Introduction

Science and literacy integration has received considerable attention in recent years due to U.S. educational reform documents such as the *Common Core State Standards* (National Governors Association Center for Best Practices and Council of Chief State School Officers (CCSSO) 2010) and *A Framework for K-12 Science Education* (National Research Council (NRC) 2012). Both reforms focus on preparing students for postsecondary and career opportunities in the 21st century and call upon teachers to support students in developing advanced abilities to read, write, and communicate across the disciplines.

Although science has traditionally been regarded solely as an empirical subject area (Fang, 2013), literacy is inherent to students’ participation in inquiry and the development of conceptual knowledge in science. Professional scientists read research and topic-based content to build background in an area of study, communicate their findings and ideas through documented field notes and journal articles, and participate in academic science-based conversations with others (Grant, Fisher, and Lapp, 2015). Since reading, writing, and listening are all essential to the work of professional scientists, it is critical that literacy practices be a central component of science instruction (Howes, Lim, & Campos, 2009).

According to The National Science Teachers Association (NSTA, 2002), teachers in the elementary grades must provide students with opportunities to develop the skills and knowledge required to function efficiently as problem-solvers in a scientific and technological society. This involves engaging elementary school students in first-hand investigation, fostering the development of inquiry skills, and supporting conceptual knowledge building in science. However, challenges related to science in the elementary grades are well documented and include a lack of materials and high-quality curriculum, lack of confidence and proficiency in science among teachers, and insufficient instructional time (Zembal-Saul, McNeil, &
Hershberger, 2013). Since the annual testing requirements of No Child Left Behind were enacted, elementary schools across the nation have reduced time spent on subjects other than reading and mathematics (Lapp, Grant, Moss, & Johnson, 2013). On average, less than 10% of instructional time is devoted to teaching science in the elementary grades (Mantzicopoulos, Samarapungavan, & Patrick, 2009). Furthermore, the little science instruction that does occur often places an overemphasis on fun, hand-on activities without any focus on developing deep conceptual knowledge in science (Zembal-Saul et al., 2013).

With a lack of quantity and quality science instruction taking place in elementary classrooms, it is not surprising that results from the 2015 National Assessment of Educational Progress (NAEP) reported only 38% of all fourth graders performed at or above the proficient level in science. Although reaching proficiency in science presents a challenge to all students, those who are of color (e.g., African-American and Hispanic), who come from low-socio-economic status (SES) backgrounds, who are English language learners (ELLs), and who are female may be at a substantial disadvantage when it comes to acquiring scientific literacy (Tong et al., 2014). For example, the 2015 NAEP results revealed that the average science score for White fourth-grade students was 33 points higher than their Black peers and 27 points higher than their Hispanic peers. Similar score gaps were reported between students eligible and not eligible for the National School Lunch Program (NSLP) and between ELLs and non-ELLs. In addition to such an achievement gap in science, ethnicity and SES are also significant factors associated with performance in reading. For example, in fourth grade, 18% of the African-American students and 21% of the Hispanic students scored at or above the proficient level on the NAEP reading assessment, when compared with 46% White students. On the scale scores, there is a 24-point difference between the White and Hispanic students and a 26-point difference between the White and African-American students. These statistics suggest there is a pressing
need to design more purposeful instruction to support all students in developing proficiency in both science and literacy, including ameliorating instructional inequality for students from ethnically diverse and low-income backgrounds.

As a former elementary classroom teacher and current doctoral student (i.e., first author) with a specific interest in the development of scientific literacy and as a reading professor who works with preservice and inservice teachers on literacy integration in science (i.e., second author), we set out to explore the existing empirical literature base on science and literacy integrated instruction in K-5 science classrooms. Because policy makers and funding agencies have been placing much attention to experimental and quasi-experimental research related to the topic, we decided to conduct a systematic literature review on studies that used such research designs. This systematic literature review was guided by two research questions:

(1) What evidence exists in the literature on the effectiveness of K-5 integrated science and literacy instruction in supporting both science and literacy learning from studies using an experimental or quasi-experimental design?

(2) What is the quality of the extant research on integrated science and literacy interventions for students in grades K-5 according to the What Works Clearinghouse (WWC) design standards?

Findings from this systematic literature review are used to explore the contributions of existing research using the aforementioned design to our knowledge about science and literacy integrated instruction. In addition, the authors discuss related implications for future research and provide research-based recommendations for classroom practice.
Science and literacy learning are not meant to compete with one another in the classroom. On the contrary, there are intersections and synergy between science and literacy teaching and learning (Pearson, Moje, & Greenleaf, 2010). Science and literacy instruction can (and should) both be implemented through inquiry-based instruction (Greech & Hale, 2006). Planning effective literacy and science integrated instruction requires recognizing the significant relationship between literacy development and science learning in the elementary grades. This symbiotic relationship can be closely analyzed through the lenses of the Next Generation Science Standards (NGSS Lead States, 2013a), which are based on the NRC’s (2012) K-12 Framework for Science Education, and the Common Core Anchor Standards (CCSSO, 2010). Although the symbiotic relationship between CCSS and NGSS is not the sole purpose for the integration of literacy and science in elementary classrooms, the following examination of related standards function as evidence for why there is so much policy-related attention on the topic. The following standards also provide implications for instructional practice and experiences elementary school age students need to have in learning and doing literacy and science.

First, we will consider the relationship between literacy and science in regard to speaking and listening. The six College and Career Readiness (CCR) Anchor Standards for Speaking and Listening accentuate the need for students to be able to engage in conversations and collaborations with others, present ideas precisely and in a variety of formats, and acquire and use a range of domain-specific vocabulary words and phrases. For example, CCR Speaking and Listening Anchor Standard 1 requires students to “prepare for and participate effectively in a range of conversations and collaborations with diverse partners, building on others’ ideas and expressing their own clearly and persuasively.” (CCSSO, 2010). The ability to use precise
language is also emphasized in the NGSS. One of the Science and Engineering Practices of the NGSS focuses on students’ ability to ask questions and define problems. As stated in A Framework for K-12 Science Education, “Students at any grade level should be able to ask questions of each other about the texts they read, the features of the phenomena they observe, and the conclusions they draw from their models or scientific investigations” (NRC, 2012, p. 56). The ability to pose specific questions, build on others’ ideas, or request additional information others are all central to building scientific knowledge and conducting investigations.

Next, we will examine this relationship in regard to reading. The CCSS places emphasis, as early as Kindergarten, on informational text reading, reading across multiple texts, critical reading of text, and supporting claims with textual evidence. For example, CCR Reading Anchor Standard 1 focuses on close reading and students’ ability to cite textual evidence to support inferences drawn from the text (CCSSO, 2010). CCR Reading Anchor Standard 8 requires that students “evaluate the argument and claims in a text, including the validity of the reasoning as well as the relevance and sufficiency of the evidence.” (CCSSO, 2010). Here we see a clear connection between the NGSS and the CCR Reading Anchor Standards. The Science and Engineering Practices of the NGSS highlight the role evidence plays in formulating questions, gathering information, and in reporting findings. While evidence may certainly be collected through formal laboratory experiments, it can also be gathered from close reading of science texts. By reading about science, students become better able to generate questions and ideas and critically evaluate information about the natural world (Grant, Fisher, & Lapp, 2015).

Lastly, let’s take a closer look at the multiple points of overlap between the NGSS and the CCSS in regards to writing. The ability to produce domain-specific text is a requirement of both scientists and engineers. The Science and Engineering Practices of the NGSS emphasize the importance of developing students’ ability to obtain, evaluate and communicate information.
Similarity, Writing Anchor Standard 2 of the CCSS requires students to produce informative/explanatory texts to examine and convey complex ideas and information (CCSSO, 2010). Additionally, Writing Anchor Standard 8 focuses on gathering relevant information from a wide variety of sources, assess the reliability of a source, and integrate the information without plagiarizing (CCSSO, 2010). Collecting data across a wide range of sources to support claims is a central activity of professional scientists.

For the purpose of this paper, we addressed sample connections between the NGSS and CCSS. However, Appendix M (2013b) of the NGSS outlines several other points of intersection between the CCSS Literacy Anchor Standards and the Science and Engineering Practices of the NGSS. The CCCS must be integrated with the NGSS to plan meaningful, rigorous science instruction. Dual implementation of the CCSS and NGSS ensures that students will have both the scientific and literacy knowledge bases to pose questions, design investigations, collect evidence, and support scientific arguments (Grant et al., 2015).

**Method**

Articles selected for this systematic literature review were located by: (a) a formal search of the following electronic data bases – the Education Resource Information Center (ERIC), Education Source, and Psych INFO; (b) exploration of federal websites, including Common Core State Standards Initiative, Department of Education, and National Center for Education Statistics; and (c) examination of web sites of national and international organizations related literacy education, including the Alliance for Excellent Education, American College Testing, and Carnegie Corporation of New York. Search terms included ‘literacy’, ‘science education’, ‘integration’, and ‘inquiry’. The abstracts of articles garnered from the electronic search were carefully reviewed to determine eligibility for inclusion.
Article Inclusion Criteria

Studies were selected for analysis and coding if they met the following criteria:

1. The study was published in a peer-reviewed journal between 2007 and 2017 and was conducted in the United States. Dissertation studies and unpublished manuscripts were excluded.

2. Participants were in grades K through 5.

3. The study investigated the efficacy of a literacy and science integrated intervention.

4. The study was an experimental or quasi-experimental design.

5. Studies reported either reading or science outcomes or both measures as a dependent variable.

Coding Procedures

A comprehensive coding protocol was adapted from a previously published synthesis (Kang, McKenna, Arden, & Ciullo, 2016) to organize information from all studies. The coding protocol focused on the following categories: participant characteristics (e.g., grade level), study methodology (e.g., experimental, quasi-experimental), a description of treatment (e.g., procedures for intervention), comparison conditions (e.g., business as usual), and dependent variables (e.g., reading comprehension measures, writing outcome measures). Studies that met inclusion criteria were coded and evaluated according to the WWC Procedures and Standards Handbook Version 4.0 (Institute of Education Sciences, 2017). Randomized group designs were evaluated according to the following: (1) degree of sample attrition and (2) the degree of similarity of the intervention and control groups at baseline (prior to the intervention). Studies in which group assignment was randomized were assigned either a rating of “meets group design standards without reservations,” “meets design standards with reservations,” or “does not meet standards.” For group designs without random assignment to treatment and control conditions,
the degree of baseline equivalence was assessed, and studies were assigned either a rating of “meets design standards with reservations” or “does not meet group design standards.”

To establish coding reliability, each article was double coded using the comprehensive coding protocol previously described as well as appropriate WWC criteria. Studies were independently coded and initial interrater agreement exceeded 90 percent.

**Results**

This systematic review of literature was developed to identify studies that included treatment of an integrated science and literacy intervention in grades K-5. First, an overview of studies that met inclusion criteria are reported. Next, the analysis of each study according to relevant WWC design standards is described. Lastly, a detailed discussion of each study is provided, including major findings and related limitations.

**Study Overview**

A total of 675 articles were identified as potentially relevant. The 675 papers were screened using the inclusion/exclusion criteria established in the protocol. Five studies met inclusion criteria. Two of the five studies used an experimental group design (40 percent) while the remaining three used a quasi-experimental group design (60 percent). Table 1 describes study characteristics including design, sample size, grade level, treatment duration and fidelity, and dependent measures.

Table 1

*Descriptive Characteristics of Selected Studies*

<table>
<thead>
<tr>
<th>Study</th>
<th>Research Design</th>
<th>Sample Size</th>
<th>Grade Level</th>
<th>Treatment</th>
<th>Duration</th>
<th>Treatment Fidelity</th>
<th>Dependent Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cervetti et al. (2012)</td>
<td>Experimental</td>
<td>Analyses were based on a randomly selected subset of students who participated in 4th</td>
<td>Researcher Developed Integrated Science-Literacy Unit</td>
<td>40 sessions, 45-60 minutes each</td>
<td>NR</td>
<td>Pre/Post Tests</td>
<td></td>
</tr>
</tbody>
</table>
In terms of intervention length, four studies reported information regarding the number of intervention sessions (Cervetti et al., 2012; Girod & Twyman, 2009; Mantzicopoulos et al., 2009; Wright & Gotwals, 2017), with a mean of 35 sessions. Four studies reported information on session duration (Cervetti et al., 2012; Girod & Twyman, 2009; Tong et al., 2014; Wright & Gotwals, 2017), which ranged from 45 minutes to 85 minutes with a mean of 59 minutes per session. Of the five studies, two (Mantzicopoulos et al., 2009; Wright & Gotwals, 2017) reported treatment fidelity data, all of which were reported at acceptable levels. In addition, two studies (Girod & Twyman, 2009; Tong et al., 2014) reported treatment fidelity procedures (e.g., classroom observations and field notes), but did not provide any specific data. One study (Cervetti et al., 2012) did not provide any information regarding treatment fidelity.

**Evaluation According to WWC Group Design Criteria**

Each study was evaluated according to WWC group design standards. Table 2 provides a description of this analysis. Two studies met group design standards without reservations (Cervetti et al., 2012; Tong et al., 2014) and two studies met group design standards with
reservations (Girod & Twyman, 2009; Wright & Gotwals, 2017). One study (Mantzicopoulou et al., 2009) did not meet design standards because they did not establish baseline equivalence according to WWC standards.

Table 2

Results of WWC Screening and Design Criteria

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Eligible design</td>
<td>Randomized</td>
<td>Quasi</td>
<td>Quasi</td>
<td>Randomized</td>
<td>Quasi</td>
</tr>
<tr>
<td>Random process?</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Sample Attrition:</td>
<td>No</td>
<td></td>
<td></td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Baseline equivalence:</td>
<td>Yes</td>
<td>No</td>
<td></td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Determination:</td>
<td>Meets WWC Standards</td>
<td>Meets WWC Standards With Reservations</td>
<td>Does Not Meet WWC Group Design Standards</td>
<td>Meets WWC Standards Without Reservations</td>
<td>Meets WWC Standards With Reservations</td>
</tr>
</tbody>
</table>

In-Depth Review of Selected Studies

In this section, all studies that met inclusion criteria are reported according to the study design. Characteristics of the intervention, participants, major findings, and related limitations of each study are discussed in the following section.

Experimental group designs. Two of the five included studies (40 percent) used an experimental group design (Cervetti et al., 2012; Tong et al., 2014). Cervetti et al. (2012) studied the efficacy of an integrated science and literacy curriculum approach to instruction as compared to isolated, content-comparable science instruction and literacy instruction at the upper-elementary level. Ninety-four fourth grade teachers (60 in Fall 2007, 34 in Spring 2008) were recruited with help from district-level science coordinators from 16 school districts in one Southern U.S. state. Teachers were randomly assigned to either a treatment or comparison group. The treatment unit was 40 sessions in total, comprised of four investigations, each with 10 sessions. In each 10-session investigation, four sessions were devoted to hands-on exploration, two sessions to reading, two sessions to writing, and two sessions to discourse, reflection and assessment. Science texts served specific roles in supporting students’ involvement in hands-on investigations and were also designed to develop students’ fluency, vocabulary, comprehension, and understanding of informational text features and structures. Teachers in the comparison group were asked to teach the content of their state science standards related to the topic of light, using a “business-as-usual” approach for the same amount of time each week and for the same duration. An assessment of student knowledge of science vocabulary, reading comprehension, and science understanding related to the topic of light was administered to all students in treatment and comparison classrooms both before and after the light unit was taught. Additionally, a pre- and post-assessment of student writing was administered to students in both groups. The writing assessment asked students to respond to the
following open-ended prompt: “How does light interact with materials? Give three examples.”

A rubric was used to score each written response on use of evidence, introduction, clarity, conclusion, vocabulary definition, and science content.

Researchers collected pretest and posttest data from a total of 467 students and found that students in the treatment group made significantly greater gains on measures of science understanding, science vocabulary, and science writing. Students in both groups made comparable gains in reading comprehension. This study met WWC standards without reservations.

Tong et. al. (2014) explored the effect of a one-year literacy-integrated science intervention delivered to fifth-grade Hispanic and African-American low socioeconomic students on their science and reading achievement. This study took place in a Southeast Texas school district, in which 85% of students qualified for free or reduced lunch. Out of a total of 10 intermediate schools in the district, four schools were randomly assigned to either the treatment group (n=2) or comparison group (n=2). The treatment intervention was delivered in English for 85 minutes each day and included: (a) Daily Oral and Written Language in Science (DOWLS) activity, in which students were asked to respond to a science-related prompt; (b) CRISSELLA, designed to build students’ science knowledge through vocabulary activities and science-related expository texts; and (c) Written and Academic oral language Vocabulary development in English in Science (WAVES), in which students created an individual glossary of academic science vocabulary. The treatment intervention also consisted of strategies for ELLs. The curriculum used in the comparison classroom was district-developed and aligned with state standards. The researchers observed that the use of science notebooks was limited in comparison classrooms. Additionally, the focus on reading vocabulary and questioning strategies varied and was inconsistent.
Measures consisted of both state-standardized tests and district-developed benchmark tests in science and literacy. Data was collected in the fall and spring of the school year from a total of 94 students in the treatment group and 194 students in the comparison group. Only data collected from low-SES non-ELLS was included in the analysis. Findings revealed that students who received the literacy-integrated science instruction were more likely than students in the control classrooms to pass the district-wide science and reading benchmark tests and the state-standardized tests of science, reading, and English fluency. In addition to a positive treatment effect on science achievement, results showed that the students in the treatment group also developed faster in oral reading fluency and were more likely to pass one of the reading benchmark tests than students in control classrooms. Tong and his colleagues also observed gender and ethnicity differences regarding science and reading achievement. Specifically, male students in both treatment and control conditions demonstrated higher academic achievement in science than did female students. Additionally, it was concluded that African-American students had a lower chance of successfully mastering science concepts and performing above the state standards when compared with Hispanic students across gender and condition. These findings are consistent with exiting research in the field (e.g., Kohlhass, Lin, & Chu, 2010). This study met WWC standards without reservations and provides evidence to suggest that literacy and inquiry-based science integrated instruction can promote both students’ science and reading achievement. The findings also call for increased attention to providing economically disadvantaged students of color with greater access to developing science proficiency.

**Quasi-experimental group designs.** Three studies (60 percent) used a quasi-experimental group design (Girod & Twyman, 2009; Mantzicopoulos et al., 2009; Wright & Gotwals, 2017). Girod and Twyman (2009) explored the added value of an integrated science and literacy curriculum over an inquiry-oriented science curriculum in two second grade
classrooms. Participants included 53 second graders from one small elementary school in a Northwest state. The experimental curriculum was a hands-on, integrated unit developed at the Lawrence Hall of Science, the public science education center at the University of California at Berkeley, with funding from the National Science Foundation under a grant titled, “Seeds of Science: Roots of Reading” (Seeds of Science/Roots of Reading, 2006). The Seeds curriculum was designed to support the development of literacy and science learning simultaneously with an emphasis on the discursive practices of science. The materials in this unit encouraged students to think, write, read, and talk in ways similar to professional scientists. The comparison curriculum, Great Explorations in Math and Science (GEMS) (2005), also developed at Lawrence Hall of Science, maintained the same curricular focus as the blended Seeds unit, but lacked many of the literacy components such as interpretation and critique of scientific text.

Students in both classrooms, prior to and following instruction, responded to scales for affect, interest, efficacy, and identity; measures of understanding of the nature of science (NOS); and conceptual understanding. No statistically significant differences were found between the experimental Seeds classroom and the comparison classrooms in regards to affect, interest, and efficacy. However, statistically significant differences were found in favor of the science and literacy blended unit on measures of identity, student understanding of the NOS and conceptual understanding. Detailed field notes were taken during observations of each period of science instruction in both classes and revealed noteworthy differences between the two curriculums in regards to reading, writing, and language use. For example, students had few opportunities in the GEMS classroom to generate claims and support it with evidence. On the other hand, argumentation and problem solving were very common in the Seeds classroom. This study meets WWC standards with reservations.
Mantzicopoulos et al. (2009) explored the effects of participation in the Scientific Literacy Project (SLP), an integrated science inquiry and literacy program (Mantzicopoulos, Patrick, & Samarapungavan, 2005), on kindergarten children’s emerging social meanings about science as compared to students who experienced only regular kindergarten curriculum, without SLP activities. Participants included 193 kindergarten children in four different schools located in a Midwestern suburban school district. Students in schools one and two (n=123) were selected to participate in science activities associated with the SLP. Students in schools three and four were selected as the comparison group (n=70). The treatment group was engaged in activities aimed at apprenticing young students into the culture of science using a guided inquiry framework. Within each unit, activities were grouped into three phases: pre-inquiry (e.g., generating questions, making predictions), inquiry (e.g., collecting data), and post-inquiry (e.g., whole-group discussion). Throughout all phases, students participated in both reading and writing activities (i.e., science notebook writing, shared book reading). Information from the comparison classrooms was acquired from classroom observations and teacher interviews. The researchers concluded that there were noticeable differences between instruction in the comparison classrooms and science activities associated with the SLP. For example, in comparison classrooms teachers often engaged students in hands-on activities without making connections to big science ideas.

Narrative data was collected from open-ended interview questions during the administration of the Puppet Interview Scales of Competence in and Enjoyment of Science (PICES). The PICES was administered to students in the treatment group at three separate time periods: prior to SLP activities (fall), mid-point (winter), and at completion of the intervention (spring). The comparison group only participated in spring assessment.
Analysis of narrative data showed that students in the SLP group viewed themselves as part of a community of science learners. Through their narratives, students in the SLP group referenced both the content of science and the epistemic processes that are integral to participating in science (e.g., building knowledge through reading science texts, asking questions, conducting experiments, drawing conclusions, sharing results). In contrast to the SLP group, the majority of the students from the comparison group reported that they did not learn science in kindergarten and that they did not know what science entailed. Also, over one-half of the children in the comparison group noted that science is not appropriate for kindergarteners. Although these results provide promising information about young children’s experiences with science and literacy integrated instruction, several limitations must be highlighted. First, although the researchers made attempts to match the SLP and comparison schools on a range of characteristics, there is no assurance that comparable samples of children and teachers were obtained. Second, due to the integrated nature of the inquiry and literacy activities, the authors were unable to determine which aspects of the intervention made a greater contribution to student’s understandings about science. This study did not meet the WWC design standards.

More recently, Wright and Gotwals (2017) conducted a quasi-experimental pilot study to explore the development of children’s (N=147) science talk in the context of an integrated science and disciplinary language and literacy curriculum for kindergarten. Teachers were recruited from one rural district in a large Midwestern state. Kindergarten teachers in two schools were invited to participate, and all 13 agreed to take part in the study. Teachers from one school (n = 6) were assigned to continue with business-as-usual science instruction while teachers in the second school (n = 7) implemented the SOLID Start (Science, Oral Language, and Literacy Development from the Start of School) curriculum. The intervention consisted of two four-week units aligned with both the NGSS and some strands of the CCSS-ELA standards for
kindergarten. Each unit comprised of 20 lessons that lasted roughly 45 minutes per day. Each lesson was guided by a driving question and included opportunities for active engagement with science phenonema as well as discussion about how these investigations answer the driving question. Teachers also read aloud from science tradebooks, provided explicit vocabulary instruction before or during the read-aloud, and faciliatated discussion relating ideas from the text to the driving question. Teacher measures consisted of teacher surveys, online instructional logs, teacher focus groups and individual interviews, and classroom observations. Measures of children’s learning included the Peabody Picture Vocabulary Test (PPVT-4), the Expressive Vocabulary Test (EVT-2), and a curriculum-based interview consisting of four sections: claims, evidence-based support, receptive vocabulary, and vocabulary application in a science context. While there were no differences at pretest, results showed Kindergarteners who received the SOLID Start curriculum outperformed children who received business-as-usual science instruction on all four sections of the curriculum-based interview. Although this study meets WWC with reservations, the findings provide initial evidence that with adequate teacher scaffolding and support, kindergarteners can engage in sophisticated science talk and meet the challenges of rigorous science and disciplinary literacy standards.

Discussion

The purpose of this systematic literature review was to explore the existing empirical literature base on science and literacy integrated instruction in K-5 science classrooms. The following questions guided this review: (1) What evidence exists in the literature on the effectiveness of K-5 integrated science and literacy instruction in supporting both science and literacy learning from studies using an experimental or quasi-experimental design? and (2) What is the quality of the extant research on integrated science and literacy interventions for students in grades K-5 according to the What Works Clearinghouse (WWC) design standards?
Overall, the reviewed studies demonstrated promising trends on both science and literacy outcomes, thus providing evidence that integrated approaches can be mutually supportive of both students’ science and literacy development. The curriculum utilized in these studies were grounded on the premise that science and literacy share similar cognitive processes and that reading and writing can enhance conceptual knowledge building in science. Therefore, students were engaged in activities designed to support science and literacy growth simultaneously, rather than using one domain to meet the learning goals of another. Additionally, inquiry was at the center of each model of science and literacy integration. In each study, students were engaged in asking science questions, making predictions, observing and recording data, and using evidence to answer their questions.

Unfortunately, purposeful inquiry-based science instruction is often postponed until later in school due to the belief that young children do not yet possess the skills required to participate in complex thinking (Mantzicopoulos et al., 2009). However, the studies included in this review provide evidence that with sufficient teacher scaffolding, students as young as Kindergarten can begin to read, write, and talk like scientists. This seems to suggest it is both possible and appropriate to involve elementary students in contextually rich inquiry experiences, as opposed to engagement with discrete sets of process tasks (e.g., categorizing). The young students who served as participants in these studies were engaged in a variety of effective and developmentally appropriate activities for prompting science and literacy. These included asking questions, engaging with science phenomena, participating in interactive read-alouds, writing in science journals, and partaking in meaningful science-based discussions. Understanding how to better support young students in science and disciplinary literacy learning is vital as we consider the learning progressions that are embedded within the educational standards (Wright & Gotwals, 2017).
Future Research

With the new literacy demands of the 21st century combined with the emphasis on disciplinary literacy called for by the CCSS and NGSS, there is surprisingly a lack of experimental studies to inform policy and practice in this area. Of the five studies that met selection criteria, only two studies met WWC criteria without reservations. It is apparent from this search that discussion regarding integrated approaches which are equally supportive of students’ science and literacy learning are emerging, but there is an urgent need for further investigation, especially at the elementary level where quality science instruction often takes a backseat to helping children prepare for standardized examination in mathematics and reading (Howes et al., 2008). Furthermore, future studies should employ stronger experimental designs by adhering to specific WWC quality indicators (Institute of Education Sciences, 2017).

The studies reviewed contained a variety of methodological limitations. For example, although the researchers made attempts to match the experimental and control groups on a range of characteristics, there is no guarantee that equivalent samples of children and teachers were obtained (e.g., Mantzicopoulos et al., 2009). Additionally, there is little assurance the teachers in the control groups carried out instruction that was content-comparable to the treatment units. Many of the studies reported very limited information about the “business-as-usual” group. Consequently, future studies should strive to ensure both equivalence and comparability by collecting and reporting baseline data for all groups.

Second, while the teachers in these studies were offered some pedagogical supports, they mostly received minimal professional development. Without adequate training, teachers may have not used curricular materials as intended. It is possible that increased curriculum-aligned professional development opportunities would have further improved instruction and student learning outcomes. Therefore, future efficacy studies should plan for several weeks of in-depth,
hands-on teacher training sessions to learn how to effectively use all curriculum materials prior to start of the intervention.

Lastly, the curriculum in most of these studies was only implemented for a short amount of time and with a small number of students, possibly limiting its full impact. Much larger scale replication studies are needed to cross-validate the results from these studies. Furthermore, future research should strive to examine student learning across a full year of implementation instead of only a short duration of time. Longitudinal studies could also shed insight on how early involvement in integrated science and literacy instruction contributes to students’ continued interest and performance in science.

Ensuring equity and excellence in literacy and science, regardless of race, ethnic group, gender, language, or disability should be a top priority (Freeman & Taylor, 2006). Very few studies have explored the impact of literacy and science integration on specific student subgroups (Tong et al., 2014). For this reason, further experimental research should focus on identifying teacher actions that are supportive of helping all students develop scientific literacy, including students with disabilities, ELLs, and students from low SES backgrounds.

**Implications for Practice**

Consistent with NSTA’s call for inquiry-based science instruction, the authors of the reviewed studies recognized inquiry as central to the learning of science. Through scientific inquiry, students learn to ask questions, conduct investigations, collect evidence across a variety of sources, and use that evidence to construct scientific explanations. To effectively implement integrated science and literacy instruction, teachers must recognize the synergetic relationship between inquiry-based science and literacy learning (Girod & Twyman, 2009). Inquiry-based science and literacy share common goals, functions, and strategies that can be capitalized upon as central features of integrated instruction (Cervetti et al., 2012). For example, the practices of
questioning, predicting, inferring, and drawing conclusions are vital to both scientific inquiry and comprehension when reading.

To help elementary teachers take advantage of this synergy, ongoing science-specific professional development is essential. Professional development should be focused on helping elementary teachers design inquiry-based science instruction, select and evaluate instructional materials, create a classroom environment conducive for learning science through inquiry, and recognize the role literacy plays in enhancing inquiry-based science. Additionally, effective integrative approaches to science are only possible with support (e.g., classroom assistance, materials, adequate planning time) (Mantzicopoulos et al., 2009). School administrators must provide ample amount of time for professional development opportunities, teacher collaboration, and instructional planning.

**Conclusion**

This systematic review of the literature, although it is not exhaustive, provides sample evidence about how elementary teachers can design and implement instruction supportive of both students’ science and literacy learning. However, several questions remain unanswered. For example: How are preservice elementary teachers being prepared to support students’ development of scientific literacy? What types of professional development is needed to support in-service teachers in the implementation of successful science and literacy integrated instruction? What instructional challenges do teachers encounter when employing science and literacy integrated instruction? And most importantly, how can teachers carry out quality literacy and science instruction in an atmosphere of high-stakes testing and accountability? These questions warrant further investigation to inform the field about what is needed for elementary science classrooms to become effective communities of inquiry that combine science teaching and learning with opportunities for students to develop critical thinking skills,
interact with text in authentic ways, engage in the specific literacy demands of the discipline, and meet both the goals of the NGSS and CCSS.

The waning prevalence of quality science instruction in U.S. elementary classrooms, intensified by the pressures of high-stake assessments, threatens to produce a future generation of science-illiterate individuals (Girod & Twyman, 2009). While not all students will grow up to become professional scientists, the ability to engage in discussions about science-based issues and solutions is a critical skill of all successful students. Therefore, existing research efforts must be continued and expanded to fully understand how to support teachers in the integration of these two synergetic disciplines, science and literacy, for the purpose of supporting the development of all students’ scientific literacy.
References


