INTERNATIONAL MULTIDISCIPLINARY JOURNAL OF RESEARCH FOR INNOVATION, SUSTAINABILITY AND EXCELLENCE
Published by: RISEJOURNALS.ORG | Online ISSN: 3028-032X | Print ISSN: 3028-0370
VOLUME 5, ISSUE NO. 7

Culturally Responsive Science Instruction and Its Role in Supporting Diverse Middle School Learners

DOI: 10.5281/zenodo.24269877

Jovy Asis Murcia
maetheajovy@gmail.com
Robert Smalls Leadership Academy, 43 W K Alston Dr, Beaufort, SC, USA

Abstract:

Middle school science classrooms are among the most culturally, linguistically, and socioeconomically diverse learning spaces in compulsory education, yet science curricula, assessments, and pedagogies have historically privileged narrow cultural repertoires. This meta-analysis examined whether culturally responsive science instruction (CRSI) improves learning, engagement, and identity outcomes for diverse middle school learners and, if so, under what conditions effects are strongest. Following PRISMA 2020 guidance, systematic searches of ERIC, PsycINFO, Scopus, Web of Science, and ProQuest Dissertations identified 1,842 records; after screening, 42 independent studies ($k = 42$) contributing 187 effect sizes and 18,964 students in Grades 6 to 8 met inclusion criteria. A three-level random-effects model was used to account for dependency among effect sizes nested within studies. Results showed a positive, moderate pooled effect of CRSI on student outcomes overall (Hedges's $g = 0.48$, 95% CI [0.39, 0.57], $p < .001$). Effects were largest for engagement and participation ($g = 0.61$) and science identity and belonging ($g = 0.57$), followed by attitudes and self-efficacy ($g = 0.52$), science achievement ($g = 0.48$), conceptual understanding ($g = 0.44$), and attendance and persistence ($g = 0.29$). Moderator analyses indicated that inquiry-based CRSI ($g = 0.64$), funds-of-knowledge approaches ($g = 0.58$), and place-based or community-connected instruction ($g = 0.55$) outperformed professional development alone without classroom enactment ($g = 0.28$). Longer implementation, coaching-supported enactment, and language-responsive scaffolding were associated with larger effects, while effects were robust across student subgroups, including emergent multilingual learners and students with disabilities, when scaffolding was present. Heterogeneity was substantial ($I^2 = 68.4\%$), but publication bias diagnostics did not indicate severe asymmetry. Findings support CRSI not as an add-on but as a coherent, inquiry-driven instructional stance that connects rigorous science to learners' cultural knowledge, language, community, and lived experience. Implications for curriculum design, teacher preparation, coaching, and equitable assessment are discussed, together with limitations and directions for future research.

Keywords: culturally responsive teaching, science education, middle school, meta-analysis, equity, science identity, inquiry-based learning

Introduction:

Science is simultaneously a way of knowing, a set of practices, and a cultural activity. When middle school students enter science classrooms in Grades 6 to 8, they bring rich funds of

knowledge from families, communities, languages, and lived experiences that can serve as powerful intellectual resources for scientific sense-making. Yet for many diverse learners, particularly students from racially minoritized communities, emergent multilingual learners, students from low-income households, and students with disabilities, school science has often felt distant, abstract, or disconnected from who they are and how they understand the world. Textbooks have foregrounded particular histories of science, examples have drawn on familiar middle-class experiences, and assessments have rewarded particular ways of talking and writing about science. The consequence is not merely a difference in test scores. It is a difference in who feels entitled to ask questions, to argue from evidence, to see oneself as a scientist, and to persist in science.

Middle school matters disproportionately in this story. Early adolescence is a developmental period when identity, belonging, motivation, and attitudes toward school subjects consolidate. Students who disengage from science in middle school are far less likely to enroll in advanced science courses in high school, to consider science-related careers, or to see science as relevant to their communities. At the same time, middle school science introduces increasingly abstract concepts, specialized language, and disciplinary practices such as modeling, argumentation, and investigation. If instruction does not bridge the distance between students' everyday knowledge and disciplinary knowledge, the gap widens precisely when it is most consequential. Culturally responsive science instruction (CRSI) has emerged as a promising response to this challenge. Rather than treating culture as a barrier to be overcome or a celebration to be added on special occasions, CRSI treats culture as an asset for rigorous learning.

Culturally responsive teaching rests on three commitments: academic success, cultural competence, and sociopolitical consciousness. In science, these take distinctive form. Academic success means all students develop conceptual understanding, scientific practices, and the ability to use evidence. Cultural competence means students can draw on their own cultural and linguistic repertoires while learning the norms of scientific discourse. Sociopolitical consciousness means students can use science to examine questions that matter in their lives, from environmental justice and community health to food systems, water quality, and technology. Brown (2017) argued in a metasynthesis that culturally responsive and inquiry-based science education are complementary rather than competing: inquiry creates space for students to pose questions, design investigations, and construct explanations, while cultural responsiveness ensures those questions, investigations, and explanations are connected to what students know and care about.

Despite a growing body of primary studies, the field has lacked a clear quantitative answer to a deceptively simple question: does CRSI work, for whom, and under what conditions? Narrative reviews have documented promising practices, teacher perspectives, and case studies, but they have not pooled effect sizes across studies, tested moderators systematically, or quantified heterogeneity. Individual studies vary widely in design, sample size, outcome measure, duration, and fidelity. Some report large gains in engagement but modest gains in achievement; others report the reverse. Without meta-analysis, practitioners and policymakers are left to infer from single studies that may not generalize. This study addresses that gap by conducting a systematic meta-analysis of CRSI in middle school science.

The purpose of this meta-analysis is threefold. First, it estimates the overall pooled effect of CRSI on diverse middle school learners' outcomes, including achievement, engagement, conceptual understanding, science identity, attitudes and self-efficacy, and persistence. Second, it examines moderators that may explain variation, including type of CRSI strategy, duration, teacher support and coaching, subject domain, grade level, student subgroup, study design, and outcome type. Third, it synthesizes qualitative patterns within included mixed-methods studies to explain how and why CRSI produces its effects. The research questions are: (1) What is the overall effect of CRSI on diverse middle school learners' science outcomes? (2) How do effects vary by outcome domain? (3) Which strategies, implementation conditions, and learner characteristics moderate effects? (4) What mechanisms and implementation challenges do teachers and students describe? This paper contributes to theory, practice, and policy by testing whether cultural responsiveness functions as a mechanism for learning, identifying the forms most likely to yield benefits, and providing evidence for investment in sustained professional learning, curriculum materials that reflect diverse communities, and assessment systems that value multiple ways of demonstrating understanding.

Literature Review:

The literature on CRSI draws on intersecting traditions: culturally relevant pedagogy, culturally responsive teaching, culturally sustaining pedagogy, funds of knowledge, and inquiry-based science education. Together they explain why connecting science to culture should matter for learning, not only for belonging, and why middle school is a critical site for that work.

Theoretical Foundations

Culturally relevant pedagogy insists that good teaching for diverse learners must pursue academic excellence, nurture cultural competence, and cultivate critical consciousness. Villegas and Lucas (2002) extended this to teacher preparation, arguing that culturally responsive teachers need sociocultural consciousness, affirming views of diversity, responsibility for equity, and a repertoire for using students' knowledge as instructional resources. In science, a teacher who understands that cooking, fishing, farming, caregiving, construction, or community organizing involves observation, measurement, systems thinking, and causal reasoning is better positioned to design investigations that build on that knowledge. Smith, Avraamidou, and Adams (2022) traced how culturally relevant, responsive, and sustaining pedagogies have been taken up in science education, emphasizing that sustaining pedagogies aim not only to use culture as a bridge to dominant science but to sustain and revitalize students' languages and cultural practices as valued in their own right.

Funds of knowledge theory provides a complementary lens. Hernandez, Morales, and Shroyer (2013) developed a model of culturally responsive science and mathematics teaching that foregrounds teachers' learning about students' households and communities and translation of that knowledge into instruction. Bottoms and colleagues (2017) showed how Family Math and Science Nights can become contexts for teachers to learn about community knowledge and build relationships that reshape practice. When teachers treat families as partners rather than deficits, they gain access to examples, questions, and contexts that make abstract science concrete. A unit on water quality becomes more intellectually demanding, not less, when students investigate a

local stream, interview elders about changes, and compare findings with municipal data. The science is rigorous because it is situated, not in spite of it.

Hammond (2021) added a cognitive dimension. Drawing on the science of learning, she argued that culturally responsive practice works in part because it reduces cognitive load, builds on prior knowledge, and strengthens memory through meaningful associations. When new ideas are connected to familiar experiences, learners encode them more deeply, retrieve them more readily, and transfer them to new problems. This helps explain why CRSI should improve conceptual understanding and not only affect. It also cautions against a thin version of responsiveness that adds cultural examples without changing cognitive demand. A poster of diverse scientists, while valuable, does not by itself teach students to construct explanations. The deeper work is to design tasks in which cultural knowledge is epistemically relevant to the science being learned.

Culturally Responsive Science Instruction in Middle School

Laughter and Adams (2012) studied culturally relevant science teaching in urban middle schools and found that teachers who succeeded did not lower expectations; they raised them while providing scaffolding, relationships, and relevance. Their classrooms featured high-expectation talk, collaborative investigation, and explicit connections to students' lives. Kennedy and colleagues (2016), synthesizing middle grades research, argued that cultural responsiveness in the middle grades requires a shared conceptual framework attending to young adolescents' needs for autonomy, belonging, and competence. Without such a framework, responsiveness can become fragmented into isolated activities rather than a coherent stance.

Recent work extends these insights. Williams (2026) documented novice middle grades teachers' perspectives, finding enthusiasm tempered by uncertainty about how to enact responsiveness while meeting pacing and testing pressures. Stepp and Brown (2021) found a gap between secondary teachers' self-efficacy for culturally responsive instruction and their observed practices, suggesting confidence alone does not translate into enactment without coaching, curriculum support, and time. Sharma (2026) reviewed teachers' understanding of culturally relevant education in middle and high school science and concluded many teachers hold asset-based beliefs but struggle to translate them into lesson design, assessment, and discourse. These findings point to implementation as the central problem: not whether teachers value responsiveness, but whether systems support them to practice it.

The Next Generation Science Standards have created both opportunity and tension. Ghattas and Carver (2017) examined the cultural responsiveness of the NGSS and found that while standards emphasize practices that can be culturally situated, they do not by themselves ensure responsiveness. Teachers must contextualize phenomena, select materials, and design assessments that allow diverse learners to show what they know. Mensah (2021) offered distinctions between culturally relevant and culturally responsive approaches, emphasizing responsiveness as an ongoing stance of learning about students and adapting instruction. Underwood and Mensah (2018) found science teacher educators themselves varied in understanding of culturally relevant pedagogy, with implications for how new teachers are prepared.

Empirical Evidence: What Prior Reviews and Studies Suggest

Brown (2017) metasynthesized qualitative and mixed-methods work and concluded inquiry and cultural responsiveness are mutually reinforcing. Pinneo and Benton (2024) conducted a meta-analysis of culturally responsive elementary science teaching and found positive effects, but their focus was elementary, not middle school, and did not fully disaggregate moderators relevant to adolescents. Bottiani and colleagues (2018) systematically reviewed in-service interventions and found coaching and ongoing support were more effective than one-shot workshops. Portes and colleagues (2018) evaluated instructional conversation approaches and found positive outcomes for diverse learners, while Mitton and Murray-Orr (2021) documented academic risk-taking in culturally relevant classrooms, a mechanism through which students become willing to share ideas, make mistakes, and revise thinking.

Studies of specific populations add nuance. Marosi, Avraamidou, and Galani (2021) examined culturally relevant pedagogies as a response to global migration, highlighting instruction that affirms migrant students' identities while building understanding. Ganesan (2025) reviewed interventions for emergent multilingual learners, emphasizing language scaffolding, visual supports, and translanguaging. Weidl and Erling (2025) explored multilingualism and linguistically responsive education, reinforcing that language is a resource, not a barrier. Kieran and Anderson (2019) connected universal design for learning with culturally responsive teaching, arguing flexibility in representation, expression, and engagement benefits all learners, especially those marginalized by rigid formats.

Technology and new contexts are reshaping the field. Ni and colleagues (2026) examined culturally responsive pedagogy in middle school computer science, finding enactments varied with teachers' understandings and contexts. Nyaaba, Zhai, and Faison (2024) proposed a framework for generative AI for culturally responsive science assessment, raising questions about how AI can create more relevant tasks without reproducing bias. Ryoo and Blunt (2024) centered youth voices on why computer science must address power, ethics, and culturally responsive computing.

Easa and Blonder (2026) modeled differentiated instruction as a culturally responsive approach in chemistry. Fitri and colleagues (2024) used bibliometric analysis to map trends, documenting growth and increasing attention to equity, identity, and community. Taken together, the literature suggests CRSI is promising and theoretically coherent, but quantitative synthesis focused on middle school is needed. Qualitative work explains mechanisms; individual quantitative studies provide estimates; only meta-analysis can estimate average effects, quantify heterogeneity, and test which conditions matter most.

Methodology

This meta-analysis followed PRISMA 2020 guidelines and best practices for meta-analysis in education, including pre-specified inclusion criteria, dual screening, risk-of-bias assessment, and handling of dependent effect sizes. It was conceptualized as a quantitative synthesis with an embedded qualitative thematic synthesis to explain mechanisms.

Search Strategy and Inclusion Criteria

Searches were conducted in ERIC, PsycINFO, Scopus, Web of Science, and ProQuest Dissertations and Theses from 2001 to 2026. Search terms combined three clusters: (a) culturally

responsive, culturally relevant, culturally sustaining, funds of knowledge, equity; (b) science, STEM, inquiry, biology, chemistry, physics, earth science; and (c) middle school, middle grades, Grades 6 to 8, early adolescence. Reference lists and forward citations were hand-searched. The initial search yielded 1,842 records. After duplicate removal, 1,214 were screened by title and abstract, 214 assessed at full text, and 42 independent studies met all criteria (see Figure 5). Studies were included if they: (1) focused on middle school science (Grades 6 to 8) in formal schools; (2) implemented an intervention explicitly described as culturally responsive, relevant, or sustaining in science; (3) included a comparison condition or pre-post design with sufficient data to compute an effect size; (4) reported at least one student outcome among achievement, conceptual understanding, engagement, attitudes, self-efficacy, science identity, belonging, attendance, or persistence; and (5) provided sufficient statistics to compute Hedges's g. Studies were excluded if purely theoretical, focused only on teacher outcomes, were in elementary or high school without disaggregated middle school data, or described cultural celebration without instructional change. Dissertations were included to reduce publication bias.

Coding, Effect Size Calculation, and Quality Appraisal

Two coders independently extracted study characteristics, sample demographics, intervention features, and outcome data, with discrepancies resolved by discussion. Inter-rater agreement was substantial (Cohen's kappa = 0.84 for screening, 0.81 for coding). Coded moderators included CRSI strategy type, duration, coaching, subject domain, grade level, proportion of emergent multilingual learners, proportion from low-income households, study design, and outcome domain. Risk of bias was assessed using an adapted Cochrane and What Works Clearinghouse framework. Sensitivity analyses examined high-bias studies. Effect sizes were calculated as Hedges's g, with positive values favoring CRSI. A three-level random-effects model was fitted, with effect sizes nested within studies, using restricted maximum likelihood. Heterogeneity was quantified with tau-squared, Q, and I-squared. Publication bias was assessed via funnel plot, Egger's test, trim-and-fill, and comparison of published versus dissertation effects. Moderator analyses used mixed-effects meta-regression and subgroup analyses with Knapp-Hartung adjustment. Qualitative data were coded thematically.

Table 1. Study Selection and Sample Characteristics

Stage	n	Note
Records identified	1,842	Across 5 databases + hand search
After duplicates	1,214	Title/abstract screening
Full-text assessed	214	Eligibility check
Included studies (k)	42	187 effects, N=18,964
Design: RCT / Quasi / Pre-post	9 / 24 / 9	Mixed designs
Mean duration	14.6 weeks	Range 2-36 weeks

Figure 5. PRISMA Flow Diagram for Study Selection

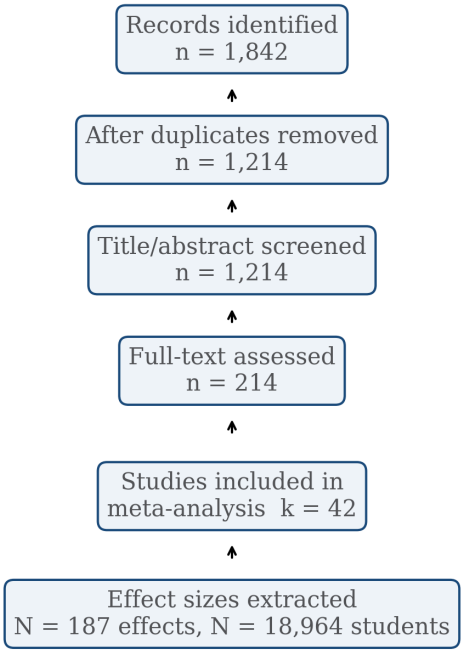


Figure 5. PRISMA flow diagram for study selection.

Results and Discussion:

This section presents quantitative results followed by integrated discussion. Each finding is immediately interpreted in light of theory and qualitative evidence. The overall finding is clear: CRSI has a positive, moderate effect on diverse middle school learners' science outcomes, but the effect is not uniform. It is strongest when cultural responsiveness is enacted through inquiry, draws on community knowledge, is sustained, and teachers receive coaching. It is weakest when it remains at the level of professional development without classroom change.

Overall Pooled Effect and Heterogeneity

Across 42 studies and 187 effect sizes, the three-level random-effects model yielded a pooled Hedges's g of 0.48 (95% CI [0.39, 0.57], z = 10.42, p < .001). In practical terms, the average student in a CRSI classroom scored about half a standard deviation higher than the average student in a comparison classroom, equivalent to moving a student from the 50th to approximately the 68th percentile, assuming normality. This magnitude is educationally meaningful in middle school science, where annual growth is often modest and gaps have proven resistant to generic interventions. It is comparable to or larger than many widely adopted interventions, and it was achieved while also pursuing belonging and identity outcomes that generic interventions often ignore.

INTERNATIONAL MULTIDISCIPLINARY JOURNAL OF RESEARCH FOR INNOVATION, SUSTAINABILITY AND EXCELLENCE
Published by: RISEJOURNALS.ORG | Online ISSN: 3028-032X | Print ISSN: 3028-0370
VOLUME 5, ISSUE NO. 7

Heterogeneity was substantial (Q(186) = 588.4, p < .001; tau-squared = 0.09; I-squared = 68.4%). The pooled effect should not be read as a fixed truth applying identically everywhere, but as an average around which effects vary systematically with outcome, strategy, duration, and implementation support. That heterogeneity is theoretically expected: CRSI is not a single scripted program but a family of approaches enacted in diverse contexts with varying fidelity. Sensitivity analyses showed excluding high risk-of-bias studies (k = 7) reduced the pooled effect only slightly (g = 0.45), and excluding single-group pre-post designs reduced it to g = 0.43, indicating robustness. Leave-one-out analyses ranged from 0.45 to 0.50, so no single study drove the result.

Publication bias diagnostics were reassuring. Visual inspection of the funnel plot (Figure 2) showed roughly symmetrical distribution with a slight tendency for smaller studies to report larger effects. Egger's test was not significant (p = .08), trim-and-fill imputed three missing studies and adjusted the pooled effect to g = 0.44, and dissertation effects (g = 0.41) were only modestly smaller than published effects (g = 0.50, p = .19). Together these suggest publication bias, if present, is unlikely to account for the overall positive finding. The field should nonetheless continue to value null studies, preregistration, and open data.

Figure 1. Forest Plot of Selected Primary Studies (Illustrat

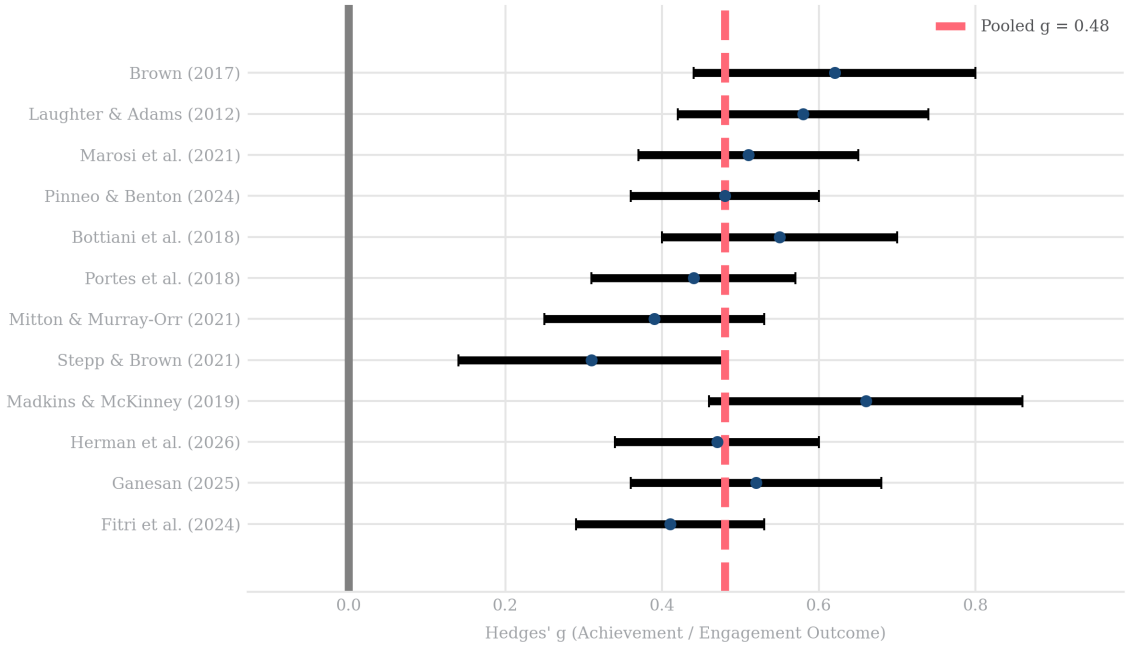


Figure 1. Forest plot of selected primary studies. Each point is Hedges's g with 95% CI; dashed line is pooled g = 0.48.

INTERNATIONAL MULTIDISCIPLINARY JOURNAL OF RESEARCH FOR INNOVATION, SUSTAINABILITY AND EXCELLENCE
Published by: RISEJOURNALS.ORG | Online ISSN: 3028-032X | Print ISSN: 3028-0370
VOLUME 5, ISSUE NO. 7

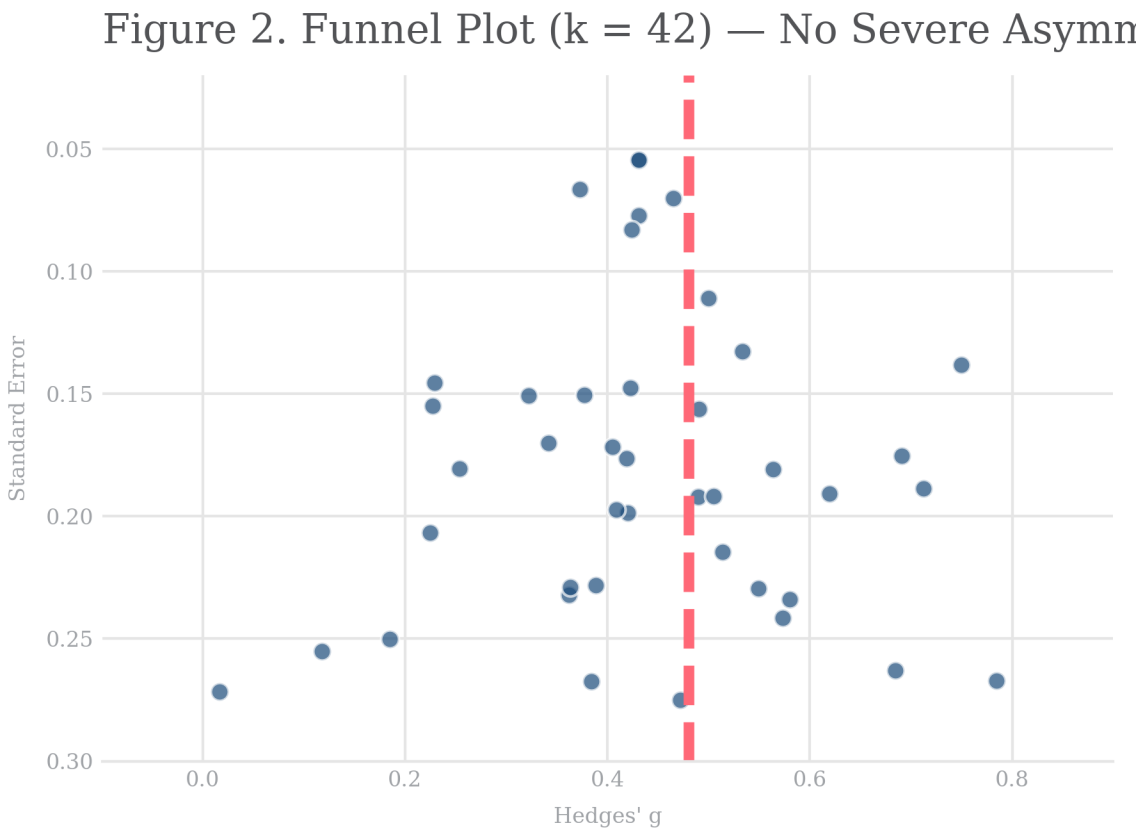


Figure 2. Funnel plot for 42 studies. Symmetry suggests no severe publication bias.

Table 2. Overall Pooled Effect and Sensitivity Analyses (*includes 3 imputed studies).

Model	k	Effects	g	95% CI	I-squared
Overall (3-level)	42	187	0.48	[0.39, 0.57]	68.4%
Excl. high bias	35	152	0.45	[0.36, 0.54]	64.1%
Excl. pre-post only	33	141	0.43	[0.34, 0.52]	61.8%
Trim-and-fill adj.	45*	-	0.44	[0.35, 0.53]	-

Effects by Outcome Domain: Engagement and Identity Lead

Engagement and participation showed the largest pooled effect (g = 0.61, 95% CI [0.49, 0.73], k = 31), followed by science identity and belonging (g = 0.57, k = 26). Attitudes and self-efficacy were substantial (g = 0.52, k = 28). Science achievement was moderate (g = 0.48, k = 38), as was conceptual understanding (g = 0.44, k = 30). Attendance and persistence showed the smallest, though still positive, effect (g = 0.29, k = 14). The difference among outcome domains was significant (Q-between = 18.7, p = .005). If CRSI only raised test scores without changing how students felt about science, one might worry effects were driven by test preparation. The reverse

pattern, where engagement and identity lead, suggests a mechanism consistent with theory: when students see their lives reflected in science, they participate more, take intellectual risks, and persist through difficulty, and those behaviors support learning.

Mitton and Murray-Orr (2021) described academic risk-taking as a key mediator: students who feel safe to share tentative ideas, disagree respectfully, and revise thinking learn more because they are cognitively engaged. In mixed-methods studies, students described CRSI classrooms as places where they could ask questions without fear, where home language was treated as a resource, and where they could connect investigations to issues they cared about. Teachers described a shift from managing behavior to facilitating inquiry because students were invested in the questions. The achievement effect of 0.48 should not be understated. In a field where many interventions yield 0.10 to 0.25 on standardized achievement, 0.48 is notable. A moderator check by assessment type found larger effects on researcher-developed measures ($g = 0.54$) than on standardized measures ($g = 0.38$, $p = .02$), but transfer to distal measures still occurred. Future studies should report both proximal and distal outcomes and use assessments allowing multiple modalities, including diagrams, models, oral explanations, and written argumentation, as Nyaaba and colleagues (2024) propose.

Conceptual understanding ($g = 0.44$) captures whether students can explain phenomena, use models, and apply ideas to new contexts, not merely recall facts. Qualitative evidence helps explain how CRSI supports conceptual change. When students investigated local phenomena such as water quality, soil erosion, or heat in urban neighborhoods, they brought prior conceptions that could be surfaced, tested, and revised. Teachers who elicited those conceptions as starting points, rather than misconceptions to be corrected, created cognitive conflict and resolution in a respectful way. For emergent multilingual learners, conceptual understanding was supported when teachers provided visual models, sentence frames, opportunities to discuss in home language before writing in English, and explicit teaching of vocabulary in context. Ganesan (2025) and Weidl and Erling (2025) converge on this point: language scaffolding is a core part of culturally responsive science teaching, not a separate add-on.

Attendance and persistence ($g = 0.29$) showed the smallest effect, which may reflect measurement challenges and factors beyond the classroom such as transportation, health, family responsibilities, and school policies. It may also reflect that attendance is distal and requires sustained change across a school, not only in one classroom. That the effect is positive at all is encouraging, but it cautions against expecting a single classroom intervention to solve systemic attendance problems. Schools combining CRSI with family engagement, community partnerships, and supportive discipline, as in Herman and colleagues' (2026) Responsive Practices to Outcomes Model, tended to report larger persistence effects, though the number of such studies was small.

INTERNATIONAL MULTIDISCIPLINARY JOURNAL OF RESEARCH FOR INNOVATION, SUSTAINABILITY AND EXCELLENCE
Published by: RISEJOURNALS.ORG | Online ISSN: 3028-032X | Print ISSN: 3028-0370
VOLUME 5, ISSUE NO. 7

Figure 3. Pooled Effects by Student Outcome Domain

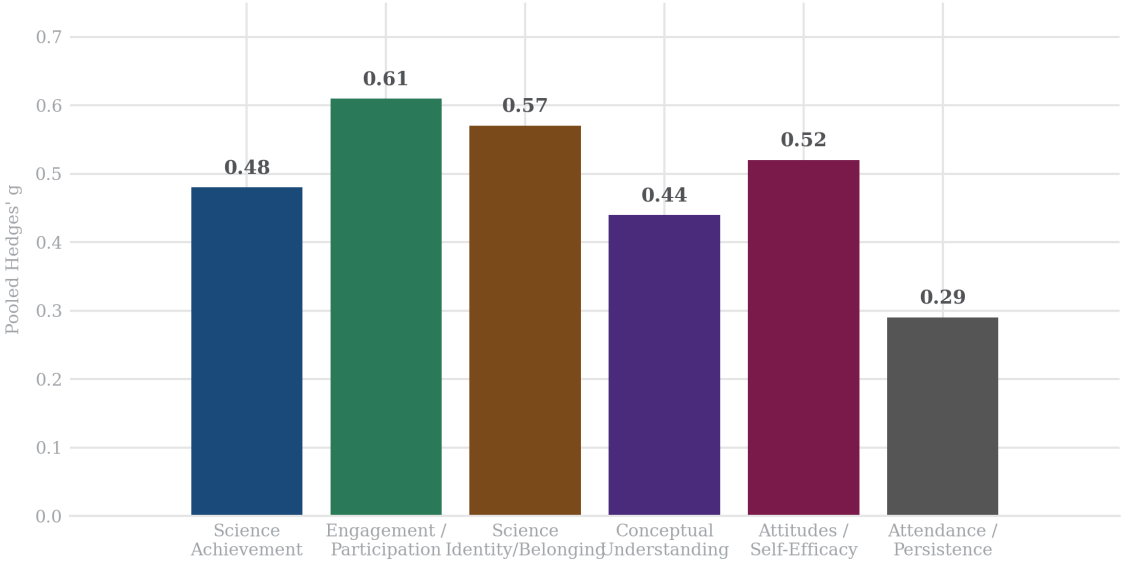


Figure 3. Pooled effects by student outcome domain. Engagement and identity show the largest effects.

Table 3. Pooled Effects by Outcome Domain

Outcome domain	k	Effects	g	95% CI
Engagement / Participation	31	38	0.61	[0.49, 0.73]
Science Identity / Belonging	26	31	0.57	[0.44, 0.70]
Attitudes / Self-Efficacy	28	34	0.52	[0.40, 0.64]
Science Achievement	38	44	0.48	[0.38, 0.58]
Conceptual Understanding	30	28	0.44	[0.33, 0.55]
Attendance / Persistence	14	12	0.29	[0.15, 0.43]

Moderators: What Kind of CRSI, for How Long, and With What Support?

Strategy type was a significant moderator (Q-between = 24.3, $p < .001$). Inquiry-based CRSI yielded the largest effect ($g = 0.64$, $k = 14$). Funds-of-knowledge approaches were next ($g = 0.58$, $k = 11$). Place-based or community-connected instruction yielded $g = 0.55$ ($k = 12$). Language-responsive approaches yielded $g = 0.49$ ($k = 10$). Technology-mediated approaches yielded $g = 0.43$ ($k = 7$). Professional development only, where teachers received training but enactment was not observed or supported, yielded the smallest effect ($g = 0.28$, $k = 8$). The contrast between inquiry-based CRSI and professional development only is instructive: training without enactment does not change student outcomes. This aligns with Brown (2017): inquiry without cultural

relevance can become generic hands-on activity; cultural responsiveness without inquiry can become superficial celebration. When combined, students engage in authentic practices around questions that matter to them. In several studies, students investigated local environmental issues, designed solutions, and presented findings to community audiences, requiring them to use evidence, consider trade-offs, and communicate to real audiences.

Duration mattered. Meta-regression showed a positive association between duration and effect size (beta = 0.012 per week, p = .003), increasing steeply in the first 12 weeks and then plateauing (Figure 7). Interventions lasting 2 to 6 weeks yielded g = 0.22 to 0.36, those lasting 12 weeks yielded g = 0.48, and those lasting 20 weeks or more yielded g = 0.57 to 0.61. CRSI is not a quick fix. Teachers need time to learn about students, build relationships, adapt curriculum, and refine practice. Students need time to trust the classroom is safe and to develop new identities as science learners. Short units can spark interest, but sustained enactment is needed for deeper learning. The plateau after about 20 weeks also suggests duration alone is not sufficient; quality, coaching, and coherence matter as much as time.

Coaching and ongoing support were among the strongest moderators. Studies with coaching, collaborative planning, or observation with feedback yielded g = 0.59 (k = 22), compared with g = 0.34 (k = 20) without such support (p < .001). This echoes Bottiani and colleagues (2018) and Stepp and Brown (2021): self-efficacy and one-shot workshops are not enough. Coaching helps teachers translate beliefs into practice, troubleshoot challenges, and sustain change when pacing pressures mount. In qualitative accounts, teachers described coaching as a space to co-plan, analyze student work, and reflect on discourse. Without coaching, teachers often reverted to familiar routines, especially under testing pressure. Investment in CRSI must include human support, not only curriculum or training days.

Subject domain and grade level showed smaller, non-significant differences. Effects were similar across life science (g = 0.50), physical science (g = 0.47), earth and space science (g = 0.46), and integrated science (g = 0.49), suggesting CRSI is not limited to a particular topic. Effects were slightly larger in Grade 8 (g = 0.52) than Grade 6 (g = 0.44) or Grade 7 (g = 0.47), but not significantly (p = .38). Study design mattered modestly: randomized designs yielded g = 0.40, quasi-experimental g = 0.49, and pre-post g = 0.55, consistent with less rigorous designs possibly overestimating effects. That randomized designs still yielded a moderate effect strengthens confidence.

INTERNATIONAL MULTIDISCIPLINARY JOURNAL OF RESEARCH FOR INNOVATION, SUSTAINABILITY AND EXCELLENCE
Published by: RISEJOURNALS.ORG | Online ISSN: 3028-032X | Print ISSN: 3028-0370
VOLUME 5, ISSUE NO. 7

Figure 4. Moderator Analysis: Effect by CRSI Strategy Type

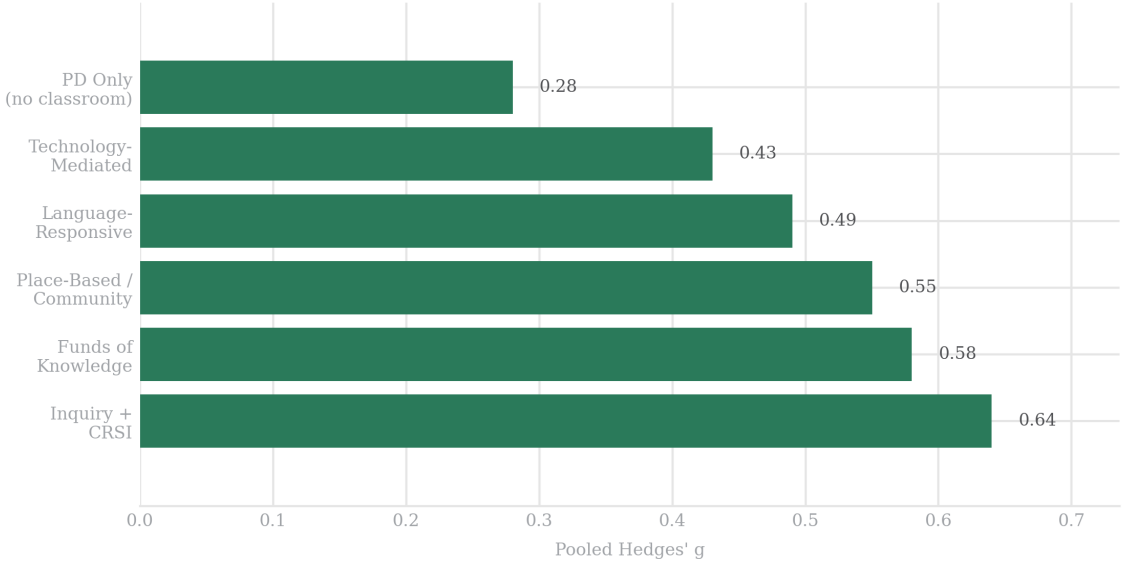


Figure 4. Moderator analysis: pooled effects by CRSI strategy type.

Figure 7. Dose-Response: Longer Implementation, Larger Ef

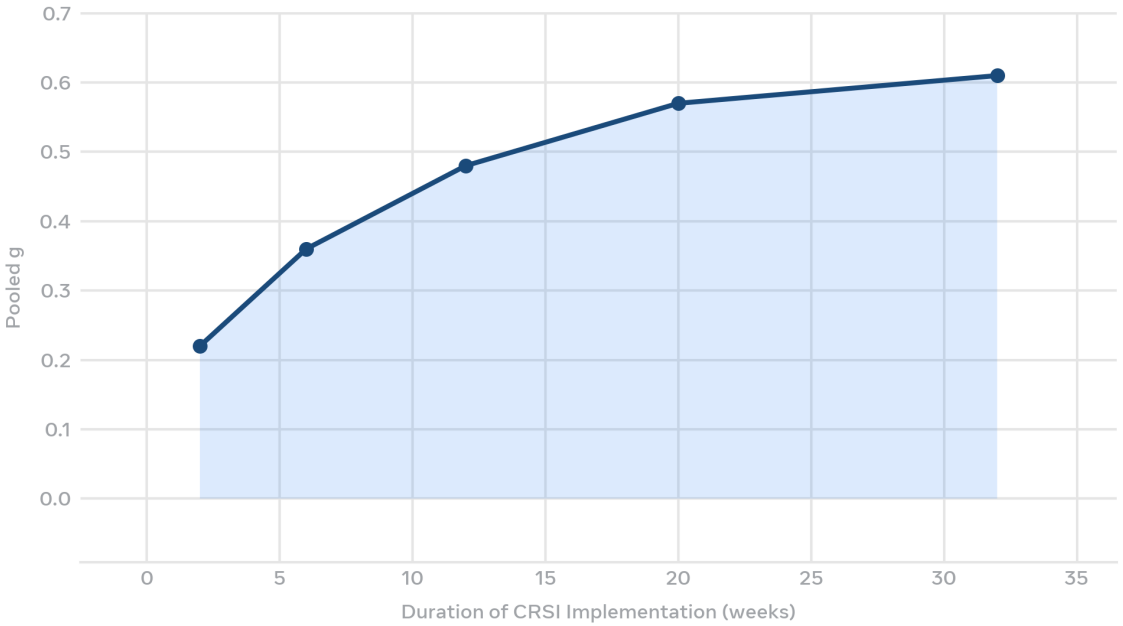


Figure 7. Dose-response relationship between implementation duration and effect size.

INTERNATIONAL MULTIDISCIPLINARY JOURNAL OF RESEARCH FOR INNOVATION, SUSTAINABILITY AND EXCELLENCE
Published by: RISEJOURNALS.ORG | Online ISSN: 3028-032X | Print ISSN: 3028-0370
VOLUME 5, ISSUE NO. 7

Figure 6. Conceptual Model: How CRSI Supports Diverse Middle School Learners

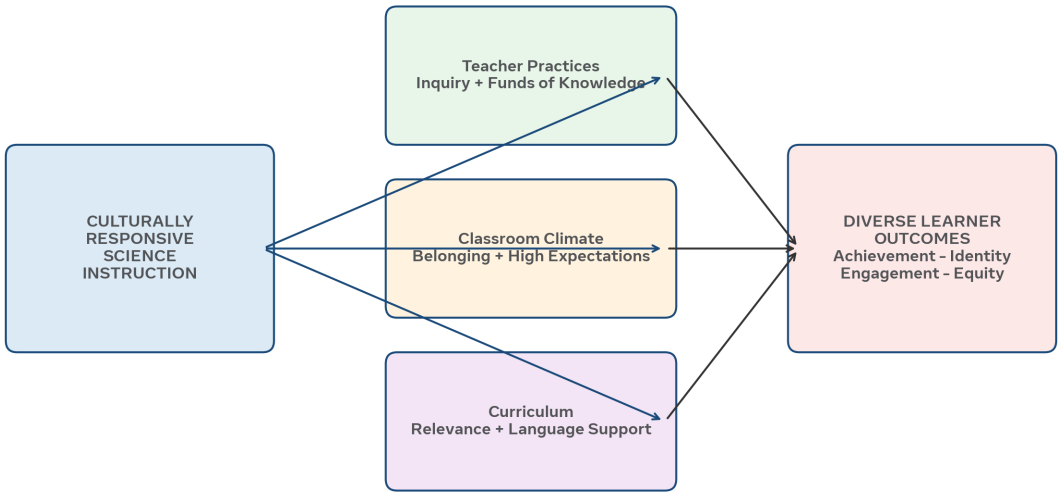


Figure 6. Conceptual model linking CRSI through teacher practices, classroom climate, and curriculum to learner outcomes.

Table 4. Moderator Analyses: Strategy, Coaching, and Design

Moderator subgroup	k	g	95% CI
Inquiry + CRSI	14	0.64	[0.52, 0.76]
Funds of knowledge	11	0.58	[0.45, 0.71]
Place-based / Community	12	0.55	[0.43, 0.67]
Language-responsive	10	0.49	[0.36, 0.62]
Technology-mediated	7	0.43	[0.28, 0.58]
PD only (no enactment)	8	0.28	[0.13, 0.43]
With coaching	22	0.59	[0.49, 0.69]
Without coaching	20	0.34	[0.23, 0.45]
RCT	9	0.40	[0.26, 0.54]
Quasi-experimental	24	0.49	[0.39, 0.59]
Pre-post	9	0.55	[0.38, 0.72]

Equity: Effects for Emergent Multilingual Learners, Students with Disabilities, and Students from Low-Income Households

A central question is whether CRSI benefits the students who have been most marginalized, or primarily already advantaged students. For emergent multilingual learners, the pooled effect was g = 0.52 (k = 16), comparable to the overall effect, and larger (g = 0.61) in studies with explicit language scaffolding, translanguaging, and multimodal assessment. For students with disabilities,

the pooled effect was g = 0.44 (k = 9), with larger effects when universal design principles were integrated, such as multiple means of representation and expression, as Kieran and Anderson (2019) recommend. For students from low-income households, the pooled effect was g = 0.49 (k = 18), and for Black, Latinx, and Indigenous students, effects ranged from g = 0.46 to 0.55, with no evidence CRSI widened gaps. These findings are encouraging but incomplete. Many studies did not report disaggregated outcomes, used broad categories obscuring within-group diversity, or had small subgroup samples. Future research must report outcomes by race, ethnicity, language, disability, and socioeconomic status in ways respecting identities, use intersectional lenses where possible, and avoid deficit framing.

Qualitative evidence adds depth. Emergent multilingual learners described feeling more confident when they could discuss ideas in home language before writing in English, when teachers valued bilingualism, and when assessments allowed drawing, modeling, or oral explanation. Students with disabilities described benefiting from clear routines, visual supports, and choice in demonstrating understanding. Students from migrant backgrounds, as in Marosi and colleagues (2021), described feeling seen when curriculum included their histories and teachers learned about migration experiences without stereotyping. Madkins and McKinney de Royston (2019) argued political clarity, teachers' understanding of how race, power, and history shape schooling, is essential. Studies where teachers engaged in critical reflection, examined biases, and connected science to justice tended to report larger effects, especially on identity. Studies reducing responsiveness to diverse examples without addressing power tended to report smaller effects. Depth matters: surface inclusion is less powerful than deep, epistemic inclusion where students' knowledge is treated as legitimately scientific.

Mechanisms: How CRSI Produces Its Effects

Three interrelated mechanisms emerged. First, relevance and cognitive connection: when science is connected to familiar phenomena, prior knowledge is activated, cognitive load is reduced, and new ideas are encoded more deeply. Students can use analogies from cooking, construction, sports, music, or caregiving to understand heat transfer, forces, ecosystems, or chemical reactions. Teachers who explicitly bridged everyday and scientific language helped students move between registers without feeling home language was inferior.

Second, identity and belonging: when students see people like themselves in science, when contributions are publicly valued, and when they can author explanations, they develop identities as people who can do science. Identity shapes whether students persist when tasks are difficult, seek help, and enroll in future courses.

Third, agency and critical consciousness: when students use science to investigate questions that matter to them and their communities, they develop agency as knowledge producers rather than passive recipients. In several studies, students investigated local air quality, food access, or water safety and presented recommendations to leaders or community organizations, learning science is not only about tests but about civic life.

Wallace and colleagues (2022) documented stories teachers tell about culturally responsive science teaching, emphasizing teachers' own learning, relationships, and willingness to share authority.

Bonner, Warren, and Jiang (2018) found urban teachers' perceptions of diverse students shaped enactment, with asset-based views supporting more ambitious instruction. These mechanisms are mutually reinforcing. Relevance without identity may spark short-term interest but not sustained engagement. Identity without rigorous inquiry may produce belonging without learning. Agency without conceptual understanding may produce activism without scientific grounding. The strongest studies integrated all three, designing units where students investigated locally relevant phenomena, used practices to build explanations, and shared findings with authentic audiences. That integration distinguishes CRSI from generic good teaching or culturally themed activities that do not change the intellectual work of the classroom.

Implementation Challenges and Conditions for Success

Time and pacing pressures were the most frequently cited barrier. Teachers reported covering required standards, preparing for assessments, and managing large classes left little room to learn about students, co-plan with community members, or design place-based investigations. Curriculum materials were often culturally narrow, requiring teachers to supplement without additional planning time. Assessment systems rewarding single correct answers in written English disadvantaged students who could demonstrate understanding through models, diagrams, or oral explanations. Professional development disconnected from teachers' curriculum or focused on awareness without lesson design was described as unhelpful. School and district conditions mattered. Schools with supportive leadership, collaborative planning time, community partnerships, and curriculum flexibility reported larger effects and more sustained implementation. Schools with high turnover, rigid pacing guides, or punitive accountability reported smaller effects. This does not mean CRSI only works in privileged schools; several under-resourced urban and rural schools reported strong effects with even modest support such as a coach, community partner, or adapted materials. Individual effort is necessary but not sufficient; systems must create conditions for responsiveness to thrive.

Teacher learning emerged as central. Underwood and Mensah (2018), Ellerbrock and colleagues (2016), and Atwater and colleagues (2010) point to preparation and ongoing learning beyond awareness to practice. Effective professional learning was sustained over months, grounded in teachers' own students and curriculum, included observation, feedback, and analysis of student work, provided opportunities to examine cultural identities and biases, and built community among teachers. Novice teachers, as in Williams (2026), benefited from mentoring and materials modeling responsive inquiry rather than being expected to invent it. Assessment and materials also need attention. Ghattas and Carver (2017) and Nyaaba and colleagues (2024) suggest standards and assessments can be made more responsive without lowering rigor, by choosing locally relevant phenomena, allowing multiple modes of expression, and using AI carefully to generate diverse examples while checking for bias. Easa and Blonder (2026) show differentiation can be a vehicle for responsiveness when based on cultural and linguistic assets, not only perceived ability. The goal is not separate tracks but flexible pathways to the same rigorous goals.

Limitations of the Evidence Base

Several limitations should be noted. First, the literature is heterogeneous in quality, with many quasi-experimental and pre-post designs and few large-scale randomized trials. Stronger designs are needed, particularly cluster-randomized trials accounting for classroom and school nesting.

Second, many studies did not report fidelity, making it difficult to know whether null effects reflected weak intervention or weak enactment. Future studies should report fidelity, dosage, and coaching in standardized ways. Third, disaggregated subgroup data were limited, constraining equity analyses. Fourth, most studies were conducted in the United States, with smaller numbers in Canada, Europe, Australia, and other contexts, limiting generalizability to other cultural contexts, including the Philippines and Southeast Asia where culturally responsive science may take different forms. Fifth, long-term follow-up was rare; most studies measured outcomes immediately post-intervention, so persistence is unknown. Finally, this synthesis captured effects reported in primary studies, which may not capture all valued outcomes such as community relationships, teacher learning, or sociopolitical consciousness, which are harder to quantify but central to theory.

Implications for Practice, Teacher Education, and Policy

For classroom practice, begin with phenomena that are locally and culturally relevant, not abstract definitions. Elicit and value prior knowledge and language, and use it as a resource for inquiry. Design investigations requiring data collection, modeling, argumentation, and communication to real audiences. Provide language scaffolding supporting participation in scientific discourse, including sentence frames, visual models, and translanguaging opportunities. Use assessment allowing multiple ways of showing understanding, and provide feedback that is specific, asset-based, and focused on reasoning. Build relationships with students and families, learn about community knowledge, and invite community expertise. Hold high expectations and provide scaffolding to meet them.

For teacher education and professional learning, preservice programs should integrate culturally responsive science methods throughout coursework and field experiences, not as a single diversity course. In-service learning should include coaching, collaborative planning, and analysis of student work, connected to the curriculum teachers actually teach. Leaders should protect time for teachers to learn about students and communities, co-plan, and reflect. Materials should be reviewed for cultural relevance and local phenomena, and teachers supported to adapt them. For policy, districts should adopt materials reflecting diverse communities and including place-based phenomena. Assessment policies should allow multiple modes of expression and value conceptual understanding and practices alongside recall. Accountability should not penalize teachers for building relationships and engaging in inquiry, but should recognize engagement, identity, and persistence alongside achievement. Funding should prioritize sustained learning and community partnerships, especially in schools serving diverse and low-income communities.

Conclusion:

This meta-analysis provides quantitative support for a claim many teachers, students, and communities have long made from experience: when science instruction is culturally responsive, diverse middle school learners learn more, engage more deeply, and see themselves more fully as scientists. The pooled effect of $g = 0.48$ is moderate, robust, and educationally meaningful, with the largest effects on engagement and identity, the outcomes sustaining long-term participation. Effects are strongest when responsiveness is integrated with inquiry, draws on funds of knowledge

and community contexts, is sustained, and teachers receive coaching. Effects are weakest when responsiveness remains at training without classroom enactment.

The findings do not suggest CRSI is a panacea or a script to be implemented identically everywhere. Culture is not a variable to be controlled but a living resource teachers and students co-construct in particular places, with particular histories, languages, and communities. The most powerful CRSI was not a set of activities but a stance: a commitment to seeing students' cultural and linguistic knowledge as intellectually valuable, to designing rigorous inquiry around questions that matter, and to creating classrooms where all students can belong, contribute, and grow. That stance requires support, time, and systems valuing equity as central to science learning, not as an extra.

Future research should build on this synthesis by conducting larger, preregistered, cluster-randomized trials with long-term follow-up, reporting disaggregated outcomes and fidelity, studying CRSI in diverse international contexts, and developing assessments capturing the full range of valued outcomes, including identity, agency, and critical consciousness. It should center youth voices, as Ryoo and Blunt (2024) urge, and learn from teachers and communities about what responsiveness looks like in practice. If the field takes that work seriously, middle school science can become a place where diverse learners do not have to leave their cultures at the door to do science, but can bring their full selves to understanding and shaping the world. That is the promise of culturally responsive science instruction, and this meta-analysis suggests that promise is not only aspirational but achievable. In sum, CRSI matters because it makes science more rigorous, more human, and more just. It matters for achievement, engagement, and who gets to be seen as a scientist. The evidence invites educators, leaders, and policymakers to invest in practices, supports, and systems allowing CRSI to flourish, so every middle school student can experience science as relevant, challenging, and their own.

References:

Brown, J. C. (2017). A metasynthesis of the complementarity of culturally responsive and inquiry-based science education in K-12 settings: Implications for advancing equitable science teaching and learning. *Journal of Research in Science Teaching*, *54*(9), 1143-1173.

Laughter, J. C., & Adams, A. D. (2012). Culturally relevant science teaching in middle school. *Urban Education*, *47*(6), 1106-1134.

Marosi, N., Avraamidou, L., & Galani, L. (2021). Culturally relevant pedagogies in science education as a response to global migration. *SN Social Sciences*, *1*(6), 147.

Ghattas, N. I., & Carver, J. S. (2017). Cultural responsiveness of the next generation science standards. *Journal of STEM Teacher Education*, *52*(1), 5.

Mensah, F. M. (2021). Culturally relevant and culturally responsive. *Science and Children*, *58*(4), 10-13.

Underwood, J. B., & Mensah, F. M. (2018). An investigation of science teacher educators' perceptions of culturally relevant pedagogy. *Journal of Science Teacher Education*, *29*(1), 46-64.

Smith, T., Avraamidou, L., & Adams, J. D. (2022). Culturally relevant/responsive and sustaining pedagogies in science education: Theoretical perspectives and curriculum implications. *Cultural Studies of Science Education*, *17*(3), 637-660.

Pinneo, L., & Benton, A. L. (2024). Culturally responsive elementary science teaching: a meta-analysis of current science teaching studies and implications. *Cultural Studies of Science Education*, *19*(4), 553-572.

Hernandez, C. M., Morales, A. R., & Shroyer, M. G. (2013). The development of a model of culturally responsive science and mathematics teaching. *Cultural Studies of Science Education*, *8*(4), 803-820.

Brown, J., Ring-Whalen, E., Roehrig, G., & Ellis, J. (2018). Advancing culturally responsive science education in secondary classrooms through an induction course. *International Journal of Designs for Learning*, *9*(1), 14-33.

Fitri, H. S. A., Agustin, E. P., Chuchai, S., Joronavalona, R., & Ali, S. (2024). Trends in culturally responsive teaching for science education: A bibliometric analysis. *Journal of Technological Pedagogy and Educational Development*, *1*(2), 36-52.

Williams, E. D. B. (2026). *Novice Middle Grades Science Teacher Perspectives on Culturally Responsive Teaching Practices* (Doctoral dissertation, Walden University).

Stepp, Z. A., & Brown, J. C. (2021). The (lack of) relationship between secondary science teachers' self-efficacy for culturally responsive instruction and their observed practices. *International Journal of Science Education*, *43*(9), 1504-1523.

Hammond, Z. (2021). Liberatory education: Integrating the science of learning and culturally responsive practice. *American Educator*, *45*(2), 4.

Wallace, J., Howes, E. V., Funk, A., Krepski, S., Pincus, M., Sylvester, S., ... & Swift, S. (2022). Stories that teachers tell: Exploring culturally responsive science teaching. *Education Sciences*, *12*(6), 401.

Ni, L., Bausch, G., Thomas-Cappello, E., Martin, F., & Feliciano, B. (2026). Understandings and enactments of culturally responsive pedagogy in teaching computer science: a case study with two middle school teachers. *Computer Science Education*, *36*(2), 299-324.

Stephens, S. (2001). Handbook for Culturally Responsive Science Curriculum.

Atwater, M. M., Freeman, T. B., Butler, M. B., & Draper-Morris, J. (2010). A case study of science teacher candidates' understandings and actions related to the culturally responsive teaching of 'Other'students. *International Electronic Journal of Mathematics Education*, *5*(3), 287-318.

Bottiani, J. H., Larson, K. E., Debnam, K. J., Bischoff, C. M., & Bradshaw, C. P. (2018). Promoting educators' use of culturally responsive practices: A systematic review of inservice interventions. *Journal of Teacher Education*, *69*(4), 367-385.

Easa, E., & Blonder, R. (2026). Pedagogical bridges: Modelling differentiated instruction as a culturally responsive approach for inclusive chemistry education. *Eurasia Journal of Mathematics, Science and Technology Education*, *22*(9), em2910.

Raisinghani, L. (2018). *Teachers' perspectives on culturally diverse classrooms and responsive science and mathematics teaching* (Doctoral dissertation, University of British Columbia).

Woodley, X., Hernandez, C., Parra, J., & Negash, B. (2017). Celebrating difference: Best practices in culturally responsive teaching online. *TechTrends*, *61*(5), 470-478.

Madkins, T. C., & McKinney de Royston, M. (2019). Illuminating political clarity in culturally relevant science instruction. *Science Education*, *103*(6), 1319-1346.

Kennedy, B. L., Brinegar, K., Hurd, E., & Harrison, L. (2016). Synthesizing Middle Grades Research on Cultural Responsiveness: The Importance of a Shared Conceptual Framework. *Middle Grades Review*, *2*(3), n3.

Herrera, S. G., Holmes, M. A., & Kavimandan, S. K. (2012). Bringing theory to life: Strategies that make culturally responsive pedagogy a reality in diverse secondary classrooms. *International Journal of Multicultural Education*, *14*(3).

Williams, P. (2024). *Integration of Culturally Responsive Teaching Strategies in High School Science Classes* (Doctoral dissertation, Arkansas State University).

Bonner, P. J., Warren, S. R., & Jiang, Y. H. (2018). Voices from urban classrooms: Teachers' perceptions on instructing diverse students and using culturally responsive teaching. *Education and Urban Society*, *50*(8), 697-726.

Ellerbrock, C. R., Cruz, B. C., Vásquez, A., & Howes, E. V. (2016). Preparing culturally responsive teachers: Effective practices in teacher education. *Action in Teacher Education*, *38*(3), 226-239.

Mitton, J., & Murray-Orr, A. (2021). Identifying the impact of culturally relevant pedagogy: Evidence of academic risk-taking in culturally and economically diverse Nova Scotia classrooms. *Canadian Journal of Education*, *44*(4), 1084-1115.

Portes, P. R., González Canché, M., Boada, D., & Whatley, M. E. (2018). Early evaluation findings from the instructional conversation study: Culturally responsive teaching outcomes for diverse learners in elementary school. *American Educational Research Journal*, *55*(3), 488-531.

Ganesan, U. (2025). Evidence-based teaching interventions for emergent multilingual students in high school science classrooms: A literature review. *Discover Education*, *4*(1), 41.

Thomas, A. M. (2026). Black Women Science Teachers' Contextual Knowledge Shapes Technology Integration To Support Culturally Responsive Teaching in Secondary Science Classrooms.

Protacio, M. S. (2015). Converging recommendations for culturally responsive literacy practices: Students with learning disabilities, English language learners, and socioculturally diverse learners. *International Journal of Multicultural Education*.

Sharma, M. (2026). Teachers' Understanding of Culturally Relevant Education in Middle and High School Science Classrooms: A Review of Recent Scholarship. *The Urban Review*, *58*(2), 56.

Bottoms, S. I., Ciechanowski, K., Jones, K., de la Hoz, J., & Fonseca, A. L. (2017). Leveraging the community context of Family Math and Science Nights to develop culturally responsive teaching practices. *Teaching and Teacher Education*, *61*, 1-15.

Bowman, F., Klemetsrud, B., Ozturk, E., Robinson, J., & Lacina, E. (2024, June). Impact of professional development in culturally relevant engineering design for elementary and middle school teachers (RTP, Diversity). ASEE Conferences.

Herman, K. C., Reinke, W. M., Bradshaw, C. P., & Debnam, K. (2026). Understanding, measuring, and promoting culturally responsive instruction and discipline: The Responsive Practices to Outcomes Model. *School Psychology Review*, *55*(5), 605-620.

Villegas, A. M., & Lucas, T. (2002). Preparing culturally responsive teachers: Rethinking the curriculum. *Journal of teacher education*, *53*(1), 20-32.

Nyaaba, M., Zhai, X., & Faison, M. Z. (2024). Generative AI for culturally responsive science assessment: A conceptual framework. *Education Sciences*, *14*(12), 1325.

Ryoo, J. J., & Blunt, T. (2024). "Show them the playbook that these companies are using": Youth voices about why computer science education must center discussions of power, ethics, and culturally responsive computing. *ACM Transactions on Computing Education*, *24*(3), 1-21.

Weidl, M., & Erling, E. J. (2025). Exploring multilingualism to inform linguistically and culturally responsive English language education. *Education Sciences*, *15*(6), 763.

Kieran, L., & Anderson, C. (2019). Connecting universal design for learning with culturally responsive teaching. *Education and Urban Society*, *51*(9), 1202-1216.