

Using Generative AI To Develop Maths Tasks With Focus On Social Justice Perspectives

Sabahat Awan

Australian Catholic University
sabahat.awan@myacu.edu.au

Chrissy Monteleone

La Trobe University
c.monteleone@latrobe.edu.au

Tracey Muir

Australian Catholic University
Tracey.Muir@acu.edu.au

Kimberley Wilson

University of the Sunshine Coast
kwilson7@usc.edu.au

This paper reports on the use of ChatGPT to develop Social Justice Mathematics (SJM) tasks for year 5 and 6 students. Two educators engaged in an iterative prompting process, drawing on a task design framework. Findings indicate that while ChatGPT can produce tasks situated in authentic, equity-oriented contexts, outputs frequently described mathematical activity rather than operationalising it. This requires teachers to apply their pedagogical content knowledge to evaluate and re-direct AI generated content. The SJM task framework proved useful both as a prompt structure and as an evaluative lens for teachers to use for designing authentic social justice-based tasks.

Introduction

Generative Artificial Intelligence (AI) models, such as ChatGPT, can be powerful tools for enhancing teachers' professional learning and pedagogical practices (Cameron & Mesiti, 2024). These technologies present both affordances and challenges for primary school teachers in mathematics education. AI tools have the potential to support teachers in creating curricular materials and designing meaningful learning activities that encourage critical thinking and effective collaboration among students (Baidoo-Anu & Ansah, 2023). However, a key challenge for teachers is knowing how to craft questions or prompts that align with the broader pedagogical intentions of a lesson. At present, the use of AI to generate prompts for developing rich tasks that emphasise the importance of critical and informed thinking when addressing social issues (Skilling & Park, 2025) is limited in educational research and practice.

A large body of literature in mathematics education calls for broadening the school mathematics curriculum to include instructional approaches that engage students with contemporary social issues (e.g. Frankenstein, 2009; Gutstein, 2016; Skovsmose, 2023). More specifically, it is argued that students should be offered opportunities to critically examine complex real-world problems (Steflitsch & Kolloosche, 2025) and develop their identities as reflective and responsible citizens (English, 2025).

This paper explores how generative AI tools can be used to adapt maths tasks to engage with personal and societal issues using a Social Justice Mathematics (SJM) framework informed by the literature (e.g., Gutstein, 2016; Skovsmose, 2023). We report on an iterative prompting process used to refine ChatGPT's output for the same mathematics task across two trials. The following research questions are addressed in this paper:

What prompts are useful for developing SJM tasks using ChatGPT?

To what extent did the tasks generated by ChatGPT align with SJM task principles?

Background

This paper will consider two core themes including Generative AI and task design, and social justice mathematics tasks. To help frame the understanding, each theme will be explained in the following sections.

AI and Task Design

Generative AI tools, particularly large language models such as ChatGPT, are increasingly being explored as supports for teachers in the design of instructional materials and learning tasks. These models are trained on large corpora of publicly available and licensed texts and generate responses by identifying patterns and associations across this training data, producing structured text in response to user prompts (Cameron & Mesiti, 2024). In educational contexts, AI can function as a support for primary school teachers by acting as a ‘guide on the side’ or as a co-designer of curriculum materials, including lesson plans and learning sequences. However, while AI can generate coherent and structured outputs, the decision-making processes underpinning these outputs remain largely opaque to users (Daugherty & Carter, 2025), and the pedagogical expertise required to evaluate and refine AI-generated teaching and learning materials continues to reside with teachers. Effective use therefore requires new forms of professional knowledge, including the capacity to craft prompts, interpret outputs critically, and iteratively refine generated content so that it aligns with pedagogical intent (Cameron & Mesiti, 2024).

Within this emerging area, attention has increasingly turned to the role of prompts in shaping the nature and quality of AI-generated outputs. Prompt engineering is commonly described as the practice of designing inputs that guide a language model to generate precise, relevant, and task-appropriate responses (Spasić & Janković, 2023). Practitioner-oriented literature further suggests that prompts are often structured around a defined task, supported by additional elements such as instructions, roles, or key terms that help focus the generated response (Ibrahim, 2023). These techniques highlight how relatively small changes in prompt structure can influence the specificity and orientation of AI generated content.

More systematic investigation of prompt design has begun to appear in research literature. Spasić and Janković (2023), for example, examined the use of standard prompting techniques in the preparation of lesson plans and found that the inclusion of role specification, explicit instructions, and seed words led to more detailed and targeted outputs. Their work highlights that the structure of a prompt can directly shape the form and focus of the materials generated by AI, suggesting that careful prompt construction is an important consideration when using these tools for educational task design.

Features of Social Justice Mathematics (SJM) Tasks

Students’ ability to critically evaluate complex information, make inferences, and consider multiple solutions to real-world problems, requires engaging them with hands-on activities that address authentic challenges (Daugherty & Carter, 2025). Within SJM pedagogy, students learn to use mathematics to justify decisions and examine societal issues (Gutstein, 2006). This approach encourages critical engagement with complex social problems while supporting students to develop agency and capacity for action through mathematical problem-solving (Ernest, 2007). In SJM classrooms, mathematics tasks often draw on cultural, socio-political, and economic contexts (D’Ambrosio & D’Ambrosio, 2013), enabling students to explore issues of justice, such as inequitable access to resources (Ödmo et al., 2023).

One way of implementing SJM pedagogy is by designing mathematics tasks that promote collaborative problem-solving. Through this approach, students can work together to negotiate

alternative strategies to addressing complex problems, justify their solutions, and learn from their classmates' perspectives (Stefflitsch & Kollosche, 2025). Maintaining the cognitive demand of tasks is essential, to ensure that all students engage with rich mathematical concepts and connect solutions with their prior mathematical knowledge (Matthews et al., 2022). Creating contextualised tasks derived from students' life-worlds can further strengthen engagement, motivating students to use mathematics to interpret and question the world around them (Gutstein, 2006). At the same time, tasks framed around broader real-world problems can deepen students' understanding of both mathematical concepts and issues of justice and equity by exposing them to social realities beyond their immediate experience (Frankenstein, 2009).

Despite these benefits, designing such tasks can be challenging. Teachers often face constraints related to time, curriculum demands and the complexity of developing appropriate instructional activities (Daugherty & Carter, 2025). Guidelines or frameworks that can be used as evaluative tools for teachers to develop or select appropriate SJM tasks can help with addressing these barriers. Table 1 provides a framework for designing tasks with a SJM focus, based on elements commonly identified in the literature as being characteristic of SJM tasks (e.g., Frankenstein, 2009; Gutiérrez, 2012; Gutstein, 2006; Skovsmose, 2012).

Table 1

A Framework for Designing Maths Tasks with a Social Justice Focus

Design principles	Associated literature
1. Embedded in social or equity-based issues	(Gutiérrez, 2012; Gutstein, 2006; Skovsmose, 2012)
2. Mathematically rich & cognitively demanding	(Matthews et al., 2022)
3. Promotes critical reflection and action	(Gutstein, 2006; Stinson & Wager, 2012)
4. Encourage multiple perspectives and solutions	(Canogullari & Radmehr, 2025; Sullivan et al., 2012)
5. Student agency and decision making	(Gutstein, 2007; Wright, 2021)
6. Contextualised tasks	(Frankenstein, 2009; Gutstein, 2016)
7. Uses real-world data	(Bragg, 2021; Frankenstein, 2009)

The framework informed the analysis of mathematics tasks generated by ChatGPT in this study and is used in the Results section to evaluate the alignment of the generated tasks with principles of SJM pedagogy.

Methodology

Research Design

This study employed a comparative qualitative case study design (Stake, 1995) to examine how iterative prompting influenced the generation of SJM tasks using ChatGPT. Two educators independently entered an identical initial comprehensive prompt (see Figure 1) requesting a mathematically rich, socially critical task addressing inequitable access to healthcare in Australia. When devising the prompt, we were informed by Spasić and Janković's (2023) findings which advised specifying a role for ChatGPT, along with instructions and seed words. Each participant then engaged in iterative prompting, refining the generated output until the task was judged to be coherent, implementable, and aligned with principles of socially critical mathematics pedagogy. The participants worked independently during this phase.

All prompts and corresponding ChatGPT outputs were retained and treated as textual artefacts. Consistent with principles of document analysis (Bowen, 2009), each prompt–response pair constituted a unit of analysis. Within-case analysis traced shifts across iterations, focusing on linguistic accessibility, mathematical specificity, structural clarity, and classroom readiness. Iterative changes were summarised in tabular form to document refinements and limitations, which are outlined in the Results section of this paper. The final task produced in

each case was then evaluated against established characteristics of SJM tasks. Cross-case comparison is presented in the Discussion section.

Figure 1

Prompts used in ChatGPT

Can you apply a mathematical critical thinking and social justice lens on a critical societal issue like accessing healthcare in different regions of Australia. Generate a maths task by making social justice lens explicit in the task. Ensure that the task is mathematically rich and cognitively demanding, introduces a context from students' familiar settings and engage students in socio-political enquiries that support students' understanding of the relationship between mathematics and social justice issues of individual and global relevance. You could include teaching prompts such as: do you know the number of people living in metropolitan cities and regional/remote areas of Australia? Also, when was the last time you visited a doctor and how long did it take to reach the clinic or hospital? Then add a prompt – do you think that everyone in Australia should have equal access to primary healthcare? The task should be able to be completed within 60-90 minutes.

Results

This section presents the results from each case sequentially. For each case, we provide an overview of the iterative prompting process, including the key prompt shifts, the outputs generated by ChatGPT, and identified limitations. Due to the lengthy nature of the prompts and ChatGPT's responses, this information is summarised in tabular form to highlight changes across iterations and to distinguish refinements made, such as linguistic, structural, and mathematical modifications. A brief analytic commentary follows each table before presenting the final task artefact produced in that case. The final output for each case is then examined in relation to principles of SJM tasks referred to earlier in this paper. A cross-case comparison is presented in the discussion section.

Case 1: Single comprehensive prompt with iterative language refinement

Table 2 provides an overview of the prompts used, the outputs produced, the nature of improvement, and perceived limitations for each iteration in Case 1.

Table 2

Summary of Prompts and Outcomes for Case One

Iteration	Prompt	Output	Nature of Improvement	Limitations
1	Initial comprehensive prompt requesting socially critical, mathematically rich task	Generated full task description with authentic Australian context; explicit equity framing; suggested use of ratios, rates, modelling and representations	N/A	Abstract, teacher-facing language; no concrete dataset supplied; mathematics described rather than operationalised
2	Request to rewrite as student-facing and engaging	Reframed with relatable scenario ("Imagine you wake up feeling unwell..."); simplified language; retained ratio/representation work and fairness prompts	Increased accessibility; clearer audience alignment; maintained mathematical intent	Still lacked specific numerical data; inquiry questions potentially removed focus from modelling
3	Request for teacher rationale linking to mathematical critical thinking	Added explicit pedagogical justification articulating mathematical reasoning and equity focus	Clarified instructional intent; supported teacher implementation	Did not strengthen mathematical structure of task itself; dataset still unspecified

As can be seen from Table 2, the refinements to the task were primarily linguistic and pedagogical rather than structural in nature. The first output demonstrated strong conceptual alignment with socially critical mathematics, but lacked concrete numerical data and was too abstract for students to easily access. The second iteration improved accessibility and audience alignment while still maintaining mathematical relevance. The third iteration added pedagogical explanation but did not substantively modify the task design. While the final artefact integrated mathematics and social justice coherently, movement across iterations were based on clarification and justification, rather than increased mathematical specificity or advice on implementation.

Generated SJM Task One

The initial output indicated that the model successfully incorporated several core design features aligned with socially critical mathematics tasks. Specifically, the task situated learning in an authentic and nationally relevant context (healthcare access in Australia), foregrounded equity and fairness as central investigative questions, positioned mathematics as an analytical tool for interrogating social conditions, embedded opportunities for proportional reasoning, ratios, rates, comparison of distributions, and data representation, and structured the activity as an inquiry rather than a procedural exercise. The second iteration resulted in the task described in Figure 2, which was considered the ‘final’ version of the task, in that the third iteration involved adding a rationale, rather than further task refinement. As Figure 2 shows, while the task required students to analyse population distributions, compare service access, and reflect on fairness, no datasets were provided in the AI response, or guidance around what mathematical modelling might be required. The response did allude to social justice elements of equity-based issues along with the potential to consider multiple perspectives and solutions, but these were implied and not inherent requirements of the task.

Figure 2

The Final Version of SJM Task One

Imagine you wake up feeling unwell. You need to see a doctor. How far would you have to travel? How long would it take you to get there?

In this task, you will use mathematics to investigate whether people in different parts of Australia have fair access to primary healthcare, such as local doctors and clinics.

You will start by thinking about your own experiences.

When was the last time you went to a doctor or clinic?

How long did it take to get there?

Was it close to your home?

Next, you will explore data about where people live in Australia — in cities, regional towns, and remote areas — and where healthcare services are located. You will use maths to compare these regions and look for patterns, differences, and gaps. This might include: comparing numbers of people and clinics; working with distances, travel times, or ratios; using tables, graphs, or maps to show your thinking.

As you work, you will be asked to think about fairness, not just numbers. Should everyone in Australia have the same access to healthcare? What does fair access actually mean?

Can maths help us show when something is unfair?

By the end of the task, you will use mathematical evidence to explain what you think about healthcare access in Australia. You will be using maths not just to solve problems, but to understand a real issue that affects people's lives and to ask important questions about fairness and justice.

Case 2: Iterative development of task

Table 3 provides an overview of the prompts used, the outputs produced, the nature of improvement, and perceived limitations for each iteration in Case 2. Table 3 shows that the refinements to the task in the second case involved improvements in organisational clarity and classroom usability, with mathematical steps made more explicit. The first output generated a structured investigation with explicit data sets that potentially was already mathematically rich. Prompted by ChatGPT's offer of providing a worksheet, the task instructions were made more precise, but were still lengthy, potentially leading to cognitive overload, and hence considered not ‘classroom ready’. Improvements tended to centre on presentation and usability, rather than

conceptual strengthening, as the initial task was already quite strong in this regard. There was no need to request data sets as they were supplied in the response. Movement across the iterations reflected streamlining, rather than structural enhancement to the task. The third iteration retained the dataset and key questions and kept the focus on the mathematics. The final iteration resulted in the task shown in Figure 3.

Table 3

Summary of Prompts and Outcomes for Case Two

Iteration	Prompt	Output	Nature of Improvement	Limitations
1	Initial comprehensive prompt requesting socially critical, mathematically rich task	Request to summarise into half-page task card (and A5 layout)	N/A	Too much information; included conversational tone; formatting dense; some redundancy
2	Request to convert to printable worksheet	Reorganised into sequenced sections (calculations, modelling, representation, reflection); included tables for working	Improved organisational clarity and classroom usability; strengthened visibility of mathematical steps	Lengthy; potential cognitive overload; not classroom 'ready'
3	Request to summarise into half-page task card (and A5 layout)	Condensed to core investigative steps (measure access, model solutions, justify decision); retained dataset and key question	Increased concision and focus; preserved rigour while improving efficiency	Assumes a lot of prior knowledge; would require teacher scaffolding to support students

Generated SJM Task Two

As Figure 3 shows, ChatGPT generated a classroom ready task card that could be used directly with students. The task integrated inequity into the desired context and provided explicit numerical data that students could use to allocate resources. It required students to make decisions and justify their reasoning. The final task incorporated many, if not all, of the SJM design principles outlined in Table 1.

Figure 3

The Final Version of SJM Task Two

Maths Investigation Card

Is Healthcare Access Fair?

In Australia, where you live can affect how quickly you can see a doctor. Some people live minutes from a clinic, while others travel hours or rely on flying services for care.

Your Challenge
You are advising the government on where to build **30 new health clinics**. Use mathematics to decide what is fair.

Key Question
Does equal always mean fair? Who benefits from your decision, and who might still miss out?

Healthcare Access Data

Region	Population	Clinics	Avg Travel Time (min)
Metro	16,000,000	800	12
Inner regional	5,000,000	160	28
Outer regional	2,200,000	60	55
Remote	350,000	10	140

What to Do

- Calculate people per clinic and clinics per 100,000 people.
- Compare regions using tables, graphs, or ratios.
- Test different ways to allocate the 30 new clinics.
- Decide which plan is fairest and justify using maths.

Discussion and Conclusion

The first research question asked what is the iterative process of designing and refining ChatGPT prompts for SJM tasks, and what limitations does it present? Across both cases, the

prompting process was effective in moving outputs from abstract, teacher-facing descriptions towards student-facing tasks with accessible language and inquiry, consistent with Spasić and Janković's (2023) finding that structured prompts with explicit contextual cues generate more pedagogically coherent outputs. However, the two cases illustrate different prompting trajectories. In Case 1, revisions were primarily about the language, not the structure, and explicit prompting was required so that the task provided students with actual data to work with. In Case 2, the initial output was mathematically sound compared with Case 1. This suggests that the quality of the prompting process depends on the teacher's capacity to recognise what is missing, both pedagogically and mathematically (content).

The second research question asked to what extent tasks generated by ChatGPT aligned with the characteristics of the SJM framework. Both tasks from each case demonstrated meaningful alignment with several core design principles of the framework. However, Task 1 implied the social justice elements, rather than embedding them structurally, leaving the mathematical modelling open, whereas Task 2 came closer to meeting the dual criteria of being both socially meaningful and mathematically rich (Matthews et al., 2022). In task 2, the mathematics was explicit and intentional, with the 'real data' (Frankenstein, 2009) included.

The iterative prompting process itself offers a useful model for teacher engagement with generative AI tools. Rather than treating a single prompt as a finished product, teachers need to act as an active evaluator, requesting successive revisions to ensure their pedagogical content knowledge is driving what the AI is producing (Cameron & Mesiti, 2024). Primary school teachers could use the characteristics of the SJM task framework both as a generative prompt structure, and as an evaluative lens, when working with AI tools. By understanding what distinguishes a mathematically rich, equity-focused task from one that is merely contextualised in social language, teachers are better positioned to direct AI tools toward outputs of genuine pedagogical value. In doing so, they could use the characteristics of the SJM task framework to generate mathematical tasks that improve their students' ability to analyse complex societal problems and evaluate a range of potential solutions, thereby promoting informed citizenship.

Acknowledgements

This study forms part of a larger project approved by the ACU Human Research Ethics Committee (HREC) under application ID 2025-4029H.

References

- Baidoo-Anu, D., & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI*, 7(1), 52-62. <https://doi.org/10.61969/jai.1337500>
- Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal*, 9(2), 27-40. <https://doi.org/10.3316/QRJ0902027>
- Bragg, L. A. (2021). A social justice approach to teaching mathematics. *Australian Primary Mathematics Classroom*, 26(1), 25-29. https://drive.google.com/file/d/1JP00gD6TBVCW_0w6QwKCKiYqzAfpCFUM/view
- Cameron, S., & Mesiti, C. (2024). What Kind of Mathematics Teacher Is ChatGPT? Identifying the Pedagogical Practices Preferred by Generative AI Tools When Preparing Lesson Plans. In J. Višňovská, E. Ross, & S. Getenet (Eds.), *Surfing the waves of mathematics education. Proceedings of the 46th annual conference of the Mathematics Education Research Group of Australasia* (pp. 135-142). Gold Coast: MERGA.
- Canogullari, A., & Radmehr, F. (2025). Task design principles in mathematics education: a literature review. *International Journal of Mathematical Education in Science and Technology*, 1-33. <https://doi.org/10.1080/0020739X.2025.2457365>
- D'Ambrosio, U., & D'Ambrosio, B. S. (2013). The role of ethnomathematics in curricular leadership in mathematics education. *Journal of Mathematics Education at Teachers College*, 4(1). <https://doi.org/10.7916/jmetc.v4i1.767>

- Daugherty, M. K., & Carter, V. (2025). Ways of thinking and STEM-based problem solving: Toward the future. In L. D. English & T. Lehmann (Eds.), *Ways of Thinking in STEM-based Problem Solving* (pp. 263-272). Routledge. <https://doi.org/10.4324/9781003404989-19>
- English, L. (2025). STEM-based problem solving in a new era. In L. D. English & T. Lehmann (Eds.), *Ways of Thinking in STEM-based Problem Solving* (pp. 1-14). Routledge. <https://doi.org/10.4324/9781003404989-1>
- Ernest, P. (2007). Mathematics education ideologies and globalization. *Philosophy of Mathematics Education*, 21.
- Frankenstein, M. (2009). Developing a critical mathematical numeracy through real real-life word problems. In L. Verschaffel, B. Greer, W. Van Dooren, & S. Mukhopadhyay (Eds.), *Words and worlds: Modelling Verbal Descriptions of Situations* (pp. 111-130). Brill. https://doi.org/10.1163/9789087909383_008
- Gutiérrez, R. (2012). Context Matters: How Should We Conceptualize Equity in Mathematics Education? In B. Herbel-Eisenmann, J. Choppin, D. Wagner, & D. Pimm (Eds.), *Equity in Discourse for Mathematics Education: Theories, Practices, and Policies* (pp. 17-33). Springer Netherlands. https://doi.org/10.1007/978-94-007-2813-4_2
- Gutstein, E. (2006). *Reading and writing the world with mathematics: Toward a pedagogy for social justice*. Taylor & Francis. <https://doi.org/10.4324/9780203112946>
- Gutstein, E. (2007). So one question leads to another”: Using mathematics to develop a pedagogy of questioning. In N. S. Nasir & P. Cobb (Eds.), *Improving access to mathematics: Diversity and equity in the classroom* (pp. 51-68). <https://www.tcpres.com/improving-access-to-mathematics-9780807747285>
- Gutstein, E. (2016). “Our issues, our people—Math as our weapon”: Critical mathematics in a Chicago neighborhood high school. *Journal for research in mathematics education*, 47(5), 454-504. <https://doi.org/10.5951/jresmetheduc.47.5.0454>
- Ibrahim, J. H. (2023). *The art of asking ChatGPT for high-quality answers: A complete guide to prompt engineering techniques*. Independently published.
- Matthews, L. E., Jones, S. M., & Parker, Y. A. (2022). *Engaging in Culturally Relevant Math Tasks, K-5: Fostering Hope in the Elementary Classroom*. Corwin Press. <https://doi.org/10.4135/9781071841655>
- Ödmo, M., Chronaki, A., & Bjorklund Boistrup, L. (2023). Bringing Critical Mathematics Education and Actor–Network Theory to a Statistics Course in Mathematics Teacher Education: Actants for Articulating Complexity in Student Teachers’ Foregrounds. *Education Sciences*, 13(12). https://mdpi-res.com/d_attachment/education/education-13-01201/article_deploy/education-13-01201-v3.pdf?version=1702298673
- Skilling, K., & Park, W. (2025). STEM in a changing world: ‘Insider’ and ‘outsider’ perspectives. In L. D. English & T. Lehmann (Eds.), *Ways of Thinking in STEM-based Problem Solving* (pp. 15-28). Routledge. <https://doi.org/10.4324/9781003404989-2>
- Skovsmose, O. (2012). *An invitation to critical mathematics education*. Springer Science & Business Media. <https://doi.org/10.1007/978-94-6091-442-3>
- Skovsmose, O. (2023). *Critical mathematics education*. Springer Nature. <https://doi.org/10.1007/978-3-031-26242-5>
- Spasić, A. J., & Janković, D. S. (2023). *Using ChatGPT standard prompt engineering techniques in lesson preparation: role, instructions and seed-word prompts* In 58th international scientific conference on information, communication and energy systems and technologies. IEEE. <https://doi.org/10.1109/ICEST58410.2023.10187269>
- Stake, R. E. (1995). *The art of case study research*. Sage.
- Stefflitsch, D., & Kollosche, D. (2025). Students’ perceptions of Critical Mathematics Education: an exploration. *Educational Studies in Mathematics*, 120(2), 225-247. <https://doi.org/10.1007/s10649-025-10423-y>
- Stinson, D. W., & Wager, A. (2012). A sojourn into the empowering uncertainties of teaching and learning mathematics for social change. *Teaching mathematics for social justice: Conversations with educators*, 1(1), 3-18. National Council of Teachers of Mathematics (NCTM).
- Sullivan, P., Clarke, D., & Clarke, B. (2012). *Teaching with tasks for effective mathematics learning* (Vol. 9). Springer Science & Business Media. <https://doi.org/10.1007/978-1-4614-4681-1>
- Wright, P. (2021). Transforming mathematics classroom practice through participatory action research. *Journal of Mathematics Teacher Education*, 24(2), 155-177. <https://doi.org/10.1007/s10857-019-09452-1>