

Article

Not peer-reviewed version

Generative AI for Culturally Responsive Assessment in Science: A Conceptual Framework

[Matthew Nyaaba](#)^{*}, [Xiaoming Zhai](#)^{*}, Morgan Z. Faison

Posted Date: 17 September 2024

doi: 10.20944/preprints202409.1276.v1

Keywords: Culturally Responsive Assessment; Assessment; Prompt Engineering; Generative AI (GenAI); Artificial Intelligence (AI); Science Education



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Generative AI for Culturally Responsive Assessment in Science: A Conceptual Framework

Matthew Nyaaba ^{1,2,*}, Xiaoming Zhai ^{1,3,*} and Morgan Z. Faison ²

- ¹ AI4STEM Education Center, University of Georgia, USA
- ² Department of Educational Theory and Practice, University of Georgia, USA
- ³ Department of Mathematics Science and Social Studies Education University
- * Correspondence: matthew.nyaaba@uga.edu (M.N.); xiaoming.zhai@uga.edu (X.Z.)

Abstract: This study presents the core tenets that constitute culturally responsive science assessments (CRSciA) for K-12 education and aligned them with the capabilities of generative AI (GenAI). The tenets were identified based on their dominance in the literature, including *Indigenous Language, Race and Ethnicity, Religious Beliefs, Family and Community, and Indigenous Knowledge*. These tenets were instrumental in developing the GenAI-CRSciA framework and the subsequent prompts to generate CRSciAs aligned with the Next Generation Science Standards (NGSS) while comparing it with standard prompt strategies. The results showed that the GenAI-CRSciA approach outperformed standard prompting strategies in terms of generating CRSciAs, proving its effectiveness in addressing the ongoing challenges in science assessments. Even though the GenAI-CRSciA framework presented assessments that cover a range of the cultural backgrounds, it presents potential challenges such as compromising curricula core competencies and overgeneralization of students' backgrounds; overlooking subcultural variations within a country or region, for example differences between urban and rural students, or between different ethnic or linguistic groups within the same nation, which can lead to bias and less effective CRSciA outputs. To address this, we recommend that teachers provide comprehensive and specific background information about their students when using the GenAI-CRSciA prompting, ensuring that the generated content is truly reflective of the students' unique cultural contexts. In larger class settings, teachers could implement comprehensive local data indexing, where detailed student background information is uploaded and retrievable for use in crafting more personalized CRSciAs. We recommend educators make conscious efforts to validate the generated assessments with input from human experts, including community or family, educators, and even students who are knowledgeable about the cultural nuances. We also suggest continuous professional development would be necessary to enhance teachers' understanding of the cultural tenets.

Keywords: culturally responsive assessment; assessment; prompt engineering; generative AI (GenAI); artificial intelligence (AI); science education

1. Introduction

"Sometimes it's not the student who is failing the assessment—it might be that the assessment is failing to fully assess the abilities of the student"- [1]

In today's diverse classrooms, one of the biggest challenges educators face is creating science assessments that genuinely reflect the cultural backgrounds of every student [2]. Science education plays a vital role in equipping students to tackle critical issues like climate change, public health, and technological advancements. Therefore, the purpose of science assessment at the K-12 level, as guided by Framework for K-12 Science Assessment [3] and the Next Generation Science Standards (NGSS; NGSS Lead States, 2013), is to develop students' scientific literacy by ensuring a deep understanding of essential scientific knowledge and the ability to apply this knowledge to real-world problems [4,5]. This approach encourages students to engage in science and engineering practices in a way that is meaningful to their lives, leveraging their cultural experiences as assets in the learning process [5,6]. Specifically, researchers suggest culturally responsive science assessments (CRSciAs) to connect science education to students' cultural contexts, making learning more relevant and engaging.

However, implementing CRSciAs in K-12 classrooms remains challenging. Teachers face difficulties due to increasing classroom diversity and the dominance of traditional, western-centric assessment methods, which often disadvantage migrant, historically marginalized, and indigenous students. Furthermore, CRSciA practices, though essential, are time-consuming and complex, making them hard to scale effectively [7,8].

Recent development of Generative Artificial Intelligence (GenAI) offers a promising solution to CRSciAs by automating assessment generation and scoring, thereby easing the burden on educators and enabling more equitable, culturally tailored assessments that better reflect students' diverse backgrounds [9]-[11]. GenAI can provide multimodal learning opportunities [12], handling a variety of cultural contexts such as adapting to different languages, symbols, and local knowledge more often more effectively than a single classroom science teacher could manage alone. Specifically in science education, GenAI can assist teachers in creating NGSS-aligned assessment tasks and providing customized feedback based on students' needs [13].

Nonetheless, the use of GenAI in assessment is not without its challenges. There are legitimate concerns about the fairness and consistency of GenAI-generated assessments, especially when it comes to avoiding bias and cultural stereotypes and ensuring that all students are assessed equitably [14,15]. Without careful consideration, there is a risk that these GenAI technologies could hegemonically reinforce the very inequities they are meant to overcome.

To address these challenges, this study seeks to address these concerns by developing a conceptual framework that brings together the capabilities of GenAI and the core tenets of CRSciA for K-12 education, and further develop a GenAI-CRSciA prompt template strategy using the framework. The GenAI-CRSciA prompt template is designed to guide educators in using GenAI to generate assessment items both scientifically and culturally responsive.

2. Literature Review

There are several studies related to culturally responsive assessments in general, and consequently to CRSciA. This section provides an overview of the impact of standardized assessment, which has traditionally dominated science education. It also examines CRSciA, emphasizing the appropriateness and urgent need and the challenges that persist in implementing CRSciA, and the capabilities of GenAI in addressing these challenges.

Impact of Standardized Assessment

One of the key impacts of standardized assessment (traditional standardized assessment) is the high possibility of creating an "achievement gap". Achievement gap discourse is prominent particularly in the U.S [16]. The term refers to the disparities in standardized test scores between Native Indigenous, Black, Latina/o, recent immigrants, and White students [17]. Recent studies indicate that current science assessment practices are grounded in the language, experiences, and values of the dominant White culture [18]. While this is a concern, Ladson-Billings [16] further posited that even a focus on the gap is misplaced, instead, we need to look at the "education debt" that has accumulated historically over time. She draws an analogy with the concept of national debt, which she contrasts with that of a national budget deficit to argue the significance of historical inequities in education. Moreover, historically marginalized students including Native Indigenous, Black, Latina/o have had accumulated disadvantages, limited opportunities, and a lack of access to equal education for generations.

CRSciA, therefore aims to address the educational debt in science and not just the gap. This implies that the lack of representation of cultures in science assessments has profound and far-reaching effects across every aspect of life in wider society. A biased science assessment that affects a particular group of students can influence their career goals and limit their contributions to society [19]. Studies shows that unfair assessments have broader societal impacts such as contributing to higher school dropout rates [20]. Students who are unfairly assessed may become disengaged in the educational system, and may be pushed out (drop out) from school as described by Johnston-Goodstar and Roholt [21]. This, in turn, can result in a larger number of unproductive citizens, which

negatively affects society by increasing the burden on social services and reducing overall economic productivity [22]-[24]. However, recent research shows that unlike high-stakes assessments, culturally response assessments motivate students and promotes authentic, life-long learning [25,26]. Furthermore, students possess well-developed understanding and awareness of cultural issues are ready to engage in CRSciA [27]. GenAI's ability to generate personalized educational content demonstrates its potential to continue addressing the diverse needs of students through CRSciAs.

Culturally Responsive Assessments in Science Education

Over the years, K-12 science education has faced challenges in engaging all students in science learning. Research has constantly reported that students failed to use scientific knowledge in solving problems or designing solutions for real world problems. Moreover, traditional methods of teaching and assessment have historically led to the exclusion of diversity among successful science students [28,29]. Most assessments in K-12 science education narrow representations, leading to a growing achievements gap for minoritized students based on race, ethnicity, history, socio-economic status particularly in urban settings in countries like, the United States [30]. This call for CRSciA, which recognizes the diverse cultural perspectives and contributions in science, can aid in the achievement of underrepresented students in science.

Even though there has been an initial perception that science education is not suitable for culturally responsive assessments, this misconception has been cleared (Ladson-Billings, 2021). Recent studies have shown remarkable practices of CRSciAs, and even expanded to include Science, Technology, Engineering, Arts, and Mathematics subjects. This proves the value of culturally responsive assessments across all disciplines and educational levels (Ladson-Billings, 2021). Moreover, the developers of the *Framework for K-12 Science Education* [3] articulated a broader set of expectations to ensure that by the end of 12th grade, all students would possess sufficient knowledge of science to engage in their everyday lives, continue learning about science outside of school, and have the skills to enter careers of their choice. They emphasize on the phrase "all students" throughout this framework to provide equitable opportunities including assessment for all students to succeed in science [31].

Challenges of Implementing Culturally Responsive Assessments in Science Education

The implementation of CRSciAs faces several challenges with their effective adoption in science assessment. One major challenge arises from the limitations stemming from classroom teachers' identities and biases, as well as their limited knowledge of diverse cultural contexts [32,33]. Teachers' personal identities, including their race, ethnicity, and cultural background, can influence how they perceive and engage with CRSciAs [33]. For example, in science assessments, a teacher who lacks familiarity with the cultural experiences of their students may unintentionally introduce biases into the assessment process, either by favoring certain cultural narratives or by overlooking others.

Another challenge lies in the integration of CRSciA within the Framework for K-12 Science Education [6]. Though this is a reflective of a broader struggle within educational systems to effectively align CRSciAs with established curricular standards, it critically applies to science assessment [34]. Ladson-Billings [35] further critiques this misalignment by highlighting how many top-down initiatives to implement culturally responsive assessments often miss the mark; as states, districts, and professional organizations attempt to address cultural issues of assessments through various frameworks and guidelines, their efforts frequently fall short of the theory's original intent. More to the challenges of CRSciAs is inadequate resources. The lack of institutional support also makes it difficult for science teachers to effectively incorporate cultural responsiveness into their assessments, and the necessary tools, guidance, and resources, educators are unable to fully leverage CRSciA practices [36].

These challenges of teacher identity, bias, and misalignment with curricular standards in science education underscore the urgent need for continuous professional development. However, studies indicate that there is currently lack or insufficient continuous professional development available for teacher regarding CRSciAs [37,38]. For instance studies Harris, et al. [6] survey conducted

across 18 states in the US revealed that 86.36% of K-12 teachers view the integration of CRA with NGSS positively but suggested a more robust teacher training programs to enhance awareness and effective adoption of both NGSS and CRA in science classroom.

Generative AI and Culturally Responsive Assessment

Generative AI refers to advanced computational techniques that create new, meaningful content such as text, images, and audio from existing data [39]. These capabilities of GenAI can address significant challenges that teachers face in designing culturally responsive assessments, such as the time-intensive nature of creating materials that are both culturally relevant and pedagogically sound [12]. Furthermore, the interactive nature of GenAI-based assessments allows for real-time feedback and adaptation, providing teachers and students with immediate opportunities to learn and correct misunderstandings [40]. For instance, GPT-4's ability to interpret and generate multimodal content, including visual data like graphs and chemical diagrams, enhances its utility in crafting assessments that are aligned with cultural contexts and could engage students in ways that traditional automatic text-based assessments cannot [41]. This multimodal approach, grounded in the theory of multimedia learning can help overcome the limitations of traditional AI, which has been largely text-bound, by incorporating a broader spectrum of human experience into the assessment process [12,42]

3. Study Approach

This study adopts Jaakkola [16] model type conceptual framework approach designed to create conceptual frameworks that reveal and predict relationships between emerging key concepts. As the intersection of CRSciA and GenAI is relatively new, this approach allows for the untied exploration of this phenomenon without the constraints of data-related limitations. The primary objective of this study is to construct a Framework that predicts how GenAI can be leveraged to effectively address CRSciA [16]. The development started with the identification of key and dominant cultural tenets in science assessment, followed by exploring the potential capabilities of GenAI in relation to the cultural tenets, integrating into a GenAI-CRSciA framework and finally demonstrating with practical examples of the framework while acknowledging the challenges involved.

3. Generative AI Framework for Culturally Responsive Assessments in Science

While the spectrum of culture is broad, making it challenging to incorporate every aspect, this farmwork discusses the prevailing tenets from various studies, including *indigenous language*, *indigenous knowledge*, *ethnicity/race*, *religious beliefs*, and *community and family* (See figure 1). These tenets are the dominant cultural dimension rooted in their significant influence on students' science learning. Indigenous Language is a key medium of communication that shapes students' understanding of scientific concepts, with proficiency in the first language (L1) strongly linked to academic success in second language(L2) instructed programs, particularly in classrooms with growing numbers of English learners [17]-[19]. *Indigenous knowledge*, recognized as essential contextual knowledge, bridges traditional and scientific knowledge systems, enriching education by incorporating students' cultural perspectives [20,21]. Likewise, *ethnicity and race*, which carry deep social meanings beyond physical characteristics, are critical in addressing educational disparities and biases within science education [22,23]. Additionally, *religious beliefs* and the influence of *community and family* shape the moral and ethical frameworks within which students interpret scientific knowledge, underscoring the need for culturally attuned assessments [24]. These tenets within CRSciA collectively ensure that science assessment is equitable, inclusive, and reflective of the diversity present in classrooms globally [7,25,26].

3.1. Indigenous Language

Indigenous language plays a crucial role in culturally responsive assessment within K-12 education, particularly in science classrooms where students from diverse linguistic backgrounds are increasingly prevalent [27]. Wright and Domke [28] highlight the emphasis on language and literacy

in the Next Generation Science Standards (NGSS) and the C3 Framework for Social Studies, noting the importance of supporting students' disciplinary language development from the elementary grades. Kūkea Shultz and Englert [29] provide a compelling example of culturally and linguistically responsive assessment with the development of the Kaiapuni Assessment of Educational Outcomes (KĀ'EO) in Hawai'i. Recognizing the cultural and linguistic diversity of the Hawaiian population, the KĀ'EO assessment was designed to be culturally valid, addressing the unique needs of students in the Hawaiian Language Immersion Program. In the context of linguistically responsive assessment, the application of GenAI in educational settings has shown significant potential in addressing the diverse linguistic needs of students. In a study by Latif, et al. [30], G-SciEdBERT was developed to address the limitations of the standard German BERT (G-BERT) model in scoring written science responses. By pre-training G-SciEdBERT on a substantial corpus of German-written science responses and fine-tuning it on specific assessment items, the researchers demonstrated a marked improvement in scoring accuracy, with a 10% increase in the quadratic weighted kappa compared to G-BERT. This finding highlights the importance of contextualized GenAI in science assessment, where specialized models like G-SciEdBERT can significantly contribute to more culturally and linguistically science assessments, particularly in non-English language contexts.

3.2. Religion

Religious beliefs are a critical factor to consider in the development of CRA in K-12 science education. Mantelas and Mavrikaki [31] highlight the intersection of religiosity and scientific concepts, such as evolution, presents unique challenges for educators in CRSciA. Their study demonstrates that students with strong religious convictions may struggle with certain scientific ideas, which can negatively impact their academic performance if assessments do not account for these beliefs. This means CRSciA should allow students to demonstrate their scientific understanding without forcing them to choose between their religious beliefs and academic success [32].

Barnes, et al. [33] further stress the importance of considering religious backgrounds in CRA, particularly for students of color who may rely on religion as a critical support system. Their research shows that strong religious beliefs can influence students' acceptance of scientific theories like evolution, which in turn affects their academic success in science-related subjects. Culturally responsive assessments must therefore be designed to account for these factors, ensuring that they do not inadvertently disadvantage students whose religious beliefs differ from mainstream scientific views.

Owens, et al. [34] contribute to this discussion by advocating for a "pedagogy of difference" in teaching science, which could be extended to assessment practices in K-12 science education. This pedagogical approach encourages students to explore the relationship between their religious beliefs and scientific concepts, fostering an environment where multiple perspectives are acknowledged and valued. In the light of this, Sumarni, et al. [24] propose a holistic model for integrating religious and cultural knowledge into STEM education, which can serve as a foundation for CRA practices. Their RE-STEM model emphasizes the importance of bridging the gap between religion, culture, and science, suggesting that assessments should be designed to reflect this integration.

Although the intersection of GenAI and religious contexts remains underexplored, GenAI's potential for engaging in nuanced discussions about religious concepts is promising (Cheong, 2020). For instance, GenAI models like ChatGPT have shown the ability to participate in theological dialogues, offering responses that respect religious traditions (Oxenber, 2023). This capability allows for the development of assessments that honor students' religious beliefs, fostering a more inclusive environment. Nonetheless, ethical considerations are paramount as Ashraf (2022) warns that GenAI applications must be carefully monitored to avoid infringing on religious freedoms and to prevent bias or disrespect in digital interactions.

3.3. Indigenous Knowledge

Indigenous knowledge plays a pivotal role in shaping culturally responsive assessments in K-12 science education, offering a means to create an equitable achievement [35]. Trumbull and Nelson-

Barber [36] explore the challenges and opportunities in developing CRA for Indigenous students in the U.S., highlighting the limitations of standardized assessments that often disregard Indigenous knowledge systems. They argue that these assessments can be ineffective and even harmful, as they fail to engage Indigenous students or accurately measure their knowledge. This affirms Muhammad, et al. [37] assertion that that traditional assessments and curricula often overlook the historical and cultural contexts of Black and Brown children in the U.S., as highlighted by the National Assessment of Educational Progress (NAEP).

Therefore the concept of “culturally-valid assessment” was proposed by Trumbull and Nelson-Barber [36] to incorporate Indigenous ways of knowing and be responsive to the cultural and linguistic diversity of students. This approach is crucial for creating assessments that support the students’ academic success while also preserving and respecting their cultural identities [38]. Furthermore, Jin [39] systematically reviews educational programs aimed at supporting Indigenous students in science and STEM fields, revealing the positive impact of integrating Indigenous knowledge with Western scientific assessment. The review shows that culturally responsive assessment approaches in these programs lead to improved educational outcomes, as they allow Indigenous students to draw connections between their cultural heritage and the scientific concepts they are learning.

GenAI has the potential to challenge dominant Eurocentric narratives and promote the inclusion of Indigenous perspectives in K-12 science assessment (Washburn & McCutchen, 2024). For instance, GenAI tools can help ensure that science assessments honor and reflect Indigenous cultural identities. However, it is essential that GenAI-generated content is contextually accurate and respects the complexity of Indigenous cultures. Castro Nascimento and Pimentel (2023) emphasize the need for GenAI models to be trained on diverse cultural datasets to avoid perpetuating narrow perspectives. The deliberate integration of IK into GenAI models can significantly enhance the cultural relevance of science education.

3.4. *Race and Ethnicity*

CRSciA requires a thorough understanding of how race and ethnicity shape students’ learning experiences and outcomes. Atwater, et al. [40] emphasize that traditional science assessments often overlook the diverse cultural backgrounds of students, particularly those from African American, Latino, and Asian American communities. They therefore advocate for science assessments that are inclusive and reflective of the race and ethnicity within classrooms. Similarly, Wells [41] calls for strategic cross-sector collaboration between K-12 and higher education to address the sociocultural factors affecting diversity in education, reinforcing the need for assessments that are sensitive to the varied cultural contexts of students.

The importance of factoring race and ethnicity into science teaching and assessment is further highlighted by Riegle-Crumb, et al. [42], who found that inquiry-based learning is associated with more positive attitudes toward science among students from diverse racial and ethnic backgrounds. When assessments are designed to reflect the competencies of students’ cultural backgrounds, it allows them to demonstrate their understanding through exploration, critical thinking, and problem-solving in STEM [43]

Research by Choudhary [44] highlights the prevalence of racial bias in GenAI tools, including ChatGPT, underscoring the necessity of rigorous auditing to detect and mitigate these biases. In the context of CRA, AI tools must be designed to promote fairness and inclusivity, particularly in the assessment of students from diverse racial and ethnic backgrounds. Warr, et al. [45] provide evidence of racial bias affecting GenAI evaluations, demonstrating that racial descriptors can influence GenAI-generated scores. This underscores the importance of developing transparent and bias-tested AI tools that account for the unique cultural contexts of students. Ensuring that AI-driven assessments are equitable requires ongoing efforts to address and correct racial biases, providing a fair and accurate evaluation for all students.

3.5. *Family and Community Engagement*

While family and community might not be directly involved in the creation of assessment items, the role of family and community experts in validating cultural factors of assessments is crucial in the CRSciA framework. Family and community involvement are elements providing essential context and resources that directly impact students' academic performance and engagement. Denton, et al. [46] emphasize the significance of community cultural wealth in STEM education, particularly in K-12 settings. They argue that an assets-based approach, which recognizes the diverse forms of capital, such as familial, linguistic, and social, that students from nondominant communities bring, is essential for developing assessments that truly reflect students' backgrounds. In K-12 science assessments, this means creating assessment tools that account for the cultural and social capital that students acquire from their families and communities.

Gerde, et al. [47] provide insight into the specific ways that families contribute to science learning at the K-12 level. Their study shows that the resources and opportunities families provide, such as access to science-related books, toys, and community experiences like visits to parks and science centers, vary widely based on parents' self-efficacy and beliefs about science. This is an indication that CRSciA at this level should not only measure what students know from formal education but also integrate the informal learning experiences that occur within their family and community contexts.

Soto-Lara and Simpkins [48] further elaborate on the role of culturally grounded family support in the science education of Mexican-descent adolescents, focusing on the K-12 educational context. Their findings reveal that parents provide support through both traditional means, such as encouraging their children, and nontraditional methods, like indigenous cultural practices and leveraging extended family networks. They also found out that despite the parents facing challenges such as limited English proficiency and lack of formal science knowledge, they remain actively involved in their children's education. For K-12 assessments to be culturally responsive, they must recognize and incorporate these unique forms of family support.

Family and community involvement is a cornerstone of student success, particularly in science assessment. Garbacz, et al. [49] stress the importance of involving families and communities through ecological support. GenAI-driven tools can bridge the gap between school and home by providing culturally relevant resources and information, fostering a supportive learning environment. However, Said and Al-amadi [50] note challenges in engaging families, particularly in areas with limited digital literacy and technology access. GenAI has the potential to address these challenges by creating accessible communication channels between schools and families. The goal is to leverage GenAI not only for assessment but also to strengthen the connection between students' educational experiences and their broader family and community contexts, enhancing overall educational support and involvement.

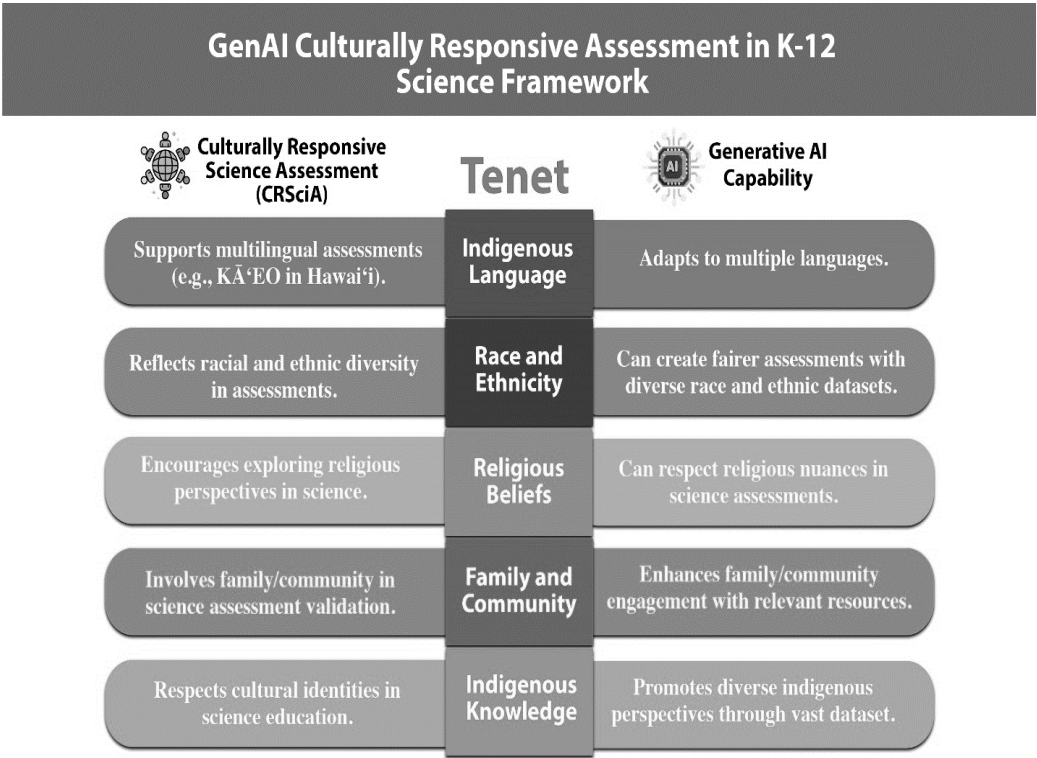


Figure 2. Generative AI Culturally Responsive Science Assessment Framework (GenAI-CRSciA).

Application of the GenAI-CRSciA Framework

In the application to determine the efficacy of the GenAI-CRSciA Framework, a seemingly experimental approach was taken to compare standard prompts, which are commonly employed by many educators when using GenAI, against the GenAI-CRSciA template-based prompts. The objective was to identify the aspects where the GenAI-CRSciA framework intersects with standard prompt strategies and to assess any differences in outcomes. Standard prompts represent the default approach that users, particularly teachers, might use without engaging in iterative refinement or technical adjustments [51]. Given that many educators may lack advanced prompt engineering skills, it is reasonable to assume that their use of GenAI would often rely on these straightforward prompts. Template-based prompts, like those developed using the GenAI-CRSciA framework, are more advanced. The GenAI-CRSciA template-based prompts incorporate the specific cultural tenets and undergo iterative refinement to produce the most culturally responsive science assessment form the NGSS Standards [52].

Use Cases of the GenAI-CRSciA Prompts

To apply the GenAI-CRSciA framework in practical examples, the NGSS Life Science questions were adapted to prompt ChatGPT 4o using both prompts, standard and GenAI-CRSciA Template-based prompts (see below). Found below the original and adapted questions for both prompts in relation to three students from Ghana, the USA, and China. The output indicated that GenAI (ChatGPT) was sensitive to the cultural backgrounds of the students, successfully generating content tailored to their national contexts.

Original Questions

NGSS standard for MS-LS2-2 [53]
MS-LS2-2_024_ Wolves and moose interactions
Ecosystem Interactions

Construct an explanation that predicts patterns of interactions among organisms across multiple ecosystems.

Activity: Wolves and moose interactions (ID# 116-04-Y03)

Construct an evidence-based argument about whether there exists a predator and prey relationship between wolves and moose.

From: [Page 1 Wolves and moose interactions \(ID# 116-04-Y03\) \(concord.org\)](#)

Source: This resource is part of the Concord Consortium's [Middle Grades \(6-8\)](#) project Task

Wolves and moose interactions (ID# 116-04-Y03)



Wolves



Moose

Wolf and moose images licensed under public domain.

Standard Prompt:

Prompt:

Create a culturally responsive science assessment question based on the NGSS standard MS-LS2-2, which involves ecosystem interactions and predator-prey relationships. The question should be suitable for middle school students in Ghana, the USA, and China. The question should ask students to construct an evidence-based argument about a predator and prey relationship, using examples that are relevant to their cultural context.

Output

Culturally Responsive Science Assessment Question:

NGSS Standard: MS-LS2-2

Concept: Ecosystem Interactions

Grade Level: Middle School

Target Regions: Ghana, the USA, China

Question:

In ecosystems around the world, animals interact as predators and prey. Think about a predator and prey relationship that is familiar in your cultural context or from your region (Ghana, the USA, or China). Construct an evidence-based argument about the relationship between these two organisms. Consider the following in your response:

Identify the predator and the prey in this relationship.

Describe the ecosystem where these interactions occur.

Provide evidence from your knowledge or experiences that shows how the predator and prey populations affect each other.

Explain how this relationship might be similar or different from other predator-prey relationships around the world.

Table 1.

GenAI-CRSciA Template-based prompt
Prompt:

Objective: Create a culturally responsive science assessment question that is scientifically rigorous and tailored to the cultural, linguistic, and educational contexts of students from [insert country/region, e.g., Ghana, the USA, China]. The question should be based on the [insert specific scientific concept or standard, e.g., ecosystem interactions, photosynthesis, forces and motion]. Instructions: Use the following guidelines to structure the assessment question: Language: Use [insert language, e.g., English, Mandarin, Twi] as the primary language for the question. Ensure the language used is appropriate for the educational level and cultural context of the students. Indigenous Knowledge: Incorporate local or indigenous knowledge relevant to the students' cultural backgrounds. Use examples that resonate with the students' experiences, such as [insert specific cultural examples related to the scientific concept, e.g., traditional ecological practices, folklore]. Race/Ethnicity: Respect and reflect the ethnic and cultural identities of the students. Include scenarios that are geographically and culturally relevant to the students' environment. Family: Encourage the student to engage with their family or community members to gather insights or stories related to the scientific concept being assessed. Ask the student to incorporate these familial insights into their response. Religious Beliefs: Allow for the exploration of religious or spiritual beliefs that might influence the student's understanding of the scientific concept. Ensure the question does not conflict with or marginalize any religious beliefs. Standards Construct a question that requires the student to engage with the scientific concept by analyzing data, making predictions, or constructing an evidence-based argument. Ensure the question aligns with the [insert specific scientific standard or concept, e.g., NGSS standard MS-LS2-2]. Provide clear instructions for the student on how to incorporate cultural elements into their response.
<i>Adapted Questions Using the GenAI-CRSciA Template-based prompt</i>
Ghana: Language: English Context: Consider how local knowledge and traditional stories in Ghana describe predator-prey relationships, such as those between lions and antelopes in the savanna ecosystem. Think about how the knowledge passed down through your family or community elders might help you understand similar relationships in other parts of the world. Task: Based on the graph and your understanding of predator-prey relationships, write an argument that shows whether you agree or disagree with Jason's explanation. Your argument should include: A statement about whether you agree or disagree with Jason's claim. Make sure to include Jason's claim if you agree with him or write your own claim if you disagree.

A statement about what is accurate or inaccurate in Jason’s evidence and reasoning.
Examples from the graph to show why Jason’s evidence is accurate or inaccurate.
Reflect on traditional Ghanaian stories about predator-prey relationships and discuss how they support or challenge Jason’s reasoning.



USA:

Language: English

Context: Reflect on how Native American perspectives describe wolves as spiritual beings and key regulators in ecosystems. Consider how this indigenous knowledge might influence your understanding of the predator-prey relationships shown in the graph.

Task:

Based on the graph and your understanding of predator-prey relationships, write an argument that shows whether you agree or disagree with Jason's explanation.

Your argument should include:

A statement about whether you agree or disagree with Jason’s claim. Make sure to include Jason’s claim if you agree with him or write your own claim if you disagree.

A statement about what is accurate or inaccurate in Jason’s evidence and reasoning.

Examples from the graph show why Jason’s evidence is accurate or inaccurate.

Incorporate Native American perspectives on wolves and discuss how these perspectives might support or challenge Jason’s reasoning.



China:

Language: Mandarin

Context: Consider how traditional Chinese views on predators like the Amur tiger can help you understand predator-prey relationships in ecosystems. Reflect on how your family’s stories or the teachings from Taoism or Buddhism might influence your interpretation of the graph.

任务:

根据图表以及你对捕食者与猎物关系的理解，写一篇论述，说明你是否同意 Jason 的解释。

你的论述应包括：

关于你是否同意 Jason 观点的陈述。如果你同意他的观点，请务必包含 Jason 的论点；如果你不同意，请写下你自己的观点。

关于 Jason 的证据和推理中哪些内容是准确的或不准确的陈述。

从图表中举例说明为什么 Jason 的证据是准确的或不准确的。

结合中国传统文化中对捕食者（如东北虎）的看法，讨论这些文化视角如何支持或挑战 Jason 的推理。

(Translation: Task:

Based on the graph and your understanding of predator-prey relationships, write an argument that shows whether you agree or disagree with Jason's explanation.

Your argument should include:

A statement about whether you agree or disagree with Jason’s claim. Make sure to include Jason’s claim if you agree with him or write your own claim if you disagree.

A statement about what is accurate or inaccurate in Jason’s evidence and reasoning.

Examples from the graph to show why Jason’s evidence is accurate or inaccurate.

Incorporate traditional Chinese perspectives on predators like the Amur tiger, and discuss how these cultural views might support or challenge Jason’s reasoning.)



4. Discussion

The alignment of the CRSciA framework with the unique cultural contexts of Ghana, the USA, and China involves careful adaptation across several key tenets. In terms of language, the questions are presented in English for both Ghana and the USA, where it is the dominant language of instruction, while in China, the questions are delivered in Mandarin [54,55]. This ensures that students can engage fully with the content in their native or dominant instructional language. For instance, the efficiency of the GenAI-CRSciA framework becomes particularly evident when addressing the challenges faced by students in linguistically diverse environments, such as China and Ghana. In China, despite policies aimed at incorporating English as a medium of instruction, research shows that Chinese students often struggle in predominantly English-medium contexts [56]. The GenAI-CRSciA framework effectively mitigates these challenges by tailoring content to specific cultural and linguistic contexts, such as translating content into Chinese and aligning it with the students’ cultural backgrounds and linguistic needs [57]. Similarly, in Ghana, where the language-in-education policy mandates the use of mother tongue (L1) as the medium of instruction in lower

primary grades (KG1 to Primary 3) and English (L2) from Primary four onwards, the GenAI-CRSciA framework demonstrates its cultural awareness by generating science assessments predominantly in English. This aligns with the reality of Ghana's educational system, where English remains the dominant medium of instruction and assessment, particularly in science education [58,59].

Indigenous knowledge is also thoughtfully integrated, with traditional Ghanaian stories and proverbs highlighting the relationship between lions and antelopes, Native American perspectives shedding light on the spiritual and ecological significance of wolves and moose, and Chinese history and folklore contextualizing the interaction between Amur tigers and sika deer. For example, in Ghana, the framework integrates culturally significant symbols like the lion and deer, which are deeply embedded in local folklore and traditions. The '*Aboakyire Festival*', celebrated by the people of Winneba, involves the symbolic hunting of a deer, an event that carries profound cultural and spiritual significance [60]. Similarly, among the Akans of Akyem Abuakwa, animal symbolism, including that of the lion and deer, plays a crucial role in cultural identity and educational practices, conveying lessons and wisdom that are essential to the community's heritage [61]. Likewise, in China, the framework demonstrates its effectiveness by incorporating ecologically significant animals such as the Amur tiger and sika deer into educational content, aligning the assessments with the local knowledge and values of communities in northeast China [62].

Further adaptations reflect the importance of race and ethnicity, where the symbolic significance of lions is considered within various Ghanaian ethnic groups, differing views on wolves and moose are acknowledged among Native American tribes and European settlers in the USA, and the cultural heritage of Amur tigers is recognized within both the Han Chinese majority and ethnic minorities [63]. Family engagement is encouraged across all regions, with students asked to gather insights and stories from their elders, enhancing the personal relevance of the assessment. Finally, religious beliefs are explored, with Ghanaian students reflecting on traditional animistic beliefs, American students considering how spiritual beliefs shape their views on wolves and moose, and Chinese students examining Taoist or Buddhist concepts of balance and harmony as they relate to the relationship between Amur tigers and sika deer [64,65]. This comprehensive approach ensures that the assessments are scientifically and deeply rooted in the cultural realities of the students.

5. Conclusions and Future Directions

Over the years, science assessment has faced significant equity challenges due to the complexity of culture, time constraints, and the skills required by teachers to generate assessments that are culturally sensitive to each student's dominant background. Despite efforts made within the NGSS, these challenges persist. This study, therefore, highlighted the core and critical tenets (including indigenous/dominant language, race/ethnicity, indigenous knowledge, religious beliefs, and family and community) that constitute CRSciA and aligned them with the capabilities of GenAI. These tenets were instrumental in developing the GenAI-CRSciA framework and subsequently GenAI-CRSciA Template-based prompt to generate assessment items from NGSS while comparing it with standard prompt strategies. The results showed that the GenAI-CRSciA approach outperformed standard prompt strategies when queried to generate CRA aligned with NGSS standards, proving its effectiveness in addressing the ongoing CRA challenges in science assessment.

Even though the GenAI-CRSciA framework presented assessment that covers range of the cultural backgrounds, it presents challenges such as the overgeneralization of students' backgrounds, compromising of curricula core competencies that vary in differences, that is, overlooking subcultural variations within a country or region, such as differences between urban and rural students, or between different ethnic or linguistic groups within the same nation which can lead to bias and less effective CRA output. To address this, we recommend that teachers provide comprehensive and specific background information about their students in the template, ensuring that the generated content is truly reflective of the students' unique cultural contexts. In larger class settings situations, teachers could implement comprehensive local data indexing, where detailed student background information is uploaded and retrievable for use in crafting more personalized assessments.

We strongly recommend educators to make conscious efforts to validate the GenAI-generated assessments with input from human experts, including community or family members, teachers, and even student, culturally responsive educators who are knowledgeable about the cultural nuances. Research indicates that current teacher education reforms have not adequately prepared teacher

candidates to meet the needs of culturally diverse student populations, as specialized training is often lacking [66,67]. We propose continuous professional development to enhance teachers' understanding of the cultural tenets and examples that may be generated by the template. This will enable them to better grasp the fundamentals of culture and assessments, ensuring they can effectively utilize the template to create equitable science assessments.

Author Contributions: "Conceptualization, Matthew Nyaaba and Xiaoming Zhai; methodology, Matthew Nyaaba; validation, Xiaoming Zhai. and Morgan Z. Faison.; writing—original draft preparation, Matthew Nyaaba; writing—review and editing, Xiaoming Zhai and Morgan Z. Faison.; visualization, Matthew Nyaaba; supervision, Xiaoming Zhai. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Frommert, C. *Creating an environment where all students see that they can be good at math*, <<https://www.edutopia.org/article/helping-students-see-themselves-good-math>> (2023).
2. Kouo, J. L. in *Research Anthology on Physical and Intellectual Disabilities in an Inclusive Society* 873-889 (IGI Global, 2022).
3. NRC. *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. (National Academies Press, 2012).
4. Mullis, I. V., Martin, M. O. & von Davier, M. TIMSS 2023 Assessment Frameworks. *International Association for the Evaluation of Educational Achievement* (2021).
5. States, N. N. L. *Next generation science standards: For states, by states*. (National Academies Press, 2013).
6. Harris, K., Sithole, A., Kibirige, J. & McCarthy, P. The Next Generation Science Standards and the quest towards culturally responsive pedagogy: Perceptions of K-12 educators. *American Academic Scientific Research Journal for Engineering, Technology, and Sciences* **49**, 192-208 (2018).
7. Nortvedt, G. A. *et al.* Aiding culturally responsive assessment in schools in a globalising world. *Educational Assessment Evaluation and Accountability* **32**, 5-27 (2020). <https://doi.org/10.1007/s11092-020-09316-w>
8. Preston, J. P. & Claypool, T. R. in *Frontiers in Education*. 679972 (Frontiers Media SA).
9. Wu, X., He, X., Liu, T., Liu, N. & Zhai, X. 401-413 (Springer Nature Switzerland).
10. Lee, G.-G. *et al.* Multimodality of ai for education: Towards artificial general intelligence. *arXiv preprint arXiv:2312.06037* (2023).
11. Li, X., Li, B. & Cho, S.-J. Empowering Chinese language learners from low-income families to improve their Chinese writing with ChatGPT's assistance afterschool. *Languages* **8**, 238 (2023).
12. Bewersdorff, A. *et al.* Taking the Next Step with Generative Artificial Intelligence: The Transformative Role of Multimodal Large Language Models in Science Education. *arXiv preprint arXiv:2401.00832* (2024).
13. Zhai, X. ChatGPT for next generation science learning. *XRDS: Crossroads, The ACM Magazine for Students* **29**, 42-46 (2023).
14. Hwang, K., Challagundla, S., Alomair, M., Chen, L. K. & Choa, F.-S. in *Workshop on Generative AI for Education*.
15. Richards, M. *et al.* Bob or Bot: Exploring ChatGPT's Answers to University Computer Science Assessment. *ACM Trans. Comput. Educ.* **24**, Article 5 (2024). <https://doi.org/10.1145/3633287>
16. Jaakkola, E. Designing conceptual articles: four approaches. *AMS Review* **10**, 18-26 (2020). <https://doi.org/10.1007/s13162-020-00161-0>
17. Doiz, A. & Lasagabaster, D. Dealing with language issues in English-medium instruction at university: A comprehensive approach. *International Journal of Bilingual Education and Bilingualism* **23**, 257-262 (2020).
18. Statistics, N. C. f. E. *English learners in public schools*, <<https://nces.ed.gov/programs/coe/indicator/cgf>> (2024).
19. Kim, M., Crossley, S. A. & Kim, B.-K. Second language reading and writing in relation to first language, vocabulary knowledge, and learning backgrounds. *International Journal of Bilingual Education and Bilingualism* **25**, 1992-2005 (2022).
20. Jovchelovitch, S. *Knowledge in context: Representations, community and culture*. (Routledge, 2019).
21. Melak, A. T. The Verbal Usage of עֵבֶר in the Book of Psalms: Exploring its Contextual Meaning and Theological Implications. *Pan-African Journal of Theology* (2024). <https://doi.org/10.56893/pajot2024v03i01.06>
22. Assari, S., Mardani, A., Maleki, M., Boyce, S. & Bazargan, M. Black-White achievement gap: Role of race, school urbanity, and parental education. *Pediatric health, medicine and therapeutics*, 1-11 (2021).
23. Zhang, W. & Koshmanova, T. Understanding the impact of race, socioeconomic status on student achievement for secondary school students. *International Journal of Education and Human Developments* **6**, 5-10 (2020).

24. Sumarni, W., Faizah, Z., Subali, B. & Wiyanto, W. The Urgency of Religious and Cultural Science in STEM Education: A Meta Data Analysis. *International Journal of Evaluation and Research in Education* **9**, 1045-1054 (2020).
25. Balzer, M. M. *The tenacity of ethnicity: A Siberian saga in global perspective*. (Princeton University Press, 2021).
26. Baker, D. *The schooled society: The educational transformation of global culture*. (Stanford University Press, 2020).
27. Rosheim, K. M., Tamte, K. G. & Froemming, M. J. Reducing Inequalities Inherent in Literacy Assessment of Multilingual Learners. *Reading Psychology* (2024). <https://doi.org/10.1080/02702711.2024.2359922>
28. Wright, T. S. & Domke, L. M. The role of language and literacy in K-5 science and social studies standards. *Journal of Literacy Research* **51**, 5-29 (2019).
29. Kūkea Shultz, P. & Englert, K. in *Frontiers in Education*. 701973 (Frontiers Media SA).
30. Latif, E., Lee, G.-G., Neuman, K., Kastorff, T. & Zhai, X. G-SciEdBERT: A Contextualized LLM for Science Assessment Tasks in German. *arXiv preprint arXiv:2402.06584* (2024).
31. Mantelas, N. & Mavrikaki, E. Religiosity and students' acceptance of evolution. *International Journal of Science Education* **42**, 3071-3092 (2020).
32. Black, P. Christian beliefs and values in science and religious education: An essay to assist the work of teachers of both subjects. *International Studies in Catholic Education* **9**, 206-222 (2017).
33. Barnes, M. E. *et al.* Relationships between the religious backgrounds and evolution acceptance of Black and Hispanic biology students. *CBE – Life Sciences Education* **19**, ar59 (2020).
34. Owens, D. C., Pear, R. S., Alexander, H. A., Reiss, M. J. & Tal, T. Scientific and religious perspectives on evolution in the curriculum: An approach based on pedagogy of difference. *Research in Science Education* **48**, 1171-1186 (2018).
35. Black, A. & Tylianakis, J. M. Teach Indigenous knowledge alongside science. *Science* **383**, 592-594 (2024).
36. Trumbull, E. & Nelson-Barber, S. The Ongoing Quest for Culturally-Responsive Assessment for Indigenous Students in the US. *Frontiers in Education* **4** (2019). <https://doi.org/10.3389/feduc.2019.00040>
37. Muhammad, G. E., Ortiz, N. A. & Neville, M. L. A Historically Responsive Literacy Model for Reading and Mathematics. *Reading Teacher* **75**, 73-81 (2021). <https://doi.org/10.1002/trtr.2035>
38. Azam, S. & Goodnough, K. Learning together about culturally relevant science teacher education: Indigenizing a science methods course. *International Journal of Innovation in Science and Mathematics Education* **26** (2018).
39. Jin, Q. Supporting indigenous students in science and STEM education: A systematic review. *Education Sciences* **11**, 555 (2021).
40. Atwater, M. M., Lance, J., Woodard, U. & Johnson, N. H. Race and ethnicity: Powerful cultural forecasters of science learning and performance. *Theory Into Practice* **52**, 6-13 (2013).
41. Wells, A. S. Racial, ethnic, and cultural diversity across K-12 and higher education sectors: challenges and opportunities for cross-sector learning. *Change: The Magazine of Higher Learning* **52**, 56-61 (2020).
42. Riegle-Crumb, C., Morton, K., Nguyen, U. & Dasgupta, N. Inquiry-based instruction in science and mathematics in middle school classrooms: Examining its association with students' attitudes by gender and race/ethnicity. *AERA open* **5**, 2332858419867653 (2019).
43. Lesseig, K., Firestone, J., Morrison, J., Slavitt, D. & Holmlund, T. An analysis of cultural influences on STEM schools: Similarities and differences across K-12 contexts. *International Journal of Science and Mathematics Education* **17**, 449-466 (2019).
44. Choudhary, T. Reducing Racial and Ethnic Bias in AI Models: A Comparative Analysis of ChatGPT and Google Bard. (2024).
45. Warr, M., Pivovarova, M., Mishra, P. & Oster, N. J. Is ChatGPT Racially Biased? The Case of Evaluating Student Writing. *The Case of Evaluating Student Writing (May 25, 2024)* (2024).
46. Denton, M., Borrego, M. & Boklage, A. Community cultural wealth in science, technology, engineering, and mathematics education: A systematic review. *Journal of Engineering Education* **109**, 556-580 (2020).
47. Gerde, H. K., Pikus, A. E., Lee, K., Van Egeren, L. A. & Huber, M. S. Q. Head Start children's science experiences in the home and community. *Early Childhood Research Quarterly* **54**, 179-193 (2021).
48. Soto-Lara, S. & Simpkins, S. D. Parent support of Mexican-descent high school adolescents' science education: A culturally grounded framework. *Journal of Adolescent Research* **37**, 541-570 (2022).
49. Garbacz, S. A., Herman, K. C., Thompson, A. M. & Reinke, W. M. Family engagement in education and intervention: Implementation and evaluation to maximize family, school, and student outcomes. *Journal of school psychology* **62**, 1-10 (2017).
50. Said, Z. & Al-amadi, A. The Role of Family Engagement in Students' Science Learning in Qatari Schools. *The Eurasia Proceedings of Educational and Social Sciences* **11**, 142-152 (2018).
51. Spasić, A. J. & Janković, D. S. in *2023 58th International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST)*. 47-50 (IEEE).
52. Seo, J. *et al.* Plain template insertion: korean-prompt-based engineering for few-shot learners. *IEEE Access* **10**, 107587-107597 (2022).

53. Consortium, C. *Wolves and moose interactions* (ID# 116-04-Y03), <<https://scienceeducation.stanford.edu/assessments>> (2024).
54. Erling, E. J., Adinolfi, L., Hultgren, A. K., Buckler, A. & Mukorera, M. Medium of instruction policies in Ghanaian and Indian primary schools: An overview of key issues and recommendations. *English as a Medium of Instruction in Postcolonial Contexts*, 18-34 (2018).
55. Loh, E. K., Tam, L. C. & Lau, K.-c. Moving between language frontiers: The challenges of the medium of instruction policy for Chinese as a second language. *Language Policy* **18**, 131-153 (2019).
56. Zhang, Z. English-medium instruction policies in China: Internationalisation of higher education. *Journal of Multilingual and Multicultural Development* **39**, 542-555 (2018).
57. Yu, S., Wang, Y., Jiang, L. & Wang, B. Coping with EMI (English as a medium of instruction): Mainland China students' strategies at a university in Macau. *Innovations in Education and Teaching International* **58**, 462-472 (2021).
58. Owu-Ewie, C. & Eshun, E. S. The Use of English as Medium of Instruction at the Upper Basic Level (Primary Four to Junior High School) in Ghana: From Theory to Practice. *Journal of Education and Practice* **6**, 72-82 (2015).
59. Owu-Ewie, C. & Eshun, E. S. Language representation in the Ghanaian lower primary classroom and its implications: the case of selected schools in the Central and Western Regions of Ghana. *Current Issues in Language Planning* **20**, 365-388 (2019).
60. Akyeampong, O. A. Aboakyer: traditional festival in decline. *Ghana Social Science Journal* **16**, 97 (2019).
61. Lumor, F. *Significance of animal symbolism among the Akans of Akyem Abuakwa traditional area* Master of Arts in Art Education thesis, Kwame Nkrumah University of Science and Technology, (2009).
62. Li, Y. *et al.* Community attitudes towards Amur tigers (*Panthera tigris altaica*) and their prey species in Yanbian, Jilin province, a region of northeast China where tigers are returning. *Plos one* **17**, e0276554 (2022).
63. Rudenko, S. & Sobolievskyi, Y. Philosophical ideas in spiritual culture of the indigenous peoples of North America. *Anthropological Measurements of Philosophical Research*, 168-182 (2020).
64. Kosoe, E. A., Adjei, P. O.-W. & Diawuo, F. From sacrilege to sustainability: the role of indigenous knowledge systems in biodiversity conservation in the Upper West Region of Ghana. *GeoJournal* **85**, 1057-1074 (2020).
65. Tavor, O. (Oxford University Press US, 2024).
66. Faison, M. Z.-J. *Let the Circle Be Unbroken: Culturally Responsive Pedagogy and Exemplary, Novice, Elementary, African American Teachers*, Emory University, (2016).
67. Ladson-Billings, G. in *The Educational Forum*. 351-354 (Taylor & Francis).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.