-based instruction that includes aesthetic reader-response, critical pedagogy, group discussion, and strategy work is as effective for basic-skills achievement scores as skills-based instruction.

However, the large body of studies related to defining high-quality literature-based instruction largely confirms each other, creating dependability, even if many are lacking full credibility. One hypothesis that arises is that the structures that teachers use to facilitate aesthetic reader-response and critical pedagogy processes are the structures that define high-quality literature-based instruction. Some studies in the Best Practices section associated these structures with high achievement scores on district tests, such as Taylor et al. (2000), Morrow et al. (1999), and Mosenthal et al. (2004).

Concerning the controversial questions about how useful aesthetic reader-response and critical pedagogy are for elementary students, and for students of different demographics and experience with reading, the studies made a strong suggestion. In the limited scope of studies here, students from diverse demographics benefited from these processes, and had the ability to use the processes to own and develop their own literacy purposes. These diverse demographics included: multi-lingual, multi-racial and multi-cultural backgrounds; all socio-economic backgrounds; urban, suburban, and rural backgrounds; and students with special needs, different abilities, and different home cultures of reading. Some studies found that students developed their own literacy purposed at the same time as developing their basic skills, those studies include: all of the Best Practices & Strategies studies, and Goldstein (1995), Goatley,

Brock & Raphael (1993), Davis (2010); Sulzby, Branz & Buhle (1993); Dahl, (1993); Dahl & Freppon (1995), and Sacks & Mergendoller (1997). One study suggested that inclusive heterogeneous groups of diverse learners enriched the implementation of aesthetic reader-response, critical pedagogy, and group discussion (Goatley, 1993). These findings contradict the idea that basic skills need come before higher-levels of engagement with literature, or that remedial education should only focus on basic skills.

In summary of the benefits for students, the next few paragraphs will state the findings for each section and the connections between the findings of all sections.

Both the Aesthetic Reader-Response section and the Critical Pedagogy section found that aesthetic reader-response and critical pedagogy included empathizing and living through the experiences of characters in a book, relating those experiences to one's own (Cox & Many, 1992; Eeds & Wells, 1989; Parsons, 2006; Souto-Manning, 2009), and clarifying one's values in context of other perspectives through collaborative discussions and writings (Arce, 2000; Flint & Laman, 2012; Hoffman, 2010; Kersten, 2006; and Lohfink & Loya, 2010), Kersten (2006) and Parsons (2006) found that discussions and writing that focus on discrepant events in the literature are most conducive to the processes of aesthetic reader-response and critical pedagogy. The processes of visualization, creating metaphors, re-writing, re-enacting, and memorializing literature, deepened students' ability to experience and voice a response to the literature (Parsons). Response journals and discussions were common structures.

Students who engaged in aesthetic reader-response and critical pedagogy had deeper understanding of the text and used higher-order thinking (such as interpretation, using textual references, understanding the pragmatic and social context, posing questions, negotiation, and critique) (Cox & Many; and Kersten). They also exhibited ownership of the reading process for their own life, and the ability to consider multiple perspectives. Goldstein (1995) suggested that in the process of making-meaning students also shared strategies and developed skills.

Students who engaged in critical pedagogy had some additional benefits. All critical pedagogy studies found that the use of texts that are culturally relevant, social justice related, or focus on problems that students experience in their lives was central to this instruction. All studies found that in the process of experiencing literature and clarifying values, students discovered their own views in context of socially unjust systems of power and privilege. In some studies they then integrated content area work as they formed plans of action to improve awareness and conditions within their community (Goldstein, 1995; and Souto-Manning, 2009). In this case, literacy became a way of examining the world and changing ones actions in the world. Interestingly, one can infer that critical pedagogy benefited from aesthetic reader-response, but also went beyond it if students decided on further research and plans of action for social justice.

The Group Discussion, Motivation & Engagement, and Best Practices sections suggested that the same structures teachers use to facilitate aesthetic reader-response and critical pedagogy, resulted in deeper cognitive processes, higher motivation, and effective literacy instruction. As a whole, the Group

Discussion section suggested that choice, student led talk, writing, structured reflection on group discussion processes, and explicit coaching on routines, rules, and expectations were key elements of quality implementation. Like aesthetic reader-response and critical pedagogy, all Group Discussion studies found that quality group discussions involved considering multiple perspectives, using textual reference, posing questions, and sharing hypotheses. One study found that justifying ideas, elaborating ideas, and critiquing texts contribute to deeper comprehension and cognition (Chinn, Anderson & Waggoner, 2001).

As a whole, the Motivation & Engagement studies found that student choice, collaborative goal setting, self-assessment conferences, hands on elaborations of texts, group discussions, and ample attention to experimenting with multiple reading strategies all positively influenced motivation to read. Such structures are consistent with aesthetic reader-response and critical pedagogy.

The Best Practices in Literacy Instruction section confirmed many of the findings mentioned previously, and all studies added that strategy instruction within meaningful literature experiences is a major qualification of effective literacy instruction. In addition, it recommended using multiple reading venues (Taylor et al, 2000; Morrow et al., 1999; and Wharton, 1998) and use of students' home-languages in initial responses to texts (Gersten, 1996).

The Strategy Instruction section found that the many structures that support meaning making should be used as a context for strategy instruction on the cueing systems and the reading process (Bauman, 1997; Taylor et al., 2003). The instruction must put meaning at the center rather than just naming the

strategy (Duffy, 1993). As a result students improved on comprehension, fluency, vocabulary and syntax (Bauman, and Block, 1993). They also had better discussion skills, problem-solving skills, more motivation, and more confidence in themselves as independent readers (Block and Taylor et al., 2003).

Finally, the Achievement and Additional Literacy Skills section suggested an ongoing debate. All studies suggested that skills instruction benefits achievement in basic skills, but they disagreed on whether a literature-based context was as effective or more effective than skills-based instruction. The findings for the debate are inconclusive because literature-based instruction was not defined the same in each study. Studies that found literature-based instruction to be equally effective on basic skills (and more effective on other skills) described structures similar to the structures that the rest of this paper names: connecting life experiences, re-enactments, strategy instruction, discussions with multiple perspectives, themes with multiple genres, and integration of reading and writing (Dahl & Freppon, 1995; Morrow, 1992; Purcell-Gates, McIntyre & Freppon, 1995; and Sacks & Mergendoller, 1997). Teachers would use these structures when facilitating aesthetic reader-response and critical pedagogy. Some of the additional benefits for students related to higher scores on written narrative register (Dahl & Freppon; and Purcell-Gates, McIntyre & Freppon) and longer persistence with challenging texts (Dahl & Freppon).

Implications for Classroom Practice

If teachers are interested in helping students own their own development as readers, and use literacy for meaningful thinking about the self and the world,

then aesthetic reader-response and critical pedagogy are good process to cultivate. This section synthesizes the findings of the research and makes suggestions for practice. The research in this paper suggests that meaningful literature-based contexts are an essential component of effective literacy instruction. This paper specifically explored how aesthetic reader-response and critical pedagogy are ways of approaching texts in order to make meaning that is relevant to real life. This section will suggest ways to implement the structures that these processes rely on and other aspects of instruction that would make these processes more successful for all students. The following synthesis addresses activities, discussion structures, strategy instruction, and the environmental context of the classroom. It then summarizes findings that are specific to some demographics of students.

Many activities contribute to the meaning-making processes of aesthetic reader-response and critical pedagogy. When students have a chance to activate their prior experiences before reading, they prepare to connect with the experience they encounter in the text. Students can articulate their experiences by telling the moving picture of the mind, illustrating, using response journals, re-enacting the story, re-writing the story, and memorializing the story in a public way. They can explore questions or break downs in meaning through discussions. They can use language charts to call attention to stylistic, genre, or syntactical structures of the works. They may further elaborate by discussing multiple perspectives with peers, or using the textual topic, language and structures of a text as they write and publish their own literature. Using a critical

pedagogy process students can engage in critiquing the injustices they experience, or that they saw while reading; from there they can do further research in content areas and the teacher can look for opportunities to facilitate their plans of action.

Aesthetic reader-response and critical pedagogy processes benefit from group discussion, and the studies made some recommendations for how to implement high quality discussions. The best group discussions involve a developmental process in creating a culture of literature talk. The developmental process includes democratically decided rules, roles and expectations. The teacher and the students can keep a reference list of the kinds of things to consider commenting on, such as style, genre, re-tellings, personal experiences, empathy, social critique, inter-textual relationships, and questions. The teacher focuses his or her comments on discussion skills, pointing out the established goals, then complimenting students on the phrases, body language, and actions the students used that helped deepen or move the discussion forward. The teacher can also model making comments to ask for clarification or to probe or extend students' comments. All studies on discussion in this paper recommended reducing the amount of teacher talk, having no predetermined meanings, and leaving the topic of the discussion open. Allowing for written or artistic response before group discussion may increase the depth and participation in the discussion because students can use their work as a reference. Although none of the studies mentioned this, written response after discussion may be a useful way for students to see how their perspective grew or

changed. In addition, the discussions improve when teachers set aside a designated time in each discussion for students to collaboratively reflect on how the discussion went and make goals for next time.

High quality literature-based instruction should include strategy instruction on the four cueing systems: pragmatics, semantics, syntax and graphophonics. Skills are embedded in the strategies, which are intended to give students many ways to problem-solve their construction of meaning. Strategy instruction also includes the process of using prior knowledge, predicting, confirming or correcting. In a final phase of the reading process, the reader sets the stage for future reading by integrating, rejecting, or withholding judgment on their experience of the text. One study said that strategy instruction includes an analysis of textual features such as the genre, characters, style, and plot of the literature. This strategy could be part of basic comprehension, but it could also lead to the critical pedagogy process of exploring and critiquing differences in stance between the reader and the author. Teachers facilitate strategy instruction through group problem-solving, guided reading, and metacognitive conferences with individuals. Graphic organizers depicting strategies may also be useful reference tools for students to use while reading.

The studies in Chapter 2 suggested some additional aspects of the environment of the classroom that contribute to quality literature-based instruction. The term “quality literature” can be hard to define, a few studies suggested that books that employ gaps in explanation, disequilibrating events, and cultural relevance could be aspects of quality. Many studies recommended

that exploring literature in themes could be a useful way to develop inter-textual integration & comparison. Using themes, the teacher can incorporate literature as one way to meaningfully consider topics in the content areas. The teacher should always distinguish literature, which is not intended to be factual, from non-fiction content area texts, which are supposedly factual. The theme can take over the entire classroom, creating a literacy rich environment for exploration. Many studies suggested that having a vast organized classroom library (including student publications) is a major asset and maximizes student choice. Studies recommended that the teacher structure several venues for reading every week: shared reading, small group guided reading, leveled reading, partner reading, and independent reading. In this way, students have highly scaffolded, collaborative, and independent opportunities to work with literature. When the teacher used conferences to discuss student self-evaluation and goals for their development as readers, students were more motivated to read. Finally, the Best Practices studies noted that teachers who have strong communication with parents and school staff could maximize resources for students, and learn about how to build on their cultural assets.

Although most studies in Chapter 2 found that literature-based instruction was beneficial for many demographics, some studies had recommendations for particular demographics. One study with subjects who had special needs suggested that explicit coaching on body language and phrases that students could use in group discussion, as well as calling attention to each student's insights, could be especially beneficial in positioning students as competent. For

students who speak a home language or dialect other than the language of instruction, providing opportunities for students to initially respond and explore ideas in their home language is vital; students can later translate their key ideas into the language of instruction. Graphic organizers were also very useful for all students. For students who entered school with little experience of books, opportunities to re-read (in teacher read-aloud, in pairs, and independently) contributed to their confidence with reading. Finally, teachers do well when they consider the culture of the students, the issues within the community, and the cultural assets of the students when they select books and themes, in this way students from all backgrounds have more opportunities to make real connections with literature than if all of the books are written from one cultural perspective.

Suggestions for Further Research

The review of the research revealed some thematic weaknesses in the body of research, and it lead to new questions. This section recommends ways that the existing research could have been better, and suggests directions for further research.

On one hand, most of the studies related to defining literature-based instruction were qualitative and were grounded in a socio-cultural-linguistic theory of literacy. On the other hand, the studies that were based in a more mechanistic view of literacy did not clearly define literature-based instruction, or defined it differently than the majority of studies in this paper. This makes it difficult to synthesize the results, and all inferences are only suggestions for further research. One recommendation might be that researchers with contrary biases

collaborate on a quantitative study and a qualitative study regarding the effects of aesthetic reader-response and critical pedagogy on achievement and motivation.

The studies should at least define the practice in detail.

Another aspect of the body of research that complicated synthesis of the findings related to the experience of the teachers. Some teachers were in their first year of a new form of instruction, while other studies mentioned that teachers had decades of experience. Many studies did not mention the experience levels of the teachers, which presents a major variable that should be addressed in the research. Ideally more studies should included teachers with many years of experience in literature-based instruction in order to evaluate it with more accuracy. A few studies looked at how teachers developmentally became more capable of literature-based instruction or critical pedagogy, more studies should be done on this process in order develop the support and education of teachers. Similarly, some studies suggested that it was initially difficult for students to switch to an active open-ended form of learning, while many studies only look at the initial experiences of students with that format. It would be more beneficial to see how students develop when they are exposed to many years of literature-based instruction. Further, none of the studies were longitudinal; studies related to future readership, academics, jobs, and societal roles would be very useful in determining the ultimate benefits that may coincide with aesthetic reader-response or critical pedagogy.

There is not very much research available on aesthetic reader-response and critical pedagogy explicitly. Nearly all of the studies on aesthetic reader-

response and critical pedagogy were action research studies, and their credibility was harmed because the researchers did not make moves to obtain a second rater or an auditor. Such research can only serve the purpose of suggesting areas for further research. A central problem with the research on literacy is that there seems to be little consideration of how aesthetic reader-response or critical pedagogy processes affect the quality of literature-based instruction. This body of research merely suggested that the structures that support aesthetic reader-response and critical pedagogy are also the structures that characterized high quality literature-based instruction. But this needs far more study. The aesthetic reader-response and critical pedagogy studies did not make connections to achievement or to the longevity of the other benefits they mentioned. Additional research should address these connections.

A new question arose as to whether teachers could effectively integrate aesthetic reader-response and critical pedagogy. Studies on this topic would develop the socio-cultural-linguistic theory of literacy and establish a framework for future research. A further way to develop the theory would be to survey how adult writers, artists, philosophers, and leaders use and interact with literature in their lives, and how that process developed for them. Studies that could address the processes and attitudes elementary students need in order to develop identity as life long readers would be very useful for a holistic reassessment of goals for literacy educators.

Conclusion

This review of literature explored the effects of aesthetic reader-response and critical pedagogy on the literacy of diverse elementary school students. It contextualized the controversy of this instruction in terms of the prevalence of deficit models of education, and the politics of privilege systems in the United States. It found that although there was questionable credibility of some studies, there was substantial dependability among studies that described the ways of literacy students used in literature-based instruction. It found that the structures that would support aesthetic reader-response and critical pedagogy are structures related to best practices in literacy education, and that they applied to students of diverse demographics. Debate still continues on whether achievement is as high or better than skills-based instruction, but this debate is impaired by lack of clear definitions of literature-based instruction. Researchers should conduct more descriptive and correlative research on how aesthetic reader-response and critical pedagogy influence the quality of literature-based instruction and the long-term engagement of students in literacy and society.

REFERENCES

- Arce, J. (2000). Developing voices: Transformative education in a first-grade two-way Spanish immersion classroom, a participatory study. *Bilingual Research Journal* 24(3), 249-260.
- Au, K. H., & Raphael, T. E. (2000). Equity and literacy in the next millennium. *Reading Research Quarterly* 35(1), 170-188.
- Baumann, J. F., Ivey, G. (1997). Delicate balance: Striving for curricular and instructional equilibrium in a second-grade, literature/ strategy-based classroom. *Reading Research Quarterly* 32(3), 244-275.
- Berry, R. A. W., Englert, C. S. (2005). Designing conversation: Discussions in a primary inclusion classroom. *Learning Disability Quarterly* 28(1), 35-58.
- Block, C. C. (1993). Strategy instruction in a literature-based reading program. *The Elementary School Journal* 94(2), 139-151.
- Bruner, J. (2002). *Making Stories: Law, Literature, Life*. Harvard University Press: Cambridge, MA.
- Chinn, C. A., Anderson, R. C., Waggoner, M. A. (2001). Patterns of discourse in two kinds of literature discussion. *Reading Research Quarterly* 36(4), 378-411.
- Cox, C., & Many, J. E. (1992). Stance towards a literary work: Applying the transactional theory to children's responses. *Reading Psychology: An International Quarterly* 13, 37-73.

- Cummins, J. (2007). Pedagogies for the poor? Realigning reading instruction for low-income students with scientifically based reading research. *Educational Researcher* 36(9), 564-572.
- Dahl, Karin, L. (1993). Children's spontaneous utterances during early reading and writing instruction in whole-language classrooms. *Journal of Literacy Research* 25(3), 279-294.
- Dahl, K. A., & Freppon, P. A. (1995). A comparison of innercity children's interpretations of reading and writing instruction in the early grades in skills-based and whole language classrooms. *Reading Research Quarterly* 30(1), 50-74.
- Davis, L. (2010). Toward a lifetime of literacy: The effect of student-centered and skills-based reading instruction on the experiences of children. *Literacy Teaching and Learning* 15(1 & 2), 53-79.
- Delpit, L. (2006). *Other People's Children: Cultural Conflict in the Classroom*. The New Press: New York.
- Duffy, G. G. (1993). Rethinking strategy instruction: Four teachers' development and their low-achievers' understandings. *The Elementary School Journal* 93(3), 231-247.
- Eeds, M., & Wells, D. (1989). Grand conversations: An exploration of meaning construction in literature study groups. *Research in the Teaching of English* 23(1), 4-29.
- Flint, A. S., Laman, T. T. (2012). Where poems hide: Finding reflective, critical spaces inside writing workshop. *Theory into Practice* 51, 12-19.

- Freire, P. (2007). *Pedagogy of the Oppressed*. New York: Continuum.
- Freire, P., Horton, M. (1990). *We Make the Road by Walking: Conversations on Education and Social Change*. Temple University Press: Philadelphia.
- Gallas, K. *The Language of Learning: How Children Talk, Write, Dance, Draw, and Sing their Understanding of the World*. Teachers College Press: New York.
- Gee, J. P. What is Literacy? Article retrieved from <http://www.ed.psu.edu/englishpds/articles/criticalliteracy/what%20is%20literacy.htm>
- Gersten, R. (1996). Literacy instruction for language-minority students: The transition years. *The Elementary School Journal* 96(3), 227-244.
- Gilles, C., Bixby, M., Crowley, P., Crenshaw, S. R., Henrichs, M., Reynolds, F. E., Pyle, D. eds. (1988). *Whole Language Strategies for Secondary Students*. Richard C. Owens Publishers Inc.: New York.
- Glod, M. (2008). Study questions "No Child" Act's reading plan. *The Washington Post, Education, May 2*.
- Goatley, V. J. (1997). Non-traditional learners' written and dialogic response to literature. *The Yearbook of the National Reading Conference* 41, 313-322.
- Goatley, V. J., Brock, C. H., Raphael, T. E. (1995). Diverse learners participating in regular education "book clubs." *Reading Research Quarterly* 30(3), 352-280.
- Goldstein, B. S. C. (1995). Critical Pedagogy in a Bilingual Spanish Education Classroom. *Journal of Learning Disabilities* 28(8), 463-475.

- Heath, S. B. (1983). *Ways with Words: Language, Life, and Work in Communities and Classrooms*. Cambridge University Press: New York.
- Hoffman, R. (2010). *The BFG and the Spaghetti Book Club: A case study of children as critics*. *Children's Literature in Education* 41, 234-250.
- Johnson, A. G. (2006). *Privilege, Power, and Difference*. McGraw Hill: New York.
- Johnson, D. D., Johnson, B. (2002). *High Stakes: Children, Testing, and Failure in American Schools*. Rowman & Littlefield Publishers, Inc.: New York.
- Juel, C., Miden-Cupp, C. (2000). Learning to read words: Linguistic units and instructional strategies. *Reading Research Quarterly* 35(4). 458-492.
- Kamps, M. D., Wills, H. P., Greenwood, C. R., Thorne, S., Lazo, J. F., Crockett, J. L., Akers, J. M., Swaggart, B. L. (2003). Curriculum influences on growth in early reading fluency for students with academic and behavioral risks: A descriptive study. *Journal of Emotional and Behavioral Disorders* 2, 211-224.
- Kantor, R., Miller, S. M., Fernie, D. E. (1992). Diverse paths to literacy in a pre-school classroom: A sociocultural perspective. *Reading Research Quarterly* 27(3), 184-201.
- Kersten, J. M. (2006). Why's everyone white? Moving toward critical pedagogy in an elementary classroom. *Journal of Urban Learning, Teaching, and Research* 2, 35-44.
- Klein, A. (2011). Stopgap Spending Bill Severs Array of Education Programs. *Education Week Blog*. Retrieved from http://blogs.edweek.org/edweek/campaign-k-12/2011/03/short-term_measure_cuts_a.html.

- Klesuis, J. P., Griffith, P. L., Zielonka, P. (1991). A whole language and traditional instruction comparison: Overall effectiveness and development of the alphabetic principle. *Reading Research Quarterly*, 30(2), 47-61.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage Publications, Inc.
- Lohfink, G., & Loya, J. (2012). The nature of Mexican American third graders' engagement with culturally relevant picture books. *Bilingual Research Journal: The Journal of the National Association for Bilingual Education* 33(3), 346-363.
- Maloch, B. (2002). Scaffolding student talk: One teacher's role in literature discussion groups. *Reading Research Quarterly* 37(1), 94-112.
- McCarthy, S. J. (2001). Identity construction in elementary readers and writers. *Reading Research Quarterly* 36(2), 122-151.
- Mosenthal, J., Lipson, M., Torncello, S., Russ, B., Mekkelsen, J. (2004). Contexts and practices of six schools successful in obtaining reading achievement. *The Elementary School Journal* 104(5), 343-367.
- Morrow, L. M. (1992). The impact of a literature-based program on literacy achievement, use of literature, and attitudes of children from minority backgrounds. *Reading Research Quarterly* 27(3), 250-275.
- Morrow, L. M., Tracey, D. H., Woo, D. G., Pressley, M. (1999). Characteristics of exemplary first-grade literacy instruction. *The Reading Teacher* 52, 462-476.

- Mosenthal, J., Lispon, M., Torncello, S., Russ, B., Mekkelsen, J. (2004). Contexts and practices of six schools successful in obtaining reading achievement. *The Elementary School Journal* 104(5), 343-367.
- Parsons, L. T. (2006). Visualizing worlds from words on a page. *Language Arts* 83(6), 492-500.
- Purcell-Gates, V., McInyre, E., & Freppon, P. A. (1995). Learning written story-book language in school: A comparison of low-SES children in skills-based and whole language classrooms.
- Ravtich, D. *The Death and Life of the Great American School System: How Testing and Choice are Undermining Education*. Basic Books: New York.
- Rose, M. (2006). *Possible Lives: The Promise of Public Education in America*. Penguin Books: New York.
- Rosenblatt, L. M. (1982). The literacy transaction: Evocation and response. *Theory into Practice* 21(4), 268-277.
- Rosenblatt, L. M. (1985). Language, literature, and values. *Language, Schooling, and Society*. Heinemann: Portsmouth, NH.
- Sacks, C. H., Mergendoller, J. R. (1997). The relationship between teachers' theoretical orientation toward reading and student outcomes in kindergarten children with different initial reading abilities. *American Educational Research Journal* 34, 721-739.
- Souto-Manning, M. (2009). Negotiating culturally responsive pedagogy through multicultural children's literature: Towards critical democratic literacy

- practices in a first grade classroom. *Journal of Early Childhood Literacy* 9, 50-74.
- Spring, J. (2011). *The American School: A Global Context from the Puritans to the Obama Administration*. McGraw Hill: New York.
- Stahl, S. A., Miller, P. D. (1989). Whole language experience approaches for beginning reading: A quantitative research synthesis. *Review of Educational Research*, (59)1, 87-116.
- Stahl, S. A. (1999). Why innovations come and go (and mostly go): The case of whole language. *Educational Researcher* 28(8), 13-22.
- Sulzby, E. Branz, C. M., Buhle, R. (1993). Repeated reading of literature and low socioeconomic status black kindergarteners and first graders. *Reading & Writing Quarterly* 9(2), 183-196.
- Taylor, B. M., Pearson, P. D., Clark, K., & Walpole, S. (2000). Effective schools and accomplished teachers: Lessons about primary-grade reading instruction in low-income schools. *The Elementary School Journal* 101(2), 121-165.
- Taylor, B. M.; Pearson, P. D.; Peterson, D. S.; & Rodriguez, M. C. (2003). Reading growth in high-poverty classrooms: The influence of teacher practices that encourage cognitive engagement in literacy learning. *The Elementary School Journal* 104(1), 3-28.
- Tindal, G., Marston, D. (1996). Technical adequacy of alternative reading measures as performance assessments. *Exceptionality*, 6(4), 201-230.
- Turner, J. (1995). The influence of classroom contexts on young children's motivation for literacy. *Reading Research Quarterly* 30(3), 410-441.

- Wharton-McDonald, R., Pressley, M., Hampston, J. M. (1998). Literacy instruction in nine first-grade classrooms: Teacher characteristics and student achievement. *The Elementary School Journal* 99(2), 101-128.
- Willis, A. I., & Harris, V. J. (2000). Political acts: Literacy learning and teaching. *Reading Research Quarterly* 35(1), 72-88.