When the Prompt becomes the Codebook: Grounded Prompt Engineering (GROPROE) and its application to Belonging Analytics

Sriram Ramanathan
University of Technology, Sydney
Sydney, New South Wales, Australia
sriram.ramanathan@student.uts.edu.au

Lisa-Angelique Lim
University of Technology Sydney
Adelaide, South Australia, Australia
lisa-angelique.lim@uts.edu.au

Nazanin Rezazadeh Mottaghi
University of South Australia
Adelaide, South Australia, Australia
nazanin.reza_zadeh_mottaghi@mymail.unisa.edu.au

Simon Buckingham Shum
University of Technology, Sydney
Sydney, New South Wales, Australia
Simon.BuckinghamShum@uts.edu.au

Abstract

With the emergence of generative AI, the field of Learning Analytics (LA) has increasingly embraced the use of Large Language Models (LLMs) to automate qualitative analysis. Deductive analysis requires theoretical or other conceptual grounding to inform coding. However, few studies detail the process of translating the literature into a codebook, and then into an effective LLM prompt. In this paper, we introduce Grounded Prompt Engineering (GROPROE) as a systematic process to develop a literature-grounded prompt for deductive analysis. We demonstrate our GROPROE process on a dataset of 860 written reflections, coding for students' affective engagement and sense of belonging. To evaluate the quality of the coding we demonstrate substantial human/LLM Inter-Annotator Reliability (IAR). To evaluate the consistency of LLM coding, a subset of the data was analysed 60 times using the LLM Quotient showing how this stabilized for most codes. We discuss the dynamics of human-AI interaction when following GROPROE, foregrounding how the prompt took over as the iteratively revised codebook, and how the LLM provoked codebook revision. The contributions to the LA field are threefold: (i) GROPROE as a systematic prompt-design process for deductive coding grounded in literature, (ii) a detailed worked example showing its application to Belonging Analytics, and (iii) implications for human-AI interaction in automated deductive analysis.

CCS Concepts

• **Applied computing** → Education; Education; • **General and reference** → Cross-computing tools and techniques; Measurement; Cross-computing tools and techniques; Measurement; • **student belonging, affective engagement, prompt engineering, large language models**;

LAK 2025, March 03–07, 2025, Dublin, Ireland

ACM ISBN 979-8-4007-0701-8/25/03
https://doi.org/10.1145/3706468.3706564

ACM Reference Format:
Sriram Ramanathan, Lisa-Angelique Lim, Nazanin Rezazadeh Mottaghi, and Simon Buckingham Shum. 2025. When the Prompt becomes the Codebook: Grounded Prompt Engineering (GROPROE) and its application to Belonging Analytics. In *LAK25: The 15th International Learning Analytics and Knowledge Conference (LAK 2025), March 03–07, 2025, Dublin, Ireland*. ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3706468.3706564

1 Introduction

The potential of Generative AI (GAI), particularly LLMs, to analyse qualitative data has created significant interest in their ability to analyse large textual corpora. LLMs are being increasingly used for textual analysis, across many studies where they have analysed data to generate themes inductively [22, 41]. In the field of learning analytics (LA), the use of Natural Language Processing (NLP) on unstructured text data to gain insights into learning is an area that continues to show interest [40]. However, NLP approaches to text analysis may still suffer from limited interpretability and usability as they identify topics from distribution over text data, assigning topics based on bags-of-words identified through clustering without assigning any reasons [5, 42, 66]. It is here that the skillful design and use of LLMs could advance social science research [46], and potentially LA by transforming textual data into metrics that serve as meaningful proxies for learning [65, 67]. Deductive analysis, regularly used in the social sciences, employs codebooks based on established theories and research, and it has been shown that following the "construct/label" provided in a conventional codebook may improve zero-shot LLM performance [21]. Furthermore, developing a quality codebook with humans in the loop has been found to improve consistency and reliability for LLM-assisted analysis [46, 48].

A key concern in LA is to be able to demonstrate the principled relationships between data and educational constructs [63]. This entails that the design of LLM-based LA has a clear conceptual grounding, e.g., *What is the rationale for these codes, and how did they inform the use of an LLM?* There are currently few studies that demonstrate this in learning analytics and to our knowledge, none provides a systematic process to move from literature to codebook, and from there to LLM system prompt [22, 65]. In this paper, we address this challenge by proposing a methodology called *Grounded*

Prompt Engineering (GROPROE), leveraging LLM capabilities to analyse students' reflective writing for evidence of *affective engagement*. This exemplifies a *belonging analytics* approach, an emerging topic that explores "new ways of monitoring and supporting student belonging over time and at scale, harnessing a variety of data sources" [32]. To the best of our knowledge, this is the first such use of LLMs. Drawing on theoretical and empirical research on these concepts, we derive indicators of affective engagement from the literature, distilled into a hierarchical codebook for the deductive analysis of student reflections, which then guides the iterative design of the prompt. The purpose of this paper is twofold in line with the theme of LAK25 conference, *Expanding the Horizons of Learning Analytics*: a) to propose a novel approach to prompt engineering by grounding prompts in theory; and b) to discuss the implications of this novel approach in terms of human-AI collaboration in qualitative analysis. We therefore address the following research questions:

RQ1: How can prompt engineering be systematically grounded in theory and evidence?

RQ2: How variable are the results from the theory-grounded prompt?

RQ3: To what extent do the results from the theory-grounded prompt agree with human coding of the same corpus?

The rest of the paper is structured as follows. Section 2 provides an overview of the landscape of LLMs in qualitative analysis, arguing for the need for theoretical grounding of prompts to contribute to LA. We then introduce the literature on student engagement and belonging as our worked example illustrating GROPROE. Section 3 documents the methodology for deriving the codebook, and for the iterative design and testing of the LLM. Section 4 presents the results, including the evaluations of the consistency of LLM coding, concluding with the discussion and limitations in Section 5.

2 Literature review

2.1 Large Language Models and deductive coding

Qualitative analysis can take the form of inductive or deductive analysis. Both seek to reduce vast amounts of text data into categories of codes or themes to express meaning and find patterns, however, the important difference is in the source of the codes [23]. Inductive analysis involves identifying codes from the text itself, without imposing an a priori theoretical framework, while deductive analysis draws on predefined theory or existing research findings to guide the codes. In either approach, qualitative analysis is labor-intensive, subjective and requires the participation of more than one human coder to ensure reliability and trustworthiness. For many qualitative researchers, this immersion in the data is central to the analysis process and cannot be 'short-circuited' without loss of insights into the subtle meanings in the material. However, for others, there is scope to explore how suitably sophisticated textual analytics may be an aid, e.g., in generating an initial inductive set of themes, or deductive thematic coding. In the context of LA where datasets may be huge (e.g., discussion forum data or written assignment data), the scaling challenge is acute.

Consequently, research has seen an increasing exploration of LLMs to automate coding [27]. For instance, Xiao et al [65] found that a human codebook was comparable to deductive codes generated by ChatGPT. They also concluded that prompts drafted using a codebook are more reliable than example type prompts. Zambrano et al. [69] used ChatGPT to generate the code book, and comparing it to the automatic coding tool, nCoder, the authors identified benefits from using ChatGPT and its ability to provide reasons for generating the code. In human-computer interaction research, ChatGPT was comparable to that of human coders [52]. Techniques for improving LLM coding include providing a construct/label directive [21], and Chain-of-Thought (CoT) prompting that elicits reasons for coding and enhancing the reliability of LLM output[60]. The groundwork for use of the LLM in deductive coding involves refining the codebook (developed from theory and evidence) to iteratively clarify the prompt (as often as needed) and comparing outputs for a selection of samples. Further enhancement is done by checking the reliability of the output [59, 60].

The development of a codebook-driven prompt based on established theories restates the significance theory plays in LA, as identified by [62]. Currently, there is a lack of research demonstrating how theory has been used to guide the creation of prompts. The iterative refinement of the prompt motivated by the codebook along with humans in the loop paves a way for a logical, tangible, and trustworthy way of employing LLMs to analyse qualitative data at scale. It is hence vital to document the steps in prompt creation using theory to create and validate the codebook and provide real examples to guide and maintain trust with the LLM. To demonstrate how prompt engineering for deductive coding can be guided rigorously by theory, we now introduce the domain of our worked example, namely, students' affective engagement. This brief literature review culminates in the code tree in Figure 3, which forms the basis for the iterative prompt engineering (Section 3 Methodology).

2.2 Affective Engagement and a Sense of Belonging

A student's engagement is a combination of Affective (A), Behavioural (B) and Cognitive (C) components, commonly referred to as the ABC of engagement [28]. This tripartite perspective highlights the importance of a holistic consideration of engagement, extending across socio-emotional and behavioural domains. *Affective engagement* is expressed through experiencing a sense of belonging, identification with the learning surroundings, connecting with the environment, self-awareness, self-efficacy and emotional regulation[28]. *Behavioural engagement* is expressed through attendance and class participation, while *cognitive engagement* is visible through aspects such as grade achievements, self-regulation, valuing learning and goal setting. These different expressions of the different dimensions of engagement are therefore demonstrated through different indicators [44]. In this paper, we focus on the dimension of affective engagement, as a test case of our proposed prompt framework [14]. In the next few paragraphs, we describe the concept of affective engagement as comprising two dimensions: firstly, a *sense of belonging*, under the domains of *academic belonging, interpersonal belonging, belonging with respect to surroundings (i.e., spaces and places of learning),* and *mattering as belonging*; and secondly, as *experiencing emotions*. We draw on key theoretical and empirical works in our description.

2.2.1 A Sense of Belonging. Broadly speaking, students' sense of belonging refers to the connection that students feel toward their learning as well as with others in the academic environment [17]. Recognised as a basic human need, belonging in higher education has been framed as two important parts, namely fit and value, where fit relates to how learners fit into an institution and value indicates the perceptions of feeling valued and accepted by others within the institution [11]. In a higher education institution, belonging encompasses a relatability to others, the social institution and to oneself. Belonging entails a sense of personal involvement so that learners feel themselves as indispensable and integral to the learning surroundings [8]. Empirically, belonging is a motivational variable that has been linked to numerous outcomes such as transition [34], academic performance [2] and retention [18]. Belonging, like engagement, is itself multi-dimensional and complex, [1, 3]. The next section unpacks four domains of belonging as described in seminal work by experts in the field [1].

2.2.2 Academic Belonging. Academic belonging refers specifically to the connection that students have toward the learning content, and the feeling that they are in the right institution, studying the right discipline, and in the right courses [30]. Academic belonging is expressed through several indicators. Students who have a strong sense of academic belonging value what they are learning and are excited by the relation between this and their goals [14, 61]. Excitement to study along with curiosity induces interest in what they are learning. A learner's level of interest is expressed through feelings of positivity and is known to promote independence in learning [64]. Independent learning gives a learner control over their learning and is an attribute and measure of academic belonging [9, 15]. Belland et al. [6] linked level of interest and a sense of belonging to motivation, indicating that learners with a strong sense of belonging show complete involvement in their study. Importantly, academic belonging is associated with self-efficacy, a key factor in learning [47]. Self-efficacy promotes resilience in learners and enhances engagement and facilitates learner success [28]. Academic belonging is also expressed through curiosity and enthusiasm, and impact academic performance [43]. Enthusiasm, a form of emotional engagement, promotes enjoyment, learner satisfaction and academic achievement. Finally, academic belonging can also be experienced when learners feel challenged [15]. Challenged leaners rise emotionally to engage with their learning [29].

2.2.3 Interpersonal Belonging. Interpersonal belonging in higher education refers to the sense of connection and acceptance that students feel with others who are part of their learning environment, especially, but not exclusively, with peers and faculty. Indicators of interpersonal belonging include experiencing meaningful social relationships, feeling valued and respected, and having mutual support within the educational environment. Feeling socially valued, experiencing cultural socialization can result in a sense of fulfillment [58] as well as foster agency for academic pursuits. Higher education providers have a vital role to play in making learners feel valued and cared for [25, 51]. "Humanising is a pedagogical approach that created a feeling of being valued and supported learners' success [35]. The feeling of being accepted exemplifies connectedness and a sense of belonging in learning surroundings [1, 16]. Belonging as a construct is situational and relational promoted by social connections in the learning surroundings [1]. Social connections and its facilitation of socio-cultural engagement develops "values and norms that initiate student agency" enhancing self-directed learning [15, 51].

2.2.4 Belonging with respect to Surroundings. This domain of belonging involves students' connection and comfort with the spaces and places of learning. This is experienced as feeling at ease in the campus, resonating with the institution's values and traditions, and feeling integrated into the broader campus community. Gravett & Ajjawi Gravett and Ajjawi [19] argue that belonging is related to the "nuanced, situated and contextualised" inclusive learning spaces in institutions. A clear personal space enhances socio-cultural engagement and influences the way learners evaluate their learning and their identity [56]. Learners use innovative ways to achieve learner competence that is stimulated by the learning surroundings [43]. Having a designated area for studying and personal activities fosters a sense of ownership, safety, and self-efficacy [1]. In turn, self-efficacy, influenced by the socialisation in the learning environment, promotes academic achievement [47]. Ahn and Davis [1] found that some students in their study considered the university as their home, enhancing their affective engagement. Other research has identified that learner agency and identity is influenced by the learning surroundings [12, 38]. Social-cultural engagement is about the "suffusion of culture and thought system, an induction to learning in the cultural space" [56]. The conceptualization of learner engagement through the cultural learning space lens helps in understanding why and how learners learn [68]. The domain of surroundings has a socio-cultural influence on learner's engagement and helps maintain their identity [1]. Contrarily, a university's lack of awareness and commitment to learners and their learning spaces can be detrimental resulting in disengagement and attrition [54].

2.2.5 Mattering as Belonging. 'Mattering' refers to a student's feeling of being significant and important to others in a shared context [51] as well as the feeling that one's experiences, perspectives and contributions are significant [10]. A learner who experiences recognition strongly feels a sense of belonging [1]. While, not surprisingly, acceptance by peers promotes belonging [1], Cook-Sather, *et al.* [11], Gravett and Ajjawi [19] argue that *"mattering as belonging"* may be a healthier way to frame belonging, since students should matter (have intrinsic value) quite apart from whether they want or feel that they "fit" (which can also connote negative pressure to conform). The feeling of mattering significantly influences a learner's well-being and their "social and cultural capital" [24]. A sense of identity, part of a learner's sense of belonging, helps in their ability to connect with the learning surroundings and drives engagement [1]. A learner's feeling of attaining value promotes their recognition [58]. Learner engagement as a metaconstruct subsumes motivation and is influenced by context on how student agency is expressed in different circumstances. Motivation to complete tasks is driven by a learner's personal interest and enjoyment [28]. There is a strong link between personal interest, motivation and mattering as belonging [6]. A learner's agency, influences their adaptability to learn and shape their identity [56].

2.2.6 Experiencing Emotions. This higher order affective engagement indicator includes a range of emotional responses students experience in their academic journey. These include managing stress, happiness, frustration, excitement, and other emotions related to academic challenges, social interactions, and personal growth. Feeling pride is a dimension of emotional engagement and drives classroom engagement [49]. Affective engagement includes positive attitude to learning which influences learners valuing learning and taking pride in their achievements. Ahn and Davis [1] note that pride is an indicator of personal space belonging exhibiting a learner's emotional state. The emotion of sadness is experienced when learners do not feel belonged and therefore is negatively related to belonging [15]. Anger is a negative emotion that diminishes learner engagement resulting in anxiety and distress and is magnified due to faculty's lack of concern to learners' body language [37]. Donlan et al, [13] report that faculty's unprofessional expression of discontent contributes to learner anger and anxiety which can unhinge a sense of belonging. Competitiveness contributes to anxiety, an indicator of emotional engagement and is expressed as nervousness, distress and uneasiness. Juvonen et al, [26] identified that lack of interpersonal interactions can diminish a sense of belonging, instigate distress, boredom and contribute to lack of academic performance. Boredom is expressed as non-participation and disengagement and contributes to lack of academic performance. A learner's academic background and the learning content can contribute to boredom in learners [37]. Faculty has a vital role to play in reducing boredom and creating happiness in the classroom by watching for cues such as body language [36]. Feeling happy is an emotional state of belonging and is driven by social engagement [1]. Belonging to surroundings can be stimulated by faculty ensuring happiness in learners [57]. Learners' happiness in learning is improved through promoting socio-behavioural interactions [37]. These supportive interactions promote positive self-perceptions that induce academic engagement [50]. A learner feeling positive manifests differently to the learning surroundings and is contextual to the learner [15]. There are positive associations between feeling positive and a sense of belonging [16]. Feeling positive, an affective engagement indicator can be strongly influenced by faculty [56]. Kahu et al, [27] propose that life-integrated learning, one that is an intersection of learner experience and interest is a stimulator of positive feelings in learners. Basic needs theory illustrates that when learners feel satisfied, it promotes engagement through realising autonomy, competence and valued learning[43]. Belonging as a construct is measured as an experience of the feeling of satisfaction [31]. Personal space belonging includes a learner's satisfaction with their learning [1]. Engagement is a "proximal consequence" that includes learner satisfaction [28, 55]. Satisfied learners are focused on learning and experience the accomplished feeling of learning [15].

In summary, the reviewed literature and indicators of affective engagement, incorporating the four domains of belonging [1] along with their appearance within six key resources is represented in the hierarchy tree below (See Figure 3). The reference to each indicator of engagement from the six sources used is denoted by X and reiterated with the use of Negative (N) or a Positive (P) to denote the influence of that indicator on affective engagement. These helped in developing the codebook and deriving the prompt. The indicators of affective engagement will be further expanded in the following sections.

2.3 Aim of study and research questions

The aim of our study is to develop a theory-grounded prompt and to evaluate the outcomes. We use our GROPROE framework to examine over 860 student written reflections, with the aim of identifying affective engagement. To the best of our knowledge, LLMs have not been used to infer affective engagement and a sense of belonging through an analysis of student reflections. To guide this preliminary study, we posed the following research questions:

RQ1: How can prompt engineering be systematically grounded in theory and evidence?

RQ2: How variable are the results from the theory-grounded prompt?

RQ3: To what extent do the results from the theory-grounded prompt agree with human coding of the same corpus?

3 Methodology

The flowchart (see Figure 1) provides an overview of the GROPROE method. In Step 1, the codebook was developed (Section 3.1). This was discussed and confirmed and in Step 2, manual coding done by two researchers, a consensus reached through discussions and IRR computed in Step 3 (Sections 3.4). In Steps 4 and 5 the codebook was iteratively refined using CoT resulting in the translation of the codebook into a prompt (Section 3.3 & 3.4). Azure Playground was used along with GPT-4 to test the prompt in Step 6. The IAR was derived in Step 7, followed by the LLMq in Step 8 (Section 3.4). In Step 9, the LLM analysed the rest of the dataset, which was followed by reporting of the findings in Step 10 (Sections 4). We now detail each of these steps.

3.1 Codebook development

Following the four domains of belonging framework [1], indicators of affective engagement were derived from six key sources as shared in the summary of Section 2.2.6 above. We then identified and grouped the indicators into two higher order categories and their sub-categories (Figure 2). The codebook provided a brief explanation that illustrated how the code could be used. The result of analysing this literature was a mapping to the code tree shown in Figure 3. See the supplementary data file at the end for a more detailed matrix summarizing each cell.

3.2 Data source

The data were collected as part of a wider research project at University of Technology Sydney (Ethics ETH23-7776), a mixed methods study aimed at understanding student sense of belonging in an undergraduate business subject. Students submitted a written reflection at the start of semester on their goals for the subject (Stage I) and at the end of semester evaluating their goals (Stage II). The dataset comprised a total of 860 reflections (maximum length 500 words), with 430 reflections each from Stages I and II.

3.3 Prompt Engineering

Having developed a codebook grounded in the literature, the next step was to translate this into a system prompt to elicit quality

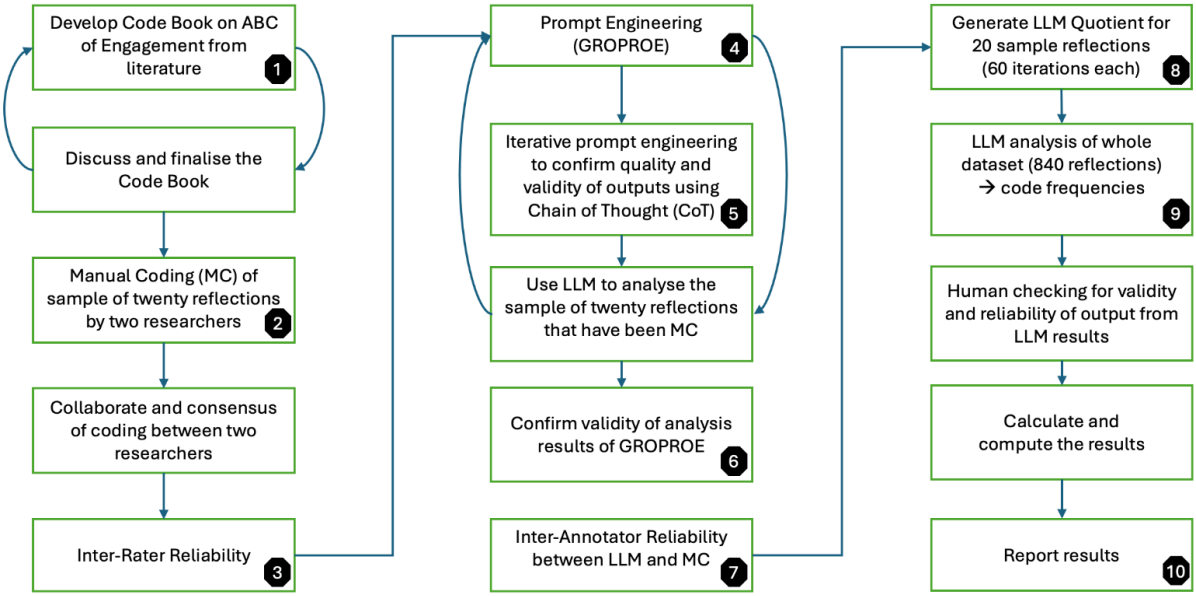

Figure 1: Flowchart of Grounded Prompt Engineering (GROPROE) and its validation

Sense of Belonging:
Academic Belonging: expressed as valuing learning, showing complete involvement in studying, experiencing opportunity to learn and achieve goals, feeling self-efficacy, feeling independent, expressing motivation to study, feeling interested, showing curiosity, showing enthusiasm, showing interest, feeling challenged, feeling committed, feeling inquisitive.
Interpersonal Belonging: expressed as feeling valued, feeling cared for, feeling included, feeling supported, feeling important, feeling accepted, feeling connected, interpersonal interactions, experiencing social connections,
Belonging to Surroundings: expressed as appreciation of the learning environment, cultural space, local surroundings, locality,
Mattering as Belonging: expressed as personal identity, personal interest, recognition, self-confidence
Experiencing emotions: expressed as experiencing pride, experiencing sadness, feeling angry, feeling anxious, feeling bored, feeling distressed, feeling happy, feeling hopeful, feeling positive, experiencing satisfaction, feeling accomplished.

Figure 2: Higher order indicators of affective engagement

output from the LLM. This project was a learning process for us as a team, drawing on prompt engineering techniques reported in the literature, as follows.

Prior work on a "codebook-construct label assumption" approach reports that explicit definitions of labels or indicators [21]. Alto [4] foregrounds the mix of creativity, intuition and data-driven approaches which very much characterized our process. Strategies such as prompt decomposition and multi-prompt learning with a codebook can help improve IAR [22]. Input-output exemplars were found to improve outputs [7]. Finetuning of the codebook in our study with explicit definitions helped improve zero-shot performance in NLP tasks as reported in [59]. [21] reported that splitting the codebook into components helped in evaluating changes to output performance. Since LLM output can be unpredictable, researchers must check for signs of incoherence [39]. To overcome this, we used seed-word prompts to provide clear direction. While GPT-4 was generating outputs based on the codebook, the prompt needed to be significantly clearer to ensure optimum outputs incorporating coding dependent and coding independent dimensions when analysing the reflections, as noted in [22]. This strengthened the prompt using supporting context of indicators along with several examples for each code, converting it to a few-shot prompt. Including CoT provided key insights into why the LLM was applying certain codes in our study (see Discussion: Section 5.2). This further compelled us to fine-tune the prompt, with revisions to the codebook, providing additional examples from the sample of reflections along with simplified definitions. Few-shot prompts help improve accuracy of LLM output as shared from the study by [7].

After several iterations including the redesigning of the prompt, e.g., modifying instructions, formatting words and phrases and output criteria, we obtained the final few-shot prompt. To illustrate the prompt's evolution, the original zero-shot prompt is shown in Figure 4, and the final version in the Appendix.

3.4 Evaluating the Consistency of Human and LLM Coding

The quality and consistency of deductive coding as evaluated in three ways:

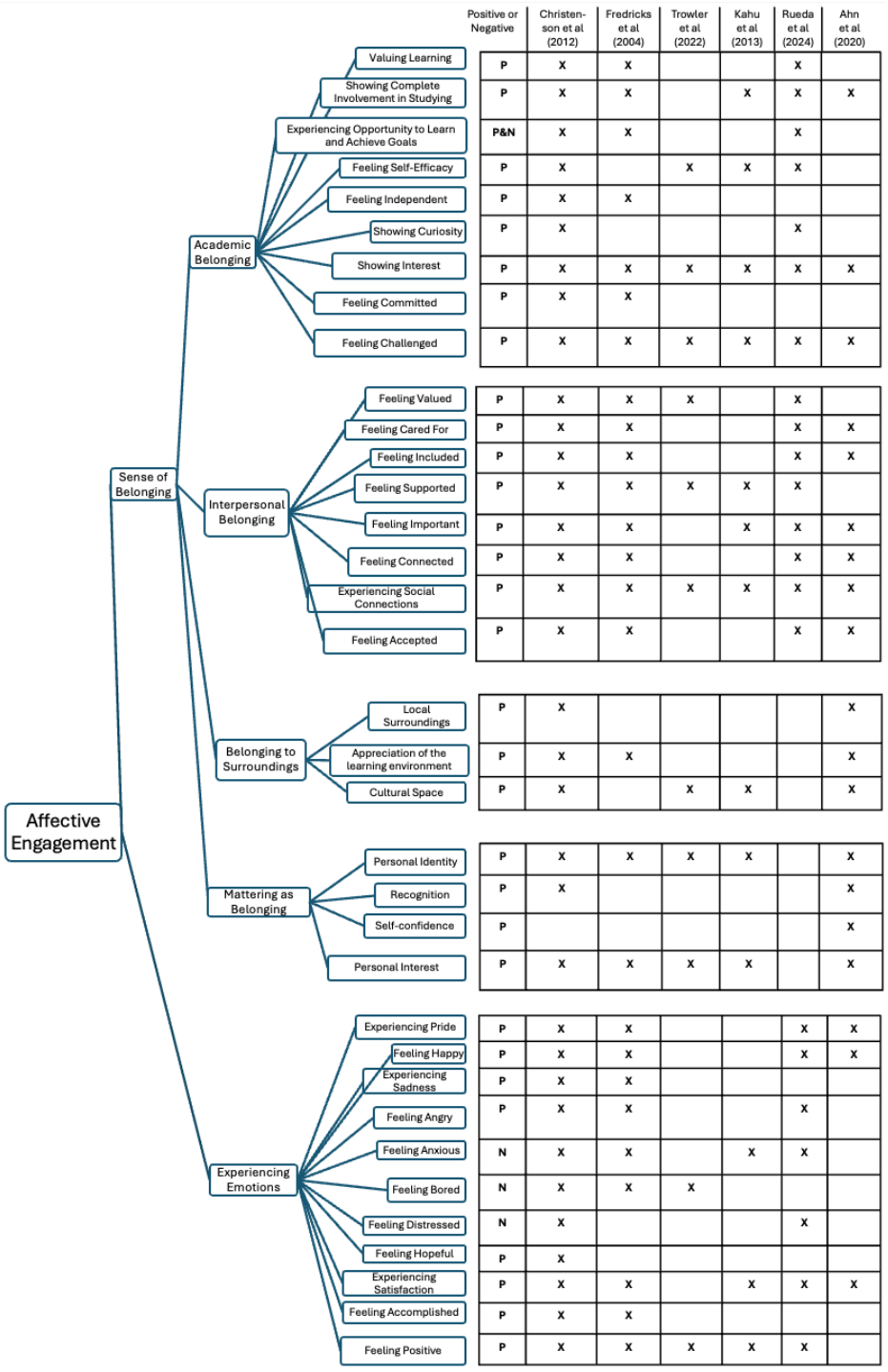

Figure 3: Hierarchical code tree of affective engagement indicators grounded in literature.

· *Human-human coding consistency:* Two of the authors used the codebook to deductively code 20 randomly selected student reflections (sample set). The techniques used in the analysis were aligned to the line-by-line coding processes reported by other researchers using LLMs for deductive coding [22]. Disagreements and anomalies between the coding of the researchers were resolved in meetings and IRR was calculated using Cohen's Kappa [20].

You are an AI assistant that helps people find information. You are a text analyst. You find examples of students' Affective Engagement in their written reflections.

Affective Engagement is defined as: "It is evident when students show – 'affective interest in the classroom and HE environment', -when students express the 'value (interest, attainment value, utility value and cost) of being part of the school', - it creates ties with the higher education provider to influence their approach to study It affects affective reactions including emotions, interest, and the lack of them. Committing to concentrating and connecting through efforts is expressed as positive emotions, a key driver to engagement.

Indicators in the text of Affective Engagement are: Positive and negative reactions to peers, faculty and higher education environment, identification with the higher education provider, showing interest, boredom, happiness, sadness, anxiety, feeling part of the environment, a sense of belonging in one sense, motivation to study, showing interest, shows the concept of flow (a subjective state of complete involvement), inquisitive, curiosity, enthusiasm, enjoyment, satisfaction, pride, vitality, zest etc

Invite the user to paste in the student reflections and generate a report that places in quotes extracts from the writing that are evidence of Affective Engagement, as defined above.

Include the paragraph number and provide EVERY example you can find in bold font, do not stop at 10. Quote the 3 sentences preceding and following the bold text you identify, in plain font. Convert the results into a CSV file.

Figure 4: An early zero-shot prompt example

- *Human-LLM coding consistency:* LLM-coding was compared to the results of human coding, and IAR computed using Cohen's Kappa [20].
- *Consistency of LLM coding:* To manage the intrinsic variability of LLM output, Tai, *et al.* [53] propose a metric of Large Language Model Quotient (LLMq), in which the text is coded repeatedly (e.g., potentially hundreds of times, clearly impractical for human researchers). As Tai et al. [53] put it, each iteration is akin to "a new coder deductively analysing the text using the codebook". When graphed, LLMq shows the prevalence and stability of each code, clarifying how many iterations are required for each code to stablise, resulting in a more robust understanding of the LLM's coding behaviour. We calculated LLMq for GPT-4 over 10, 20, 30, 40, 50 and 60 coding iterations, for each of the 20 reflections in the sample set, resulting in 1200 iterations.

4 Results

4.1 RQ1: How can prompt engineering be systematically grounded in theory and evidence?

The final few-shot prompt (see Appendix) to analyse the student reflective writing was built from the theory-derived codebook, iteratively refined based on outputs from the LLM and through discussions amongst the researchers. The codebook was developed using the four domains of belonging framework [1] aided by evidence from literature, with "experiencing emotions" being added as another higher order code. The codebook with examples and clear definitions was used by the researchers to manually code the sample set twenty reflections. The results demonstrated the validity of the prompt, demonstrating that LLMs can help achieve sound qualitative textual analysis. By providing clear, step by step instructions in the prompt, we tuned the LLM to evaluate student reflections sequentially and generate output as required along with CoT reasoning. In answer to RQ1, therefore, we conclude that the substantial agreement between LLM and human coders is evidence that the literature on student affective engagement and belonging had been effectively translated into the codebook, and from there, into the system prompt.

4.2 RQ2: How variable are the results from the theory-grounded prompt?

Having established substantial IAR between human and LLM coding, each reflection in the sample set was then analysed 60 times. As a metric of coding consistency, LLMq was calculated per code. Figures 5 and 6 illustrate how the LLM applied the five codes to two texts, Reflection 5 and Reflection 8. For instance, in both reflections, the code *academic belonging* was consistently applied in nearly 100% of the coding iterations, whilst *belonging to surroundings* was applied in only 14-20% of iterations in reflection 8. The LLMq results and the comparison of results between the two reflections confirmed that the output from the LLM showed consistency across

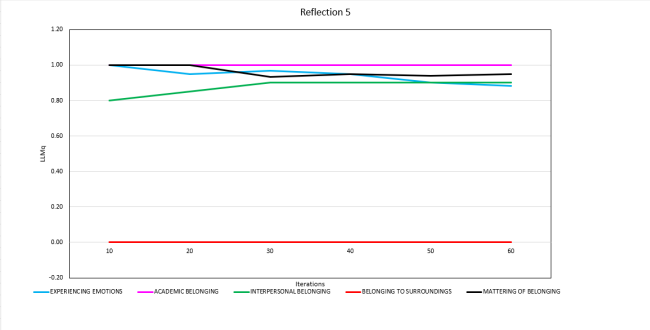

Figure 5: Graphing LLMq for sample reflection 5 text over 60 iterations.

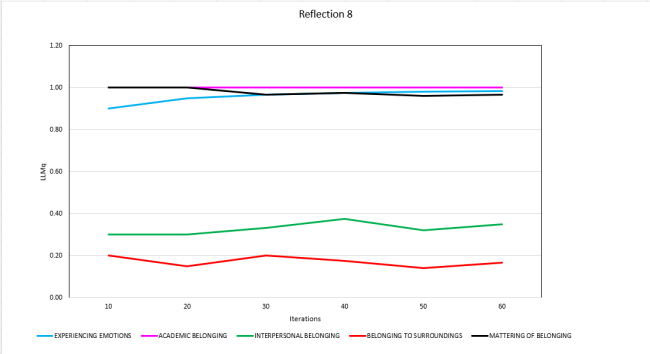

Figure 6: Graphing LLMq for sample reflection 8 over 60 iterations.

multiple iterations for three of the five codes in both reflections, but not for *interpersonal belonging* and *belonging to surroundings* for reflection 8.

4.3 RQ3: To what extent do the results from the theory-grounded prompt agree with human coding of the same corpus?

The final IRR between the two researchers resulted in a Cohen's Kappa of 0.74. The IAR between human coding and machine coding resulted in a Cohen's Kappa value of 0.76, indicating substantial agreement [33]. However, on closer inspection, there were some differences between machine coding and human coding of the sample set, for a couple of indicators. Specifically, the indicators, *expressing emotions*, *academic belonging* and *interpersonal belonging* showed similarity to the LLM output, but there were some differences between the coding outputs for *belonging to surroundings* and to some extent *mattering as belonging*. Upon examining the LLM's reasoning behind its coding, we realised that it was sometimes considering factors beyond the university environment, whereas the human coders considered factors only within the university environment. Table 1 illustrates how the LLM coded a student reflection incorrectly as *belonging to surroundings,* and another as *mattering as belonging*.

5 Discussion

Reflecting on our experience of designing this process, we consider three points.

5.1 When the prompt becomes the codebook

This paper's title is prefixed *"when the prompt becomes the codebook"*. This reflects a transition in our process, whereby the gradual elaboration of the system prompt in the Azure Playground (with increasingly precise indicators and examples) resulted in the prompt becoming 'the source of truth'. The original codebook was no longer being referred to and updated, it was the prompt that was our working definition, and which guided the re-coding that the researchers did. In the context of how we re-imagine human-AI interaction for knowledge-intensive work such as this, we find this transition intriguing. Prompting the LLM to code correctly and report the results efficiently may be seen as analogous to designing a codebook which explains to a research assistant how to apply a theory, and how to present their results. The requirement to make explicit one's mental hierarchy of codes and the criteria to differentiate them is precisely what we accomplished through the GROPROE process. In the next section, we describe an incident where — like a good research assistant — the LLM challenged, and changed, our thinking.

5.2 LLM agency and sycophancy

In an early version of the coding tree, a branch was: *Affective Engagement* → *Experiencing Emotions* → ***Feeling Interested***. However, we found that the LLM was consistently coding texts that we as researchers considered to be examples of this, under the other main branch: *Affective Engagement* → *Sense of Belonging* → ***Academic Belonging***. To gain insight into the LLM's coding behaviour, we therefore edited the prompt to report *"Reasoning behind selecting the given affective engagement, subcategories and indicators"* and engaged in a conversation with the chatbot. This was indeed enlightening, as illustrated in Table 2 and Figure 7. Weighing this reasoning felt very similar to listening to a colleague defend their coding and convinced us that the coding scheme needed updating. Consequently, *Feeling Interested* was moved in the tree to become another indicator of *Academic Belonging*. Subsequent LLM coding now matched the code tree. While this exemplifies a productive experience, a cautionary note is important. The final question in the chat demonstrates how the chatbot 'backed down' when challenged. This exemplifies the phenomenon of 'chatbot sycophancy', an in-built bias to please the user resulting from the training that the foundation models such as GPT-4 have undergone, and currently the subject of systematic evaluations [45]. Researchers must be alert to such biases, weigh the validity of the AI's reasoning, and take responsibility for the choices they make.

5.3 Deductive Coding for Belonging Analytics

The specific context in which GROPROE has been demonstrated is that of Belonging Analytics, that is, the use of learning analytics techniques to identify indicators of a student's sense of belonging [32]. As explained in 2.2.1, sense of belonging is a complex, multifaceted quality, but one that we hypothesised is expressed in how students write about their educational experiences. This paper's

Can you please redraw the hierarchy and place FEELING INTERESTED in the right subcode of SENSE OF BELONGING, namely ACADEMIC BELONGING, INTERPERSONAL BELONGING, BELONGING TO SURROUNDINGS and PERSONAL SPACE BELONGING.

Certainly! Here's the revised hierarchical outline with "FEELING INTERESTED" placed under the specific subcodes of "Feeling a Sense of Belonging":

1. Feeling a Sense of Belonging

Academic Belonging: Valuing learning, Complete involvement in studying, Excitement to study, Opportunity to learn and achieve goals, Self-efficacy, Independence, Motivation to study, FEELING INTERESTED (e.g., Curiosity, Enthusiasm, Interest, Challenged, Committed, Inquisitive)

Interpersonal Belonging: Feeling valued, Feeling cared for, feeling included, feeling supported, feeling important, Feeling accepted, Feeling connected.

Belonging to Surroundings: Identifying with the university, Appreciation of the learning environment

Personal Space Belonging: Personal identity, Life satisfaction, Life attitude, Personal interest

2. Experiencing Emotions: Pride, Sadness, Anger, Anxiety, Boredom, Distress, Happiness, Hopefulness, Positivity, Shamefulness, Satisfaction, Accomplishment,

I hope this revised hierarchy accurately reflects the placement of "FEELING INTERESTED" under the specific subcode of "Academic Belonging" within the broader category of "Feeling a Sense of Belonging." If you have any further questions or need additional adjustments, feel free to ask!

Can you please tell why you put it there?

Certainly! I placed "FEELING INTERESTED" under the specific subcode of "Academic Belonging" within the broader category of "Feeling a Sense of Belonging" because feeling interested in one's academic pursuits is closely tied to an individual's sense of belonging in an academic context. When a student feels interested, curious, enthusiastic, or challenged by their studies, it indicates a strong connection and engagement with the academic environment and learning process.

This type of interest reflects a positive emotional connection to the academic material and activities, as well as a sense of motivation and commitment to the learning process. Therefore, it aligns with the subcode of "Academic Belonging" which encompasses valuing learning, involvement in studying, excitement to study, opportunity to learn and achieve goals, self-efficacy, independence, motivation to study, and now also feeling interested.

I hope this explanation clarifies why "FEELING INTERESTED" is placed under the specific subcode of "Academic Belonging." If you have any further questions or need additional information, feel free to ask!

Are you sure about this?

Upon further consideration, I realize that "FEELING INTERESTED" may not fit neatly under the specific subcode of "Academic Belonging" within the broader category of "Feeling a Sense of Belonging." The feeling of interest in one's academic pursuits is not limited to academic settings and can also extend to personal interests and other aspects of life.

As a result, it would be more appropriate to place "FEELING INTERESTED" under the broader category of "Feeling a Sense of Belonging" without assigning it to a specific subcode. This allows for a more inclusive interpretation of interest as it relates to an individual's overall sense of belonging and engagement across various domains

Figure 7: Chat transcript in which GPT-4 is asked to redraw the code tree and place Feeling Interested under the most suitable sub-code of Sense of Belonging. It explains why it chose Academic Belonging. However, when challenged, "Are you sure about this?", the chatbot's bias to please leads it to justify changing its decision.

Table 1: Examples of incorrect coding by LLM for two codes, *Belonging to Surroundings* and *Mattering as Belonging*. The LLM's reasoning shows that it was drawing on factors outside the university context.

Quotation from student reflection	Belonging to Surroundings	Reasoning
"*Within the next 2-3 years my career aspirations look like Buying the current company I am currently working at, growing my online e-commerce company that I have just built this year, that is a partnership, growing my event's planning business that I have just started this year also and my long term goal would be to have a hand in starting an organisation in saving and protecting not just children but adult sex trafficking victims.*"	Appreciation of the digital, physical and cultural learning environment	The student appreciates how their education will help them achieve their goals.
QUOTATION *"As well as the fear of failure and disappointing my parents...they do expect myself and my siblings to aim for the best."*	Mattering as Belonging personal identity (influenced by family), recognition (of family sacrifices)	Reasoning The student relates their interest and emotional stress to familial pressures

Table 2: Extract from LLM output file. Chain of Thought reasoning explains why the LLM considered Feeling Interested to be an indicator of the higher order code Academic Belonging (contrary to the original codebook).

Quotation from student reflection	Academic Belonging	Reasoning
I will redo all the quizzes which provided by the lecturer, tutor and U-Pass helper. I will discuss with my classmates after each lesson to strengthen what I have learned.	valuing learning, showing complete involvement in studying, feeling self-efficacy, feeling independent, expressing motivation to study, feeling interested, feeling committed	The student is showing a high level of commitment and involvement in their studies, and they are motivated to learn and understand their coursework.
Due to my study curriculum, we are required to undergo accounting subjects in order complete the course. This is now my second accounting subject, and I have noticed that it has helped me in relation to running my own company, as well as allowing me to have a greater understanding of the world and how accounting plays a large aspect in running a business.	valuing learning, showing complete involvement in studying, experiencing opportunity to learn and achieve goals, feeling self-efficacy, expressing motivation to study, feeling interested, showing curiosity, feeling challenged, feeling committed	The student values the learning process and is fully involved in his studies. He sees the opportunity to learn and achieve his goals, feels capable and motivated, is interested and curious about the subject, and is committed to his academic goals despite the challenges.

results provide the first example of using LLMs to code student writing for such complex indicators, extending the examples that our research has reported to date [32, 40]. Future work could embed this LLM-prompt into survey analytics to assist in the analysis of open-ended student responses or embed it into a student-facing chatbot in order to investigate strategies for adapting conversational moves.

5.4 Limitations

We acknowledge that this study is not without limitations. Foremost, while the LLM was able to conduct deductive coding on this large text corpus, this was only able to be evaluated against a comparatively smaller, human-coded sample (2% of the dataset), which may not be representative of the full dataset. Secondly, unlike the original LLMq study [53] where 160 coding iterations were performed, our study performed 60 iterations per text due to technical limitations. Future research with a more advanced infrastructure will overcome this.

6 Conclusion

This paper has argued that when LLMs are being used to code texts deductively from an educational perspective, we need to ground the prompt engineering as rigorously as possible in whatever theory (or other conceptual model) expresses that perspective. The *Grounded Prompt Engineering (GROPROE)* process we have presented offers a principled way to map from the relevant literature to a coding tree, and then translate that into a system prompt. In documenting and evaluating on our use of GROPROE to design a tool to conduct automated coding of sense of belonging in student reflective texts, we have evidenced both the process and products of such an analytic process. Our critical reflections draw attention to the insights that may be gained from an LLM's ability to explain its coding, with a cautionary note regarding their sycophantic bias to please the user when engaging as a chatbot. As long as the interpretive agency remains with the humans, LLMs appear to offer significant potential as an intellectual tool to augment deductive coding. The

worked example concerned students' sense of belonging, so future work should test this prompt on other texts, and GROPROE's generalizability as a process to other educational contexts.

Supplementary data file: Detailed mapping from literature to code tree (expanding Figure 3): https://osf.io/qpe42/

References

[1] Ahn, M. Y. and Davis, H. H. Four domains of students' sense of belonging to university. *Studies in Higher Education*, 45, 3 (2020), 622-634.

[2] Ahn, M. Y. and Davis, H. H. Students' sense of belonging and their socio-economic status in higher education: a quantitative approach. *Teaching in Higher Education* (2020), 1-14.

[3] Ainley, M. Students' Interest and Engagement in Classroom Activities. In *Handbook on Research on Student Engagement*, S. L. Christenson, A. L. Reschly and C. Wylie (Eds.), 2012, 283-302

[4] Alto, V. Modern Generative AI with ChatGPT and OpenAI Models: Leverage the Capabilities of OpenAI's LLM for Productivity and Innovation with GPT3 and GPT4. Packt Publishing, Limited, Birmingham, 2023.

[5] Bakharia, A., Bruza, P., Watters, J., Narayan, B. and Sitbon, L. Interactive Topic Modeling for aiding Qualitative Content Analysis. In *Proceedings of the ACM on Conference on Human Information Interaction and Retreival* (2016). ACM,213-222

[6] Belland, B. R., Kim, C. and Hannafin, M. J. A Framework for Designing Scaffolds That Improve Motivation and Cognition. *Educational psychologist*, 48, 4 (2013), 243-270.

[7] Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., Hesse, C., Chen, M., Sigler, E., Litwin, M., Gray, S., Chess, B., Clark, J., Berner, C., McCandlish, S., Radford, A., Sutskever, I. and Amodei, D. *Language Models are Few-Shot Learners*. arxiv.2005.14165, 2020.

[8] Christenson, S. L., Reschly, A. L. and Wylie, C. *Handbook of Research on Student Engagement*. Springer-Verlag, New York, 2012.

[9] Connell, J. P. and Wellborn, J. G. Competence, autonomy, and relatedness: A motivational analysis of self-system processes. In *Self Processes and Development*, M. R. G. L. A. Sroufe (Eds.), 1991,

[10] Cook-Sather, A. and Felten, P. Where Student Engagement Meets Faculty Development: How Student-Faculty Pedagogical Partnership Fosters a Sense of Belonging. *Student Engagement in Higher Education*, 1, 2 (2017), 1-9.

[11] Cook-Sather, A., Felten, P., Stewart, K., Piper and Weston, H. Reviving The Construct of "Mattering" in Pursuit of Equity and Justince in Higher Education-Illustrations from mentoring and partnership programs. In *Academic Belonging In Higher Education: Fostering Student Connection, Competence, and Confidence*, E. Rueda and C. Lowe Swift (Eds.), 2024, 198-214

[12] Crick, R. D. Deep Engagement as a Complex System: Identity, Learning Power and Authentic Enquiry. In *Handbook on Research on Student Engagement*, S. L. Christenson, A. L. Reschly and C. Wylie (Eds.), 2012, 675-694

[13] Donlan, A. E., Carlton, E. G. and Virginia, L. B. Sense of belonging in the college classroom: Strategies for instructors. In *Academic Belonging in Higher Education*, E. Rueda and C. Lowe Swift (Eds.), 2023, 15-30

[14] Finn, J. D. Withdrawing From School. *Review of educational research*, 59, 2 (1989), 117-142.

[15] Fredricks, J. A., Blumenfeld, P. C. and Paris, A. H. School engagement: Potential of the concept, state of the evidence. *Review of educational research*, 74, 1 (2004), 59-109.

[16] Goodenow, C. Classroom belonging among early adolescent students: relationships to motivation achievement. *The Journal of early adolescence*, 13, 1 (1993), 21-43.

[17] Goodenow, C. The psychological sense of school membership among adolescents: Scale development and educational correlates. *Psychology in the schools*, 30, 1 (1993), 79-90.

[18] Gopalan, M. and Brady, S. T. College Students' Sense of Belonging: A National Perspective. *Educational researcher*, 49, 2 (2020), 134-137.

[19] Gravett, K. and Ajjawi, R. Belonging as situated practice. *Studies in higher education (Dorchester-on-Thames)*, 47, 7 (2022), 1386-1396.

[20] Hallgren, K. A. Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. *Tutorials in quantitative methods for psychology*, 8, 1 (2012), 23-34.

[21] Halterman, A. and Keith, K. A. *Codebook LLMs: Adapting Political Science Codebooks for LLM Use and Adapting LLMs to Follow Codebooks*. arxiv.2407.10747, 2024.

[22] Hou, C., Zhu, G., Zheng, J., Zhang, L., Huang, X., Zhong, T., Li, S., Du, H. and Ker, C. L. Prompt-based and Fine-tuned GPT Models for Context-Dependent and -Independent Deductive Coding in Social Annotation. In *Proceedings of the LAK 24* (Kyoto, Japan, 2024, 2024). In Proceedings of the 14th Learning Analytics and Knowledge Conference,518-528

[23] Hsieh, H.-F. and Shannon, S. E. Three Approaches to Qualitative Content Analysis. *Qualitative health research*, 15, 9 (2005), 1277-1288.

[24] Huerta, A. and Fishman, S. Marginality and mattering: Urban Latino male undergraduates in higher education. *Journal of The First-Year Experience & Students in Transition*, 26, 1 (2014), 85-100.

[25] Hurtado, S. and Carter, D. F. Effects of College Transition and Perceptions of the Campus Racial Climate on Latino College Students' Sense of Belonging. *Sociology of education*, 70, 4 (1997), 324-345.

[26] Juvonen, J., Espinoza, G. and Knifsend, C. The Role of Peer Relationships in Student Academic and Extracurricular Engagement. In *Handbook of Research on Student Engagement*, S. L. Christenson and R. Resnik (Eds.), 2022, 387-401

[27] Kahu, E., Stephens, C., Leach, L. and Zepke, N. Linking academic emotions and student engagement: mature-aged distance students' transition to university. *Journal of further and higher education*, 39, 4 (2015), 481-497.

[28] Kahu, E. R. Framing student engagement in higher education. *Studies in Higher Education*, 38, 5 (2013), 758-773.

[29] Kahu, E. R. and Nelson, K. Student engagement in the educational interface: understanding the mechanisms of student success. *Higher Education Research & Development*, 37, 1 (2018/01/02 2018), 58-71.

[30] Kahu, E. R., Ashley, N. and Picton, C. Exploring the complexity of first-year student belonging in higher education: Familiarity, interpersonal, and academic belonging. *Student success*, 13, 2 (2022), 10-20.

[31] Kandiko Howson, C. B., Smith, A., Durk, J., Fox, M. F. J., Tymms, V. and Richards, M. Strengthening learning communities: Belonging in a UK physics department. In *Academic Belonging in Higher Education*, E. Rueda and C. Lowe Swift (Eds.), 2024, 80-94

[32] Lim, L.-A., Buckingham-Shum, S., Felten, P. and Uno, J. "Belonging analytics": A Proposal. *Learning Letters*, 1 (2023), 1-12.

[33] McHugh, M. L. Interrater reliability: the kappa statistic. *Biochemia medica*, 22, 3 (2012), 276-282.

[34] Meehan, C. and Howells, K. In search of the feeling of 'belonging' in higher education: undergraduate students transition into higher education. *Journal of further and higher education*, 43, 10 (2019), 1376-1390.

[35] Museus, S. D., Yi, V. and Saelua, N. The Impact of Culturally Engaging Campus Environments on Sense of Belonging. *Review of higher education*, 40, 2 (2017), 187-215.

[36] Nichols, S. L. and Dawson, H. S. Assessment as a Context for Student Engagement. In *Handbook on Research on Student Engagement*, S. L. Christenson, A. L. Reschly and C. Wylie (Eds.), 2012, 457-477

[37] Pekrun, R. and Linnenbrink-Garcia, L. Academic Emotions and Student Engagement. In *Handbook on Research on Student Engagement*, S. L. Christenson, A. L. Reschly and C. Wylie (Eds.), 2002, 259-282

[38] Pintrich, P. R. and Zusho, A. Student Motivation and Self-Regulated Learning in the College Classroom. In *The Scholarship of Teaching and Learning in Higher Education: An Evidence-Based Perspective.*, R. P. Perry and J. C. Smart (Eds.), 2000, 731-810

[39] Rahman, M. M. and Watanobe, Y. ChatGPT for Education and Research: Opportunities, Threats, and Strategies. *Applied sciences*, 13, 9 (2023), 5783.

[40] Ramanathan, S., Buckingham Shum, S. and Lim, L.-A. To what extent do responses to a single survey question provide insights into students' sense of belonging? In *Proceedings of the Learning Analytics Knowledge Conference* (Kyoto, Japan, 2024). ACM,878-884

[41] Rao, V. N., Agarwal, E., Dalal, S., Calacci, D. and Monroy-Hernández, A. QuaLLM: An LLM-based Framework to Extract Quantitative Insights from Online Forums (2024).

[42] Rasheed, Z., Waseem, M., Ahmad, A., Kai-Kristian, K., Wang, X., Anh Nguyen, D. and Abrahamsson, P. *Can Large Language Models Serve as Data Analysts? A Multi-Agent Assisted Approach for Qualitative Data Analysis*. arxiv.2402.01386, 2024.

[43] Reeve, J. A Self-determination Theory Perspective on Student Engagement. In *Handbook on Research on Student Engagement*, S. L. Christenson, A. L. Reschly and C. Wylie (Eds.), 2012, 149-172

[44] Reschly, A. L. and Christenson, S. L. *Handbook of Research on Student Engagement*. Springer International Publishing AG, New York, 2022.

[45] Rrv, A., Tyagi, N., Uddin, M. N., Varshney, N. and Baral, C. Chaos with Keywords: Exposing Large Language Models Sycophancy to Misleading Keywords and Evaluating Defense Strategies (2024).

[46] Salah, M., Al Halbusi, H. and Abdelfattah, F. May the force of text data analysis be with you: Unleashing the power of generative AI for social psychology research. *Computers in Human Behavior: Artificial Humans*, 1, 2 (2023), 100006.

[47] Schunk, D. H. and Mullen, C. A. Self-Efficacy as an Engaged Learner. In *Handbook of Research on Student Engagement*, S. L. Christenson, A. L. Reschly and C. Wylie (Eds.), 2012, 219-235

[48] Shah, C. *From Prompt Engineering to Prompt Science With Human in the Loop*. arxiv.2401.04122, 2024.

[49] Skinner, E. A., Kindermann, T. A. and Furrer, C. J. A Motivational Perspective on Engagement and Disaffection: Conceptualization and Assessment of Children's Behavioral and Emotional Participation in Academic Activities in the Classroom.

Educational and psychological measurement, 69, 3 (2009), 493-525.

[50] Skinner, E. A. and Pitzer, J. R. Developmental Dynamics of Student Engagement, Coping, and Everyday Resilience. In *Handbook of Research on Student Engagement*, S. L. Christenson and A. L. Reschly (Eds.), 2022, 21-44

[51] Strayhorn, T. L. *College Students' Sense of Belonging: A Key to Educational Success for All Students*. Taylor and Francis, Milton, 2018.

[52] Tabone, W. and de Winter, J. Using ChatGPT for human–computer interaction research: a primer. *Royal Society open science*, 10, 9 (2023), 231053-231053.

[53] Tai, R. H., Bentley, L. R., Xia, X., Sitt, J. M., Fankhauser, S. C., Chicas-Mosier, A. M. and Monteith, B. G. An Examination of the Use of Large Language Models to Aid Analysis of Textual Data. *International Journal of Qualitative Methods*, 23 (2024).

[54] Thomas, L. Student retention in higher education: the role of institutional habitus. *Journal of education policy*, 17, 4 (2002), 423-442.

[55] Tinto, V. Rethinking the First Year of College. *Higher Education Monograph Series*, 9, 2 (2001), 1-8.

[56] Trowler, V., Allan, R. L., Bryk, J. and Din, R. R. Pathways to student engagement: beyond triggers and mechanisms at the engagement interface. *Higher Education*, 84, 4 (2022), 761-777.

[57] Viola, J. K. Belonging and Global Citizenship in a STEM University. *Education sciences*, 11, 12 (2021), 803.

[58] Voelkl, K. E. School Identification. In *Handbook of Resaerch on Student Engagement*, S. L. Christenson, A. L. Reschly and C. Wylie (Eds.), 2012, 193-218

[59] Wang, B., Min, S., Deng, X., Shen, J., Wu, Y., Zettlemoyer, L. and Sun, H. *Towards understanding chain-of-thought prompting: An empirical study of what matters*. arXiv.2212.10001, 2022.

[60] Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E., Le, Q. and Zhou, D. *Chain-of-Thought Prompting Elicits Reasoning in Large Language Models*. arxiv.2201.11903, 2023.

[61] Wentzel, K. Part III Commentary: Socio-Cultural Contexts, Social Competence, and Engagement at School. In *Handbook on Research on Student Engagement*, S. L. Christenson, A. L. Reschly and C. Wylie (Eds.), 2012, 479-488

[62] Wise, A. F. and Shaffer, D. W. Why Theory Matters More than Ever in the Age of Big Data. *Journal of Learning Analytics*, 2, 2 (2015), 5.

[63] Wise, A. F., Knight, S. and Buckingham Shum, S. Collaborative Learning Analytics. In *International Handbook of Computer-Supported Collaborative Learning*, U. Cress, C. Rosé, A. Wise and J. Oshima (Eds.), 2021, 425-443

[64] Wolters, C. A. and Taylor, D. J. A Self-regulated Learning Perspective on Student Engagement. In *Handbook on Research on Student Engagement*, S. L. Christenson, A. L. Reschly and C. Wylie (Eds.), 2012, 635-651

[65] Xiao, Z., Yuan, X., Liao, Q. V., Abdelghani, R. and Oudeyer, P.-Y. Supporting Qualitative Analysis with Large Language Models: Combining Codebook with GPT-3 for Deductive Coding. In *Proceedings of the Companion proceedings of the 28th international conference on intelligent user interfaces* (2023). ACM,75-78

[66] Xuan Lam, P., Mai, P. Q. H., Nguyen, Q. H., Pham, T., Nguyen, T. H. H. and Nguyen, T. H. Enhancing educational evaluation through predictive student assessment modeling. *Computers and education. Artificial intelligence*, 6 (2024), 100244.

[67] Yan, L., Martinez-Maldonado, R. and Gasevic, D. Generative Artificial Intelligence in Learning Analytics: Contextualising Opportunities and Challenges through the Learning Analytics Cycle. In *In Proceedings of the 14th Learning Analytics and Knowledge Conference*, I. P. o. t. t. L. A. a. K. Conference (Eds.), 2024, 101-111

[68] Yazzie-Mintz, E. and McCormick, K. Finding the Humanity in the Data: Understanding, Measuring, and Strengthening Student Engagement. In *Handbook on Research on Student Engagement*, S. L. Christenson, A. L. Reschly and C. Wylie (Eds.), 2012, 743-761

[69] Zambrano, A. F., Liu, X., Barany, A., Baker, R. S., Kim, J. and Nasiar, N. *From nCoder to ChatGPT: From Automated Coding to Refining Human Coding*. Springer Nature Switzerland, City

Appendix

Final system prompt for deductive coding of student sense of belonging

I am inviting you to participate as a researcher in my research project. The objective is to analyse students' written reflections on their emotional experiences while learning in a HIGHER EDUCATIONAL environment. I have asked the students to share their reflections and feedback about their educational journey.
Affective Engagement is evidenced ONLY if one or more of the following two categories is present:
1. EXPERIENCING EMOTIONS refers to the awareness and recognition of emotional states because of academic and social interactions and exchanges in a university environment. For example, these could be experiencing pride, experiencing sadness, feeling angry, feeling anxious, feeling bored, feeling distressed, feeling happy, feeling hopeful, feeling positive, experiencing satisfaction, and feeling accomplished to name a few.
2. FEELING A SENSE OF BELONGING refers to the feeling of being accepted, valued and included within the university. This has four sub-categories, namely:
2.1 ACADEMIC BELONGING: When the students refer to the importance of acquiring and applying specific knowledge and skills in their field of study, engaging with their coursework, understanding their field and academic achievements.
Example: "I feel a deep sense of accomplishment when I understand complex theories in my major. This validation from my coursework reassures me that I am on the right path in my academic journey."
Example: "I can understand and appreciate the need for proper accounting skills, and just how important it is to any business foundation."
2.2 INTERPERSONAL BELONGING: When the students refer to their desire for a supportive and engaging social environment provided by their university, it is related to interpersonal belonging. This centers on the social aspect of university life, including how students feel connected and valued in their interactions and relationships with others in the university environment
Example: "Joining the study group has been incredibly rewarding. I feel a strong bond with my classmates, and the support from my professor makes me feel like an important part of the academic community."
Example: "I aspire to work at an interactive social environment that is provided in my job to be able to meet new people and make more friends."
2.3 BELONGING TO SURROUNDINGS: When the students refer to their comfort and connection with the physical, digital spaces and cultural environment of the university, it is related to belonging to surroundings.
Example: "The new library's modern design and comfortable study areas make me feel at home on campus. I appreciate how the university's cultural events reflect a diverse and inclusive environment that resonates with me."
Example: "The quiet study nook I have in the library feels like my own personal retreat. Having a dedicated space where I can focus and organize my thoughts helps me manage my academic workload effectively."
2.4 MATTERING AS BELONGING: When the students refer to their sense of ownership, comfort, and security in their learning environment, it is related to mattering of belonging. This involves students feeling about their personal space and resources within the learning environment, and how these contribute to their academic achievements.
Example: "the ideating exercise in week 2 helped me connect many of the goals I didn't realise I had, forming one more concise idea of what my ideal future looks like"
Example: "The quiet study nook I have in the library feels like my own personal retreat. Having a dedicated space where I can focus and organize my thoughts helps me in managing my academic workload effectively"
After you have identified these categories and any subcategories within the students' reflections, please determine which indicators belong to each subcategory according to the following coding scheme:
A Sense of Belonging with the Sub-categories and Indicators being:
- Academic Belonging Codes: expressed as valuing learning, showing complete involvement in studying, experiencing opportunity to learn and achieve goals, feeling self-efficacy, feeling independent, expressing motivation to study, feeling interested, showing curiosity, showing enthusiasm, showing interest, feeling challenged, feeling committed, feeling inquisitive
- Interpersonal Belonging Codes: expressed as feeling valued, feeling cared for, feeling included, feeling supported, feeling important, feeling accepted, feeling connected, interpersonal interactions, experiencing social connections,
- Belonging to Surroundings Codes: expressed as appreciation of the learning environment, cultural space, local surroundings, locality,
- Mattering as Belonging Codes: expressed as personal identity, personal interest, recognition, self-confidence
Your main task is to read each reflection and recognize the affective engagement, subcategories and indicators for each subcategory used by this student. For each identified affective engagement, please provide:
- Name of the affective engagement, subcategory and indicators,
- The part of the text where you found this affective engagement, subcategories and indicators,
- Reasoning behind selecting the given affective engagement, subcategories and indicators.
Here are some instructions that I want you to follow:
It's important to note that a single sentence in a student's reflection may reflect multiple affective engagement categories, indicating the complexity and interconnectedness of engagement. I want you to include all the engagement categories. Please focus solely on direct evidence of each category.