

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Not peer-reviewed version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: 21 October 2024

doi: 10.20944/preprints202409.1276.v2

Keywords: Culturally Responsive Assessment; Assessment; Prompt Engineering; Generative AI (GenAI); AI; Science Education; customized GPT; Dynamic Prompt



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 Science Assessment: A Conceptual Framework

Matthew Nyaaba<sup>1,2,\*</sup>, Xiaoming Zhai<sup>1,3,\*</sup> and Morgan Z. Faison<sup>2</sup>

<sup>1</sup> AI4STEM Education Center, University of Georgia, USA

<sup>2</sup> Department of Educational Theory and Practice, University of Georgia, USA

<sup>3</sup> 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 a novel approach to automatic generation of cultural and context-specific science assessments for K-12 education using generative AI (GenAI). We first developed a GenAI Culturally Responsive Science Assessment (GenAI-CRSciA) framework that establishes the relationship between CRSciA and GenAI, by incorporating key cultural tenets such as *indigenous language, Indigenous knowledge, ethnicity/race, and religion*. The CRSciA framework along with dynamic prompt strategies were used to develop the CRSciA-Generator model within the OpenAI platform. The CRSciA-Generator allows users to automatically generate assessments tailored to students' cultural and contextual needs. In a pilot comparison test between the CRSciA-Generator and the base GPT 4o (with standard prompt), the models were tasked with generating CRSciAs that aligned with the Next Generation Science Standard on *predator and prey relationship* for students from Ghana, the USA, and China. The results showed that the CRSciA-Generator output assessments incorporated more tailored culturally and context assessment items for each specific group with examples, such as traditional stories of lions and antelopes in Ghana, Native American views on wolves in the USA, and Taoist or Buddhist teachings on the Amur tiger in China than the standard prompt out within the base GPT 4o. However, due to the background information provided, the CRSciA-Generator overgeneralized its output focusing on broad national contexts, treating entire countries as culturally homogenous and neglecting the subcultures. Therefore, we recommend that teachers provide detailed background information about their students when using the CRSciA-Generator. Additionally, we believe the pilot test did not fully validate the model's efficacy, and future studies involving human experts' review are recommended to evaluate the cultural and contextual validity of the generated assessments. We also suggest empirical studies in diverse contexts to further test and validate the model's overall effectiveness.

**Keywords:** culturally responsive assessment; assessment; prompt engineering; Generative AI (GenAI); AI; science education; customized GPT; dynamic prompt

## 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 science 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. Consequently, K-12 science assessments, guided by frameworks like the Next Generation Science Standards (NGSS), aim to foster scientific literacy in students [3,4]. This approach seeks to cultivate a profound grasp of fundamental scientific concepts and the capacity to utilize this knowledge in addressing real-world challenges [5,6]. This goal encourages the teaching and assessment of science to engage students in practices that are

meaningful to their cultural experiences as assets in their learning processes [6,7]. This has led to several initiatives, such as culturally responsive assessment of connecting science concepts to students' cultural experiences and leveraging on their background experiences as an asset to their learning [8]. Researchers suggest culturally responsive science assessments (CRSciAs) make learning more relevant and engaging to students [9]. However, implementing CRSciAs in K-12 classrooms remains challenging. Teachers face difficulties due to increasing classroom diversity and the dominance of traditional, western-centric assessments, which often disadvantage the migrant, historically marginalized, and indigenous students [10]. Furthermore, CRSciA practices, though essential, are time-consuming and complex, making them hard to scale effectively [11,12].

The recent development of Generative Artificial Intelligence (GenAI) offers a promising solution to CRSciAs by enabling more equitable, culturally tailored assessments that better reflect students' diverse backgrounds [13–15]. GenAI can provide multimodal learning opportunities [16], handling a variety of cultural contexts such as adapting to different languages, symbols, and local knowledge more often and more effectively than what 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 [17].

Nonetheless, the use of GenAI in assessments generation is not without its challenges or the current rapid use of ChatGPT for automatic assessment [18]. There are legitimate concerns about the fairness, cultural biases and inaccuracies of GenAI-generated assessments, especially when it comes to avoiding bias and cultural stereotypes to ensure that all students are assessed equitably [19–21]. For instance, Kiyak [22], in their review study on GenAI and automatic assessment, highlighted the critical issue of GenAI-generated assessments as they lack contextual knowledge which relates to languages and indigenous knowledge. Moreover, Chan et al. [23], in their study exploring ChatGPT's potential in question generation, also identified cultural biases concerning race and ethnicity, and religious beliefs in the generated questions [24]. These studies, among others, recommended the need for datasets and frameworks that can guide GenAI prompt engineering to consider the diverse cultural background of students as assets in assessment generation. These recommendations emphasize the urgent and critical importance of addressing these issues, as without careful consideration, there is a potential risk that GenAI could hegemonically reinforce the very inequities they are intended to overcome in education through assessment generation [25].

To address the gaps, this study seeks to develop a conceptual framework that brings together the capabilities of GenAI and the core tenets of CRSciA for K-12 science assessments. Specifically, we first develop a GenAI-CRSciA conceptual framework by identifying the key concepts of cultural responsive assessments with GenAI and articulating the relationships between them [26]. This lays out the foundation for the CRSciA Generator within the OpenAI model. We then integrate the conceptual framework into the GPT model to create the CRSciA Generator. This involves a dynamic prompt method based on the CRSciA-Framework. This prompt approach dynamically initiates conversation and further interacts with users to generate assessments items tailored to their students' cultural and context-specific needs based on the information provided by the user [27]. Finally, we pilot the CRSciA Generator, by comparing it with base GPT 4o and standard prompt within the cultures of US, Ghana, and China. The findings show that the CRSciA-Generator has the potential to automatically generate cultural and context-specific science assessments. This has the potential to acknowledge students' cultural backgrounds as an asset to scientific literacy.

## **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 the discourse about CRSciA, the challenges that persist in implementing CRSciA, and finally, the capabilities of GenAI in addressing these challenges.

### *2.1. 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 countries like the U.S. with diverse student populations [28,29]. The term refers to the disparities in standardized test scores between Native Indigenous, Black, Latina/o, recent immigrants, and White students [30]. While this is a concern, Ladson-Billings [28] 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 [31].

Beyond the supposed “achievement gap,” CRSciA aims to address the educational debt in science. 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 [32]. Studies show that unfair assessments have broader societal impacts, such as contributing to higher school dropout rates [33]. Students who are unfairly assessed may become disengaged in the educational system and may be pushed out (drop out) from school [33]. 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 [34–36]. However, recent research shows that, unlike high-stakes assessments, culturally responsive assessments motivate students and promote authentic, life-long learning [37]. Furthermore, students who possess a well-developed understanding and awareness of cultural issues are ready to engage in CRSciA [38]. GenAI’s ability to generate personalized educational content demonstrates its potential to continue addressing the diverse needs of students through CRSciAs.

## 2.2. *Culturally Responsive Assessments in Science Education*

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 [39]. Moreover, the *Framework for K–12 Science Education* articulates a broader set of expectations to ensure that by the end of 12th grade [3]. The assumption is 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 pursue careers of their choice. They emphasize the phrase “all students” throughout this framework to provide equitable opportunities, including assessment, for all students to succeed in science [40].

## 2.3. *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 [41]. Teachers’ personal identities, including their race, ethnicity, and cultural background, can influence how they perceive and engage with CRSciAs [42]. 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 [3]. Though this is a reflection of a broader struggle within educational systems to effectively align CRSciAs with established curricular standards, it critically applies to science assessment [11,43,44]. Ladson-Billings [39]. Although this is a reflection of a broader struggle within educational systems to effectively align CRSciAs with established curricular standards, it critically

applies to science assessment [11,43,44]. Ladson-Billings [39] further critiques this misalignment by highlighting how 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 with the necessary tools, guidance, and resources, educators are unable to fully leverage CRSciA practices [45].

These challenges of teacher identity, bias, misalignment with curricular standards in science education, and a lack of institutional support and resources for culturally responsive teaching and assessment from states, districts, and professional organizations underscore the urgent need for continuous professional development. Moreover, studies indicate that there is currently a lack of continuous professional development available for teachers regarding CRSciAs [46]. For instance, studies by Harris et al. [7] surveyed teachers 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.

#### 2.4. 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 [14]. 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 [47,48]. 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 [49,50]. 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 [16]. 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 [14,16,51].

## 4. Generative AI Framework for Culturally Responsive Assessments in Science

Our study framework was grounded in cultural tenets that are central to shaping individuals' identities and learning experiences in science as well as GenAI capabilities. At the moment, GenAI has the ability to translate over 50 languages, an essential feature that plays a significant role in assessment generation. Studies also show that, with proper prompts, GenAI could reduce race and ethnicity biases [52–54]. Furthermore, GenAI is noted for demonstrating respect for religious and cultural difference and also has the potential to enhance indigenous knowledge [55–58]. These qualities of GenAI largely influenced the selection of these cultural tenets [59]. We acknowledge other cultural tenets, such as socio-economic status and gender, and advocate for future frameworks as the field of GenAI evolves [60]. The framework specifically focused on these prevailing tenets *indigenous language, indigenous knowledge, ethnicity/race, religious beliefs, and community and family* (See Figure 1).

### 4.1. 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 [43]. Wright and Domke [61] 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 [62] 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, researchers designed the KĀ‘EO assessment 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. [63], 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 cultural and linguistic science assessments, particularly in non-English language contexts.

#### 4.2. Religion

Religious beliefs are a critical factor to consider in the development of CRA in K-12 science education. Mantelas and Mavrikaki [64] highlighted that 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 [65].

Barnes et al. [66] 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. Barnes et al. [66] 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. [67] 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 light of this, Sumarni et al. [68] proposed 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 [69]. For instance, GenAI models like ChatGPT have shown the ability to participate in theological dialogues, offering responses that respect religious traditions [58]. 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 [70] advocates that GenAI applications must be carefully monitored to avoid infringing on religious freedoms, bias and disrespect in digital interactions.

#### 4.3. 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 [71]. Trumbull and Nelson-Barber [72] explored 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 argued 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. [31] 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 [72] to incorporate Indigenous ways of knowing and to be responsive to the cultural and linguistic diversity of students. This approach is crucial for creating assessments that support students’ academic success while also preserving and respecting their cultural identities [73]. Furthermore, Jin [74] systematically reviewed 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 [75]. 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 [76] emphasized the need for GenAI models to be trained on diverse cultural datasets to avoid perpetuating narrow perspectives. The deliberate integration of indigenous knowledge into GenAI models can significantly enhance the cultural relevance of science education.

#### *4.4. Race and Ethnicity*

CRSciA requires a thorough understanding of how race and ethnicity shape students’ learning experiences and outcomes. Atwater et al. [77] emphasized that traditional science assessments often overlook the diverse cultural backgrounds of students, particularly those from African American, Latino, and Asian American communities. They, therefore, advocated for science assessments that are inclusive and reflective of the race and ethnicity within classrooms. Similarly, Wells [78] called 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 was further highlighted by Rieggle-Crumb et al. [79], 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 [80].

Research by Choudhary [81] 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, GenAI tools must be designed to promote fairness and inclusivity, particularly in the assessment of students from diverse racial and ethnic backgrounds. Warr et al. [82] provided evidence of racial bias affecting GenAI evaluations, demonstrating that racial descriptors can influence GenAI-generated assessments. This underscores the importance of developing transparent and bias-tested AI tools that account for the unique cultural contexts of students. Ensuring that GenAI-driven assessments are equitable requires ongoing efforts to address and correct racial biases, providing a fair and accurate evaluation for all students.

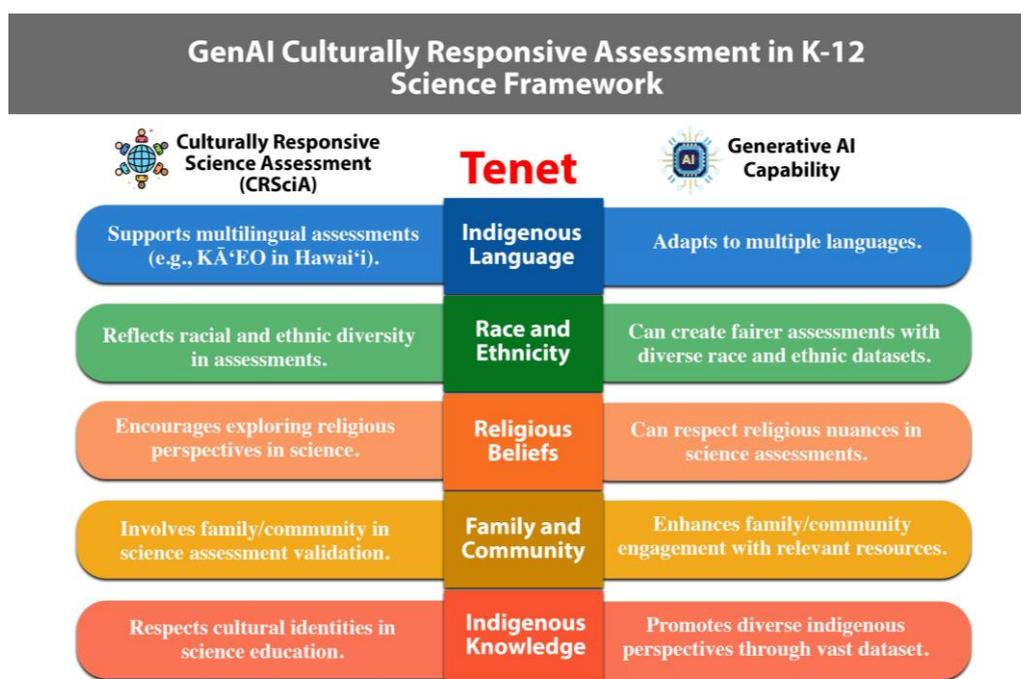
#### *4.5. Family and Community Engagement*

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. [83] emphasized the significance of community cultural wealth in STEM education, particularly in K-12 settings. They argued 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. [84] 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 and 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 [85] further elaborated 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.

Family and community involvement is a cornerstone of student success, particularly in science assessment. Garbacz et al. [86] stressed 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 [87] noted 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.

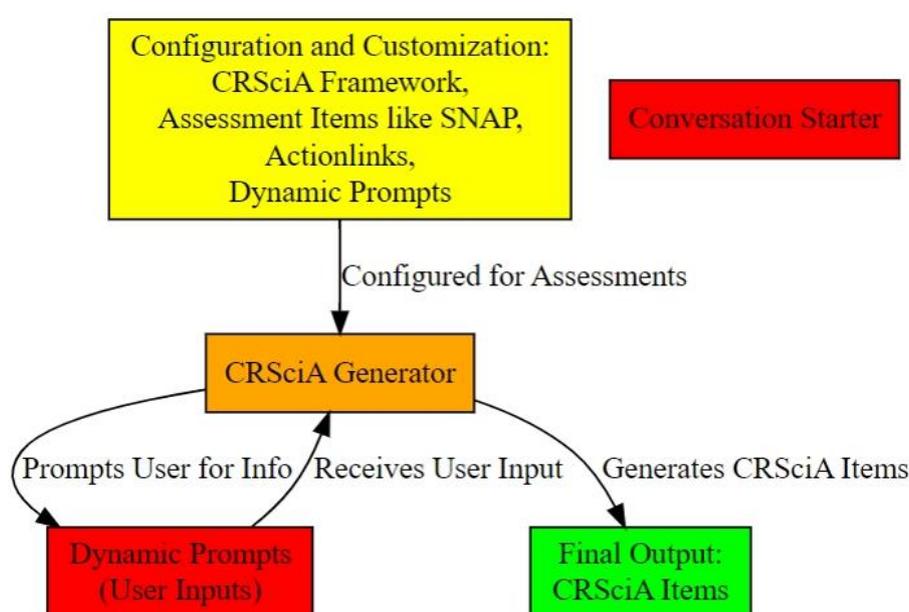


**Figure 1.** Generative AI Culturally Responsive Science Assessment Framework (GenAI-CRSciA).

## 5. Developing the CRSciA-Generator

The OpenAI platform offers the opportunity to configure and customize GPT models as specialized ChatGPTs (such as Code Tutor) for particular use cases [88]. This configuration allows users to tailor the model to meet specific needs, making it more relevant and effective for specialized tasks based on context. The process is essential for the study because base GPT models lack domain-specific focus and are not fine-tuned for specialized tasks or contexts, making them more prone to biases or hallucinations [89].

The development of the CRSciA Generator follows a three-step process: *Configuration and Customization Process*, *Prompt Engineering*, and *Final Output*. These processes were designed to ensure that teachers or users could engage seamlessly and generate CRSciA items for their class without facing inaccurate outputs. With these steps, the model was made more user-friendly, allowing users of all prompt skills to follow through and get the best possible output from the system (See Figure 2).



**Figure 2.** CRSciA-Generator System.

### 5.1. Configuration and Customization

The CRSciA Generator is grounded in rigorous configuration and customization processes, starting with the integration of the GenAI-CRSciA frameworks and SNAP items. These documents are uploaded to the model as files. A critical aspect of customization involves optimizing the uploaded files for efficient processing while staying within the token limits of the GPT model. Exceeding the token capacity could have resulted in incomplete data processing. To mitigate this, we choose to upload PDF files instead of DOC/DOCX files, primarily due to their smaller file size. PDFs are generally more compact because they are designed for viewing rather than editing, allowing for content compression without a loss in quality [90,91]. The use of PDF is particularly important for SNAP items, which contain both text and visual components.

To support the creation of dynamic and interactive assessments, the configuration involves the incorporation of advanced tools such as web search, DALL-E image generation, and a code interpreter. The web search feature enables the generator to access relevant and up-to-date scientific information, ensuring the assessments remain accurate and aligned with the latest developments. The DALL-E image generation feature provides customized diagrams or visuals to accompany certain assessment questions, adding an interactive and visual dimension, which is especially useful for science subjects requiring illustrations. Additionally, the code interpreter enhance the generator's

functionality by allowing the creation of programming-based assessment items, particularly in STEM subjects.

## 5.2. Prompt Engineering

The customization involves *dynamic* prompt strategy (*proactive user prompts*) based on the GenAI-CRSciA framework to support teachers and users in generating CRSciA items. The dynamic prompt strategy specifically involves two strategies *conversation starters* and *dynamic replies*. The *conversation starters* are designed to initiate interaction by asking key questions at the outset [92]. For example, the model prompts the user with:

*"Welcome! I am your culturally responsive science assessment generator (CRSciA Generator). I am here to help you develop science assessment items that meet the diverse cultural and context-specific needs of your students. Would you like assistance in developing a culturally responsive science assessment for your students that align with the NGSS? Please type 'Yes' or 'No' to proceed.*

Moreover, the *dynamic replies* simplify user interaction with suggested inputs and follow-up questions, such as "Yes" or "No". For example, a "yes" to the conversation ensures that users have the flexibility to tailor the assessments based on specific needs and preferences. An example might be:

*"Great! I can help you create an assessment aligned with NGSS standards. Would you like me to use the SNAP questions from the Stanford NGSS Assessment Project? Please type 'Yes' or 'No'."*

To use the OpenAI-powered API (like GPT-4) within the OpenAI Python library, see the example of how to adapt the dynamic prompt (see Table 1). In the provided Python code, parameters such as *temperature* = 0.7 and *max\_tokens* = 150 are set at default to control the randomness and length of the model's output [89]. This was to ensure consistency with the CRSciA-Generator parameters within the OpenAI platform, preventing any divergent behavior in the model's response generation.

**Table 1.** Python Code for Dynamic Prompt in the CRSciA-Generator.

Component	Code Snippet
Import Libraries	<code>import openai</code>
API Key Setup	<code>openai.api_key = "API Key"</code>
Get Response Function	<pre>def get_openai_response(prompt, model="gpt-4"):     response = openai.Completion.create(         engine=model,         prompt=prompt,         max_tokens=150,         temperature=0.7,         n=1,         stop=None     )</pre>
Return Response	<code>return response.choices [0].text.strip()</code>
Conversation Starter Function	<pre>def conversation_starter():     starter_prompt = (         "Welcome! I am your culturally responsive science assessment         generator (CRSciA Generator). "         "I am here to help you create assessment items that meet the diverse         cultural and context-specific needs of your class "         "that align with the NGSS. Let's begin with a few questions to tailor         the assessment for your class.\n"         "What science topic or NGSS standard would you like to cover?"     )</pre>
User Topic Input	<code>user_topic = input(get_openai_response(starter_prompt) + "\n")</code>

Return User Topic	<code>return user_topic</code>
User-Prompted Pathway Function	<code>def user_prompted_pathway():     language_prompt = "What are the dominant languages your students can read and write in for science?"     cultural_prompt = "Would you like to include any culturally specific knowledge or context in the assessment? (Yes/No)"</code>
Get Responses from User	<code>language = input(get_openai_response(language_prompt) + "\n") cultural_relevance = input(get_openai_response(cultural_prompt) + "\n")</code>
Cultural Context Check	<code>if cultural_relevance.lower() == 'yes':     context_prompt = "Please provide some details about the cultural context you'd like to include."     context = input(get_openai_response(context_prompt) + "\n")     return language, context</code>
Return Language and Context	<code>else:     return language, None</code>
Main Function to Generate Assessment	<code>def generate_assessment():     topic = conversation_starter()     language, context = user_prompted_pathway()</code>
Display Summary	<code>print("\n--- Assessment Summary ---") print(f"Science Topic: {topic}") print(f"Language: {language}") if context:     print(f"Cultural Context: {context}") else:     print("No specific cultural context included.")</code>
Run the Generator	<code>print("\nYour assessment will be tailored based on the information provided.") generate_assessment()</code>

### 5.3. Piloting the CRSciA-Generator

In the application to determine the efficacy of the CRSciA-Generator, a pilot test was taken to compare the base GPT 4o model with the standard prompt and the CRSciA-Generator with the embedded dynamic prompt. 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 prompt offers a simple and direct approach for instructing ChatGPT by specifying a particular task for the model to execute, often using a format like "Generate a [task]" [93]. Given that SNAP items require students to complete specific tasks, standard prompts are an ideal choice for comparison in this context. In contrast, dynamic prompts provide a more flexible and adaptive approach, adjusting based on user inputs and the specific context of the task. This responsiveness makes dynamic prompts particularly useful for generating content that aligns with the user's needs. In the CRSciA-Generator, for example, dynamic prompts start with a conversation starter, facilitating a deeper and more interactive engagement with the user [94,95].

### 5.4. Use Cases of the CRSciA Generator and Prompts

To apply the GenAI-CRSciA framework in practical examples, we adapted the NGSS Life Science questions to prompt the base ChatGPT 4o (See Table 2) using standard prompt (See Table 3) and dynamic prompt within the CRSciA-Generator (see Table 4). Find below the original and generated questions for both prompts in relation to three students from Ghana, the USA, and China.

**Table 2.** Original Questions.

<p>NGSS standard for MS-LS2-2 [96]NGSS standard for MS-LS2-2 [96]NGSS standard for MS-LS2-2 [96]NGSS standard for MS-LS2-2 [96]</p> <p>MS-LS2-2_024_ Wolves and moose interactions</p> <p>Ecosystem Interactions</p> <p>Construct an explanation that predicts patterns of interactions among organisms across multiple ecosystems.</p> <p>Activity: Wolves and moose interactions (ID# 116-04-Y03)</p> <p>Construct an evidence-based argument about whether there exists a predator and prey relationship between wolves and moose.</p> <p>From: Page 1 Wolves and moose interactions (ID# 116-04-Y03) (concord.org)</p> <p>Source: This resource is part of the Concord Consortium's Middle Grades (6-8) project</p> <p><i>Task</i></p> <p>Wolves and moose interactions (ID# 116-04-Y03)</p> <div style="display: flex; justify-content: space-around;"> <div style="text-align: center;">  <p><b>Wolves</b></p> </div> <div style="text-align: center;">  <p><b>Moose</b></p> </div> </div> <p><i>Wolf and moose images are licensed under the public domain.</i></p>
---

**Table 3.** Standard Prompt with base GPT 4o.

<p><i>Standard Prompt:</i></p> <p><i>Prompt:</i></p> <p>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.</p> <p><i>Output</i></p> <p><i>Culturally Responsive Science Assessment Question:</i></p> <p>NGSS Standard: MS-LS2-2</p> <p>Concept: Ecosystem Interactions</p> <p>Grade Level: Middle School</p> <p>Target Regions: Ghana, the USA, China</p> <p><i>Question:</i></p> <p>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:</p> <ol style="list-style-type: none"> <li>1. Identify the predator and the prey in this relationship.</li> <li>2. Describe the ecosystem where these interactions occur.</li> <li>3. Provide evidence from your knowledge or experiences that shows how the predator and prey populations affect each other.</li> <li>4. Explain how this relationship might be similar or different from other predator-prey relationships around the world.</li> </ol>
---

**Table 4.** CRSciA -Generator Conversation Starter and Dynamic Prompt.

<p><i>Prompt: Following CRSciA -Generator Conversation Starter and Dynamic Prompt Strategies</i></p> <p>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].</p> <p>Instructions: Use the following guidelines to structure the assessment question:</p> <ol style="list-style-type: none"> <li>1. Language: <ul style="list-style-type: none"> <li>○ Use [insert language, e.g., English, Mandarin, Twi] as the primary language for the question.</li> <li>○ Ensure the language used is appropriate for the educational level and cultural context of the students.</li> </ul> </li> <li>2. Indigenous Knowledge: <ul style="list-style-type: none"> <li>○ Incorporate local or indigenous knowledge relevant to the students' cultural backgrounds.</li> <li>○ Use examples that resonate with the student's experiences, such as [insert specific cultural examples related to the scientific concept, e.g., traditional ecological practices, folklore].</li> </ul> </li> <li>3. Race/Ethnicity: <ul style="list-style-type: none"> <li>○ Respect and reflect the ethnic and cultural identities of the students.</li> <li>○ Include scenarios that are geographically and culturally relevant to the student's environment.</li> </ul> </li> <li>4. Family: <ul style="list-style-type: none"> <li>○ Encourage the student to engage with their family or community members to gather insights or stories related to the scientific concept being assessed.</li> <li>○ Ask the student to incorporate these familial insights into their response.</li> </ul> </li> <li>5. Religious Beliefs: <ul style="list-style-type: none"> <li>○ Allow for the exploration of religious or spiritual beliefs that might influence the student's understanding of the scientific concept.</li> <li>○ Ensure the question does not conflict with or marginalize any religious beliefs.</li> </ul> </li> </ol> <ul style="list-style-type: none"> <li>• Standards</li> </ul> <p>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.</p> <p style="text-align: center;"><i>CRSciA-Generator Questions</i></p> <p>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:</p> <ol style="list-style-type: none"> <li>1. 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.</li> <li>2. Your argument should include: <ul style="list-style-type: none"> <li>○ 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.</li> <li>○ A statement about what is accurate or inaccurate in Jason's evidence and reasoning.</li> </ul> </li> </ol>
--

- 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:

1. 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.
2. 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

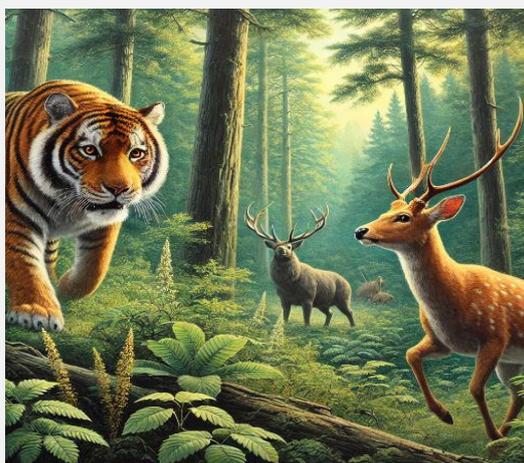
Context: Consider how traditional Chinese views on predators like the Amur tiger can help you understand predator-prey relationships in ecosystems. Reflect on how your family's stories or the teachings from Taoism or Buddhism might influence your interpretation of the graph.

## 任务:

1. 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。
2. 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:

1. 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.
2. 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.)



In evaluating the outputs, the standard prompt was found to generate suggestive approaches to developing questions that could meet the GenAI-CRSciA framework, particularly suggesting teachers or users to “Think about a predator and prey relationship that is familiar in your cultural context or from your region (Ghana, the USA, or China)”. This makes it limited in its ability to fully embrace the individual cultural nuances, which constrained the specificity in the output.

In contrast, the output from the CRSciA-Generator comprehensively addresses the GenAI-CRSciA framework by generating different sets of questions for each cultural group. For example, regarding language, the CRSciA-Generator translates certain keywords into the specific languages of each country (e.g., Akan for Ghana, Mandarin for China), ensuring linguistic accessibility and cultural connection for students. It further generates relevant examples and visuals that aligns with local species and ecological relationships, enhancing student engagement by resonating with their racial and ethnic contexts. Moreover, the CRSciA-Generator acknowledges the religious components within the framework, providing culturally specific examples, such as an antelope, a wolf, or an amur, that address beliefs or stories unique to each country.

Nevertheless, one observation about the CRSciA-Generator output is that it generalized the content for each of the countries, rather than specifying distinct regional or subcultural contexts within each country, which could potentially introduce bias by overlooking the rich diversity within subcultures. This limitation is likely related to the amount of background information provided, as the input is more of country-specific levels.

## 6. Discussion

The alignment of the CRSciA- Generator 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 [97,98]. 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 [99]. 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 [100]. 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 CRSciA-Generator 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 [101,102].

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, such as 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 [103]. 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 [104]. 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 [105].

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 [106]. Moreover, 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 [107,108]. This comprehensive approach ensures that the assessments are scientifically and deeply rooted in the cultural realities of the students. The findings demonstrate GenAI's potential to revolutionize 21st-century assessments which aligns with the conclusions of Owoseni et al. [109] who highlight GenAI's capability to enhance summative assessments by automating tasks like question generation and grading.

Again, the findings of this study suggest that customization and dynamic prompt engineering will be the future of AI literacy. This evolution highlights the need for further research and the integration of prompt engineering into education. This suggestion align with those of Knoth et al. [110], as well as Arvidsson and Axell [111], who emphasize the crucial role of prompt engineering in optimizing GenAI across different fields, including education and requirements engineering. Both works highlight that the quality of GenAI outputs is directly influenced by the precision of prompt engineering, underscoring the importance of GenAI literacy Arvidsson and Axell [111] further point out how domain-specific guidelines can improve the accuracy and utility of GenAI for specialized tasks.

## 7. Conclusions and Future Directions

This study started with the development of a GenAI-CRSciA framework which incorporated core cultural tenets such as indigenous/dominant languages, race/ethnicity, indigenous knowledge, religious beliefs, and family and community, which are essential for creating CRSciA and aligned with the capabilities of GenAI. This framework guided the customization of a CRSciA-Generator within the OpenAI platform. The CRSciA-Generator is equipped with dynamic prompts based on the GenAI-CRSciA-framework to automatically generate assessments following the conversational information from users. A pilot test compared the CRSciA-Generator with the GPT-4o base model using standard prompts. The results reveal that while the GPT-4o base model successfully generates science assessments aligned with NGSS standards, its approach remains broad and generalized, lacking region-specific cultural relevance.

In contrast, the CRSciA-Generator produced assessment items incorporating local knowledge, traditional stories, and cultural perspectives, making it more culturally responsive to students in regions such as Ghana, the USA, and China. The assessment specifically features examples such as traditional stories of lions and antelopes in Ghana, Native American views on wolves in the USA, and Taoist or Buddhist teachings on the Amur tiger in China. This culturally tailored assessment has the potential to foster deeper student engagement by connecting science concepts to students' heritage and worldviews, indicating that the CRSciA-Generator has the potential to generate equitable and inclusive CRSciAs for diverse students from different cultures and context experiences.

However, the pilot test was overgeneralized by focusing on broad national contexts, treating entire countries as culturally homogenous. Therefore, we recommend that teachers provide sufficient cultural and context background information about their students to enable the CRSciA-Generator produces more contextually accurate assessments. Additionally, we believe the pilot test does not fully validate the model's efficacy. Future studies involving human experts are necessary to review and verify the cultural and contextual specificity of the generated assessments. We recommend conducting empirical studies in diverse contexts to further use and validate the model's overall effectiveness.

**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.

## Appendix A

### Example of the Dynamic Prompt Strategy within the CRSciA-Generator

Type anything or click this starter to begin ..



Welcome! I am your culturally responsive science assessment generator (CRSciA Generator). I am here to help you develop science assessment items that meet the diverse cultural and context-specific needs of your students, aligned with the NGSS. Let us begin with a few questions to tailor the assessment for your class.

Would you like assistance in developing a culturally responsive science assessment for your students? Please type 'Yes' or 'No' to proceed.

Yes



Great! I can help you create an assessment aligned with NGSS standards. Would you like me to use the SNAP questions from the Stanford NGSS Assessment Project? Please type 'Yes' or 'No'.

Ys



Thank you! Let's start by customizing your assessment.

1. What is the main science topic or objective you would like the assessment to cover?
2. Are there specific NGSS Performance Expectations (PEs) you would like to use for this assessment?



Focus on photosynthesis and plant growth.

Sound waves in different environments assessment.



Message Culturally Responsive Science Assessment Generator



## Appendix B

### Python Code Snip of Dynamic Prompt

```
In [2]: M import openai

# Set up your API key directly in the script (remember to rotate it after use)
openai.api_key = "API Key"

# Function to get responses from OpenAI
def get_openai_response(prompt, model="gpt-4"):
    response = openai.Completion.create(
        engine=model,
        prompt=prompt,
        max_tokens=150,
        temperature=0.7,
        n=1,
        stop=None
    )
    return response.choices[0].text.strip()

# Define the conversation starter function
def conversation_starter():
    starter_prompt = (
        "Welcome! I am your culturally responsive science assessment generator (CRSciA Generator). "
        "I am here to help you create assessment items that meet the diverse cultural and context-specific needs of your class "
        "that align with the NGSS. Let's begin with a few questions to tailor the assessment for your class.\n"
        "What science topic or NGSS standard would you like to cover?"
    )

# Get the user's topic
user_topic = input(get_openai_response(starter_prompt) + "\n")
return user_topic

# Define the user-prompted pathway function
def user_prompted_pathway():
    language_prompt = "What are the dominant languages your students can read and write in for science?"
    cultural_prompt = "Would you like to include any culturally specific knowledge or context in the assessment? (Yes/No)"

# Get responses from the user
language = input(get_openai_response(language_prompt) + "\n")
cultural_relevance = input(get_openai_response(cultural_prompt) + "\n")

if cultural_relevance.lower() == 'yes':
    context_prompt = "Please provide some details about the cultural context you'd like to include."
    context = input(get_openai_response(context_prompt) + "\n")
    return language, context
else:
    return language, None

# Main function to run the CRSciA conversation
def generate_assessment():
    # Start the conversation
    topic = conversation_starter()

# Get user preferences for customization
language, context = user_prompted_pathway()

# Display the customized assessment parameters
print("\n--- Assessment Summary ---")
print(f"Science Topic: {topic}")
print(f"Language: {language}")
if context:
    print(f"Cultural Context: {context}")
else:
    print("No specific cultural context included.")

print("\nYour assessment will be tailored based on the information provided.")

# Run the generator conversation
generate_assessment()
```

## References

1. Frommert, C. *Creating an environment where all students see that they can be good at math*. 2023 Last Update Date: January 27 [cited 2024 August 5]; Available from: <https://www.edutopia.org/article/helping-students-see-themselves-good-math>.
2. Kouo, J.L., Don't Assess a Fish by Its Ability to Climb a Tree: Considerations and Strategies to Ensure Equitable Formative Assessment Practices for All Learners, in *Research Anthology on Physical and Intellectual Disabilities in an Inclusive Society*. 2022, IGI Global. p. 873-889.
3. NRC, A framework for K-12 science education: Practices, crosscutting concepts, and core ideas, ed. N.R. Council, et al. 2012: National Academies Press.
4. States, N.L., Next generation science standards: For states, by states. 2013: National Academies Press.
5. Mullis, I.V., M.O. Martin, and M. von Davier, *TIMSS 2023 Assessment Frameworks*. International Association for the Evaluation of Educational Achievement, 2021.

6. States, N.N.L., Next generation science standards: For states, by states. 2013: National Academies Press.
7. Harris, K., et al., *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, 2018. **49**(1): p. 192-208.
8. Câmara, J.N., Funds of knowledge: Towards an asset-based approach to refugee education and family engagement in England. *British Educational Research Journal*, 2024. **50**(2): p. 876-904.
9. O'Leary, E.S., et al., Creating inclusive classrooms by engaging STEM faculty in culturally responsive teaching workshops. *International Journal of Stem Education*, 2020. **7**(1).
10. Idrus, F. and M. Sohid, Teachers' expectations and challenges in using culturally responsive teaching (CRT) strategies in the ESL classroom. *Journal of Language Teaching and Research*, 2023. **14**(3): p. 629-635.
11. Nortvedt, G.A., et al., *Aiding culturally responsive assessment in schools in a globalising world*. Educational Assessment Evaluation and Accountability, 2020. **32**(1): p. 5-27.
12. Preston, J.P. and T.R. Claypool. Analyzing assessment practices for Indigenous students. in *Frontiers in Education*. 2021. Frontiers Media SA.
13. Wu, X., et al. Matching Exemplar as Next Sentence Prediction (MeNSP): Zero-Shot Prompt Learning for Automatic Scoring in Science Education. 2023. Cham: Springer Nature Switzerland.
14. Lee, G.-G., et al., Multimodality of ai for education: Towards artificial general intelligence. arXiv preprint arXiv:2312.06037, 2023.
15. Li, X., B. Li, and S.-J. Cho, Empowering Chinese language learners from low-income families to improve their Chinese writing with ChatGPT's assistance afterschool. *Languages*, 2023. **8**(4): p. 238.
16. 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.
17. Zhai, X., *ChatGPT for next generation science learning*. XRDS: Crossroads, The ACM Magazine for Students, 2023. **29**(3): p. 42-46.
18. Tan, B., et al., A Review of Automatic Item Generation Techniques Leveraging Large Language Models.
19. Hwang, K., et al. Towards AI-assisted multiple choice question generation and quality evaluation at scale: Aligning with Bloom's Taxonomy. in *Workshop on Generative AI for Education*. 2023.
20. Richards, M., et al., Bob or Bot: Exploring ChatGPT's Answers to University Computer Science Assessment. *ACM Trans. Comput. Educ.*, 2024. **24**(1): p. Article 5.
21. Andreadis, K., et al., *Mixed methods assessment of the influence of demographics on medical advice of ChatGPT*. *Journal of the American Medical Informatics Association*, 2024: p. ocae086.
22. Kiyak, Y.S. and E. Emekli, ChatGPT prompts for generating multiple-choice questions in medical education and evidence on their validity: a literature review. *Postgraduate Medical Journal*, 2024.
23. Chan, W., A. An, and H. Davoudi. A Case Study on ChatGPT Question Generation. in *2023 IEEE International Conference on Big Data (BigData)*. 2023.
24. Sembok, T.M.T. and S. Wani. *Is ChatGPT not Appropriate for Religious Use?* 2024. Singapore: Springer Nature Singapore.
25. Nyaaba, M., A.L. Wright, and G.L. Choi, Generative AI and Digital Neocolonialism in Global Education: Towards an Equitable Framework. arXiv preprint arXiv:2406.02966, 2024.
26. Jaakkola, E., *Designing conceptual articles: four approaches*. *AMS Review*, 2020. **10**(1): p. 18-26.
27. McKenney, S.R., Thomas C., *Conducting educational design research*. 2012, London: Routledge.
28. Ladson-Billings, G., From the achievement gap to the education debt: Understanding achievement in US schools. *Educational researcher*, 2006. **35**(7): p. 3-12.
29. Yerrick, R. and M. Ridgeway, CULTURALLY RESPONSIVE PEDAGOGY, SCIENCE LITERACY, AND URBAN UNDERREPRESENTED SCIENCE STUDENTS, in *Inclusive Principles and Practices in Literacy Education*, M. Milton, Editor. 2017. p. 87-103.
30. Morgan, P.L., et al., \* Racial and Ethnic Disparities in Advanced Science and Mathematics Achievement During Elementary School. *Gifted Child Quarterly*, 2023. **67**(2): p. 151-172.
31. Muhammad, G.E., N.A. Ortiz, and M.L. Neville, *A Historically Responsive Literacy Model for Reading and Mathematics*. *Reading Teacher*, 2021. **75**(1): p. 73-81.
32. Hoosain, M.S., B.S. Paul, and S. Ramakrishna, The impact of 4IR digital technologies and circular thinking on the United Nations sustainable development goals. *Sustainability*, 2020. **12**(23): p. 10143.
33. Johnston-Goodstar, K. and R.V. Roholt, "Our kids aren't dropping out; they're being pushed out": Native American students and racial microaggressions in schools, in *Microaggressions and Social Work Research, Practice and Education*. 2020, Routledge. p. 31-48.
34. Apple, M.W., *Can education change society?* 2012: Routledge.
35. Baker, D., *The schooled society: The educational transformation of global culture*. 2020: Stanford University Press.
36. Fink, D. and L. Stoll, Educational change: Easier said than done, in *International Handbook of Educational Change: Part One*. 1998, Springer. p. 297-321.

37. Benediktsson, A.I. and H. Ragnarsdóttir, *Immigrant students' experiences of assessment methods used in Icelandic universities*. Multicultural Education Review, 2020. **12**(2): p. 98-116.
38. Maaruf, S.Z., et al. Examining Knowledge, Awareness and Action Orientation Through Multicultural Assessment Proficiency (MAP) for Cultural Responsive Pedagogy in Higher Education Practices. in 2nd International Research Conference on Business and Economics (IRCBE). 2016. Semarang, INDONESIA.
39. Ladson-Billings, G. Three decades of culturally relevant, responsive, & sustaining pedagogy: What lies ahead? in *The Educational Forum*. 2021. Taylor & Francis.
40. Ghattas, N.I. and J.S. Carver, *Cultural responsiveness of the next generation science standards*. Journal of STEM Teacher Education, 2017. **52**(1): p. 5.
41. Brown, M., et al., Culturally responsive classroom-based assessment A case study of secondary schools in Ireland. *Rie-Revista De Investigacion Educativa*, 2022. **40**(1): p. 15-32.
42. Edwards, F. and R. Edwards, A story of culture and teaching: the complexity of teacher identity formation. *The Curriculum Journal*, 2017. **28**(2): p. 190-211.
43. Rosheim, K.M., K.G. Tamte, and M.J. Froemming, Reducing Inequalities Inherent in Literacy Assessment of Multilingual Learners. *Reading Psychology*, 2024.
44. Nayir, F., et al., *Assessment with and for Migration Background Students-Cases from Europe*. Eurasian Journal of Educational Research, 2019(79): p. 39-68.
45. Muñoz, J., Culturally Responsive Teaching: A 50-State Survey of Teaching Standards. New America, 2019.
46. Stepp, Z.A. and J.C. Brown, The (lack of) relationship between secondary science teachers' self-efficacy for culturally responsive instruction and their observed practices. *International Journal of Science Education*, 2021. **43**(9): p. 1504-1523.
47. IONESCU, V.M. and M.C. ENESCU. Using ChatGPT for Generating and Evaluating Online Tests. in 2023 15th International Conference on Electronics, Computers and Artificial Intelligence (ECAI). 2023. IEEE.
48. Yi, G., et al. Exploring the Capability of ChatGPT in Test Generation. in 2023 IEEE 23rd International Conference on Software Quality, Reliability, and Security Companion (QRS-C). 2023.
49. Pandey, M.K. and J. Upadhyay, Overview of ChatGPT Model Architecture, in *Advanced Applications of Generative AI and Natural Language Processing Models*. 2024, IGI Global. p. 87-107.
50. Lee, U., et al., Few-shot is enough: exploring ChatGPT prompt engineering method for automatic question generation in english education. *Education and Information Technologies*, 2023: p. 1-33.
51. Alasadi, E.A. and C.R. Baiz, *Multimodal Generative Artificial Intelligence Tackles Visual Problems in Chemistry*. Journal of Chemical Education, 2024. **101**(7): p. 2716-2729.
52. Gupta, N., et al., *Fundamentals Of Chat GPT For Beginners Using AI*. 2024: Academic Guru Publishing House.
53. Bevilacqua, M., et al., When Automated Assessment Meets Automated Content Generation: Examining Text Quality in the Era of GPTs. arXiv preprint arXiv:2309.14488, 2023.
54. Kartal, G., *Contemporary language teaching and learning with ChatGPT*. Contemporary Research in Language and Linguistics (ISSN: 2980-2253), 2023. **1**(1).
55. Arifianto, M.L., T. Thoriquttyas, and I.F. Izzudin, *Inquiring Arabic-Islamic Debatable Terms to ChatGPT: Will It Take a Side?* Digital Muslim Review, 2023. **1**(1): p. 18-35.
56. Shabbir, A., et al., Beyond boundaries: Navigating the positive potential of ChatGPT, empowering education in underdeveloped corners of the world. *Heliyon*, 2024. **10**(16).
57. Zheng, Y.D. and N. Stewart, *Improving EFL students' cultural awareness: Reframing moral dilemmatic stories with ChatGPT*. Computers and Education: Artificial Intelligence, 2024. **6**: p. 100223.
58. Oxenberg, R., A Dialogue on the Existence and Nature of God with ChatGPT.
59. Jayaweera, M., From Jargon to Clarity: Enhancing Science Communication with ChatGPT. *Vidyodaya Journal of Science*, 2023. **26**(01).
60. Elu, J., *Gender and science education in sub-Saharan Africa*. Journal of African Development, 2018. **20**(2): p. 105-110.
61. Wright, T.S. and L.M. Domke, *The role of language and literacy in K-5 science and social studies standards*. Journal of Literacy Research, 2019. **51**(1): p. 5-29.
62. Kūkea Shultz, P. and K. Englert. Cultural validity as foundational to assessment development: An indigenous example. in *Frontiers in Education*. 2021. Frontiers Media SA.
63. Latif, E., et al., G-SciEdBERT: A Contextualized LLM for Science Assessment Tasks in German. arXiv preprint arXiv:2402.06584, 2024.
64. Mantelas, N. and E. Mavrikaki, *Religiosity and students' acceptance of evolution*. International Journal of Science Education, 2020. **42**(18): p. 3071-3092.
65. 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*, 2017. **9**(2): p. 206-222.
66. Barnes, M.E., et al., Relationships between the religious backgrounds and evolution acceptance of Black and Hispanic biology students. *CBE—Life Sciences Education*, 2020. **19**(4): p. ar59.
67. Owens, D.C., et al., Scientific and religious perspectives on evolution in the curriculum: An approach based on pedagogy of difference. *Research in Science Education*, 2018. **48**(6): p. 1171-1186.

68. Sumarni, W., et al., *The Urgency of Religious and Cultural Science in STEM Education: A Meta Data Analysis*. International Journal of Evaluation and Research in Education, 2020. **9**(4): p. 1045-1054.
69. Cheong, P.H., Robots, religion and communication: Rethinking piety, practices and pedagogy in the era of artificial intelligence, in *Religion in the Age of Digitalization*. 2020, Routledge. p. 86-96.
70. Ashraf, C., Exploring the impacts of artificial intelligence on freedom of religion or belief online. *The International Journal of Human Rights*, 2022. **26**(5): p. 757-791.
71. Black, A. and J.M. Tylianakis, *Teach Indigenous knowledge alongside science*. Science, 2024. **383**(6683): p. 592-594.
72. Trumbull, E. and S. Nelson-Barber, The Ongoing Quest for Culturally-Responsive Assessment for Indigenous Students in the US. *Frontiers in Education*, 2019. **4**.
73. Azam, S. and K. Goodnough, *Learning together about culturally relevant science teacher education: Indigenizing a science methods course*. International Journal of Innovation in Science and Mathematics Education, 2018. **26**(2).
74. Jin, Q., Supporting indigenous students in science and STEM education: A systematic review. *Education Sciences*, 2021. **11**(9): p. 555.
75. Washburn, J. and J.M. McCutchen, AI Meets AI: ChatGPT as a Pedagogical Tool to Teach American Indian History. *Critical Humanities*, 2024. **2**(2): p. 2.
76. Castro Nascimento, C.M. and A.S. Pimentel, *Do large language models understand chemistry? a conversation with chatgpt*. Journal of Chemical Information and Modeling, 2023. **63**(6): p. 1649-1655.
77. Atwater, M.M., et al., Race and ethnicity: Powerful cultural forecasters of science learning and performance. *Theory Into Practice*, 2013. **52**(1): p. 6-13.
78. 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*, 2020. **52**(2): p. 56-61.
79. Riegle-Crumb, C., et al., Inquiry-based instruction in science and mathematics in middle school classrooms: Examining its association with students' attitudes by gender and race/ethnicity. *AERA open*, 2019. **5**(3): p. 2332858419867653.
80. Lesseig, K., et al., *An analysis of cultural influences on STEM schools: Similarities and differences across K-12 contexts*. International Journal of Science and Mathematics Education, 2019. **17**: p. 449-466.
81. Choudhary, T., Reducing Racial and Ethnic Bias in AI Models: A Comparative Analysis of ChatGPT and Google Bard. 2024.
82. Warr, M., et al., *Is ChatGPT Racially Biased? The Case of Evaluating Student Writing*. The Case of Evaluating Student Writing (May 25, 2024), 2024.
83. Denton, M., M. Borrego, and A. Boklage, Community cultural wealth in science, technology, engineering, and mathematics education: A systematic review. *Journal of Engineering Education*, 2020. **109**(3): p. 556-580.
84. Gerde, H.K., et al., *Head Start children's science experiences in the home and community*. Early Childhood Research Quarterly, 2021. **54**: p. 179-193.
85. Soto-Lara, S. and S.D. Simpkins, Parent support of Mexican-descent high school adolescents' science education: A culturally grounded framework. *Journal of Adolescent Research*, 2022. **37**(4): p. 541-570.
86. Garbacz, S.A., et al., Family engagement in education and intervention: Implementation and evaluation to maximize family, school, and student outcomes. *Journal of school psychology*, 2017. **62**: p. 1-10.
87. Said, Z. and A. Al-amadi, *The Role of Family Engagement in Students' Science Learning in Qatari Schools*. The Eurasia Proceedings of Educational and Social Sciences, 2018. **11**: p. 142-152.
88. Shahin, M., F.F. Chen, and A. Hosseinzadeh, *Harnessing customized AI to create voice of customer via GPT3.5*. Advanced Engineering Informatics, 2024. **61**: p. 102462.
89. Gallifant, J., et al., *Peer review of GPT-4 technical report and systems card*. PLOS Digital Health, 2024. **3**(1): p. e0000417.
90. Bast, H. and C. Korzen. A benchmark and evaluation for text extraction from PDF. in 2017 ACM/IEEE joint conference on digital libraries (JCDL). 2017. IEEE.
91. Sİregar, A.Z., *Implementation Of ADFGVX Cipher And RSA On TXT And Doc File Security*. Journal Basic Science and Technology, 2021. **10**(3): p. 105-110.
92. Janssens, R., P.W.T. Demeester, and T. Belpaeme, *Integrating Visual Context into Language Models for Situated Social Conversation Starters*. IEEE Transactions on Affective Computing, 2024.
93. Spasić, A.J. and D.S. Janković. Using ChatGPT standard prompt engineering techniques in lesson preparation: role, instructions and seed-word prompts. in 2023 58th International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST). 2023. IEEE.
94. Yang, F., et al., *Dynamic prompt learning: Addressing cross-attention leakage for text-based image editing*. Advances in Neural Information Processing Systems, 2023. **36**: p. 26291-26303.
95. Mo, W., et al. Dynamic Prompt Optimizing for Text-to-Image Generation. in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024.

96. Consortium, C. *Wolves and moose interactions (ID# 116-04-Y03)*. NGSS standard for MS-LS2-2 2024; Available from: <https://scienceeducation.stanford.edu/assessments>.
97. Erling, E.J., et al., 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*, 2018: p. 18-34.
98. Loh, E.K., L.C. Tam, and K.-c. Lau, Moving between language frontiers: The challenges of the medium of instruction policy for Chinese as a second language. *Language Policy*, 2019. **18**(1): p. 131-153.
99. Zhang, Z., *English-medium instruction policies in China: Internationalisation of higher education*. *Journal of Multilingual and Multicultural Development*, 2018. **39**(6): p. 542-555.
100. Yu, S., et al., Coping with EMI (English as a medium of instruction): Mainland China students' strategies at a university in Macau. *Innovations in Education and Teaching International*, 2021. **58**(4): p. 462-472.
101. Owu-Ewie, C. and E.S. Eshun, 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*, 2015. **6**(3): p. 72-82.
102. Owu-Ewie, C. and E.S. Eshun, 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*, 2019. **20**(4): p. 365-388.
103. Akyeampong, O.A., *Aboakyer: traditional festival in decline*. *Ghana Social Science Journal*, 2019. **16**(1): p. 97.
104. Lumor, F., Significance of animal symbolism among the Akans of Akyem Abuakwa traditional area. 2009, Kwame Nkrumah University of Science and Technology.
105. 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*, 2022. **17**(10): p. e0276554.
106. Rudenko, S. and Y. Sobolievskiy, *Philosophical ideas in spiritual culture of the indigenous peoples of North America*. *Anthropological Measurements of Philosophical Research*, 2020(18): p. 168-182.
107. Kosoe, E.A., P.O.-W. Adjei, and F. Diawuo, From sacrilege to sustainability: the role of indigenous knowledge systems in biodiversity conservation in the Upper West Region of Ghana. *GeoJournal*, 2020. **85**: p. 1057-1074.
108. Tavor, O., *In the Land of Tigers and Snakes: Living with Animals in Medieval Chinese Religions*, By Huaiyu Chen. 2024, Oxford University Press US.
109. Owoseni, A., O. Kolade, and A. Egbetokun, Applications of Generative AI in Summative Assessment, in *Generative AI in Higher Education: Innovation Strategies for Teaching and Learning*. 2024, Springer Nature Switzerland: Cham. p. 97-122.
110. Knoth, N., et al., *AI literacy and its implications for prompt engineering strategies*. *Computers and Education: Artificial Intelligence*, 2024. **6**: p. 100225.
111. Arvidsson, S. and J. Axell, *Prompt engineering guidelines for LLMs in Requirements Engineering*. 2023.

**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.