WHEN DEFINITION MEETS PROCEDURE: THE ALIGNMENT BETWEEN STUDENTS' DEFINITIONS OF EQUIVALENCE AND THEIR ALGEBRAIC PRACTICE

Mela Aziza[1,2] & Chris Sangwin[2]

[1]Universitas Islam Negeri Fatmawati Sukarno, [2]The University of Edinburgh

Robust algebraic reasoning depends on recognising when expressions or equations are equivalent and justifying that judgement. This study investigates students' definitions and their actual justification strategies, and the alignment between them. Open written responses from 13 first-year undergraduate mathematics students were analysed thematically, both in a traditional manner and using an LLM. The tasks required definitions, examples of equivalent or non-equivalent pairs, and justification for equivalence in expressions and equations. For equivalence of expressions, students primarily provided normative definitions, the same value for every x, whereas in practice, students often relied on transformations or numerical-value checks. Meanwhile, students showed greater alignment in equivalent equations.

INTRODUCTION

Equivalence plays a central role in both school and university mathematics. This concept is taught from the early use of the equal sign (Jones et al., 2012) to the equality relationship (Cook et al., 2022; Wladis et al., 2022). Equivalence has two distinct topic domains: the equivalence of expressions and equations, which is key to advanced algebraic reasoning (Cook et al., 2022). Understanding these two domains entails not only providing definitions but also applying them in practice (Wladis et al., 2022). "Practice" refers to the concrete process of justifying that two expressions are identical under all possible variable values or that two equations have the same solution set (Cook et al., 2022). Aligning definitions and practice matters for first-year university students, transitioning to higher education as they start to face greater demands on proof, use of formal mathematics, and a wider range of mathematical objects (Alcock & Simpson, 2017). When the definition does not move into practice, apparent understanding can be misleading, as practical evidence may point in the opposite direction (Vinner, 1991). Therefore, studying how definitions and practices align can inform learning, assessment, and teaching strategies in higher education.

EMPIRICAL AND THEORETICAL BACKGROUND

A previous study found that although students can provide formal definitions, they often tend to perform procedures (e.g., testing specific values or manipulating symbols) when justifying or proving mathematical claims (Knuth et al., 2006). In the context of the transition to university mathematics, the main difficulties are not simply memorising definitions but aligning definitions with proof practice, using

2026. In H. Krzywacki, M. Hähkiöniemi, E. Haataja, M.S. Hannula & P. Hästö (Eds.). *Proceedings of the 49th Conference of the International Group for the Psychology of Mathematics Education: Research Reports*, pp. 107-114. PME.

quantification explicitly, and choosing the appropriate type of justification (Alcock & Simpson, 2017). However, studies investigating the alignment between students' justifications in practices and their stated normative definitions of expression and equation equivalence at the university level remain relatively limited. Key issues, including the specific conditions that ensure operations preserve equivalence, such as valid operations and circumstances in which transformations result in equivalent or non-equivalent statements, are frequently not explicitly discussed in the literature (Knuth et al., 2006).

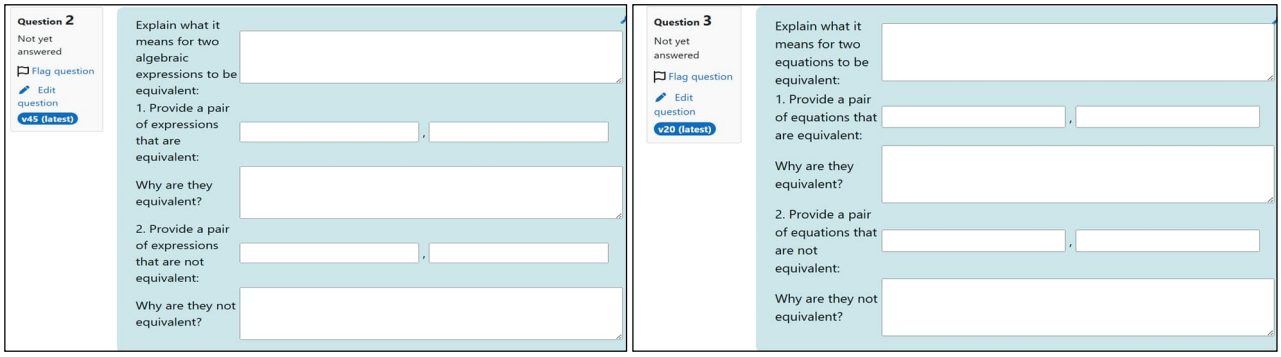
Several theoretical frameworks clarify the alignment between definitions and practices. The concept-definition (formal concepts) and concept-image (actions involved in exemplifying or justifying claims) framework of Tall and Vinner (1981) explains why students can state formal definitions but proceed according to procedural habits or their concept images. A concept definition is an accurate verbal explanation that normatively defines a concept, such as the definition of equivalence. In contrast, a concept image means an individual's complete mental images of a concept, including all associated mental pictures, personal examples, intuitive understandings, and procedures (Tall & Vinner, 1981; Vinner, 1991). Normatively, students are expected not only to state "have the same value for any possible value of x" for expression equivalence, but also to prove these identities universally; for equation equivalence, students are expected to reason about solution sets and use transformations shown to preserve solution equivalence. Similarly, the operational-structural distinction of mathematical objects (Sfard, 1991) helps explain why many students understand the equals sign operationally (results in an answer) rather than relationally (has the same value), which then influences their justification practices; often involving calculations or numerical examples, rather than general structural transformations (Knuth et al., 2006). These empirical and theoretical views motivate us to investigate and assess the alignment between the formal definition of equivalence and students' justification practices for equivalence and non-equivalence expressions and equations. To what extent do students' definitions of equivalence align with their practices for justifying the equivalence of algebraic expressions and equations?

METHODS

We developed an open-ended online questionnaire through four steps adapted from Daro and Burkhard (2012) and Melhuish (2018): (1) determining acceptance criteria; (2) finding literature and collecting potential questions; (3) classifying and sorting the questions; and (4) formulating open-ended questions. After conducting steps (1), (2), and (3), we obtained eight questions on five sub-concept groups. However, we decided not to develop questions on "the meaning of equality" and "the meaning of an equation" because they have been well studied, e.g., by Godfrey and Thomas (2008) and Steinberg et al. (1991). We still developed questions related to "the meaning of the equal sign" because of its importance for understanding other concepts in mathematics. Thus, we formulated seven open-ended questions on three sub-concepts: (1) the

meaning of the equal sign, (2) algebraic equivalence involving the equivalence of expressions and equations, and (3) transformation equivalence by adapting existing questions and creating new questions that prior research has not addressed. In this paper, we report students' responses to sub-concept (2), comprising only Questions 2 and 3 (Figure 1).

Figure 1: Questions 2 and 3 (equivalence of expressions and equations)



We piloted the questionnaire with PhD students in mathematics and then invited first-year undergraduate mathematics students to complete it. Our goal was to identify common response types and their links to underlying student conceptions, to recognise recurring patterns, to categorise responses based on conceptual understanding, and to highlight misconceptions. We analysed the data thematically, both manually and using a Large Language Model (LLM), ChatGPT-5. This combined approach (manual and LLM) enables exploration of textual patterns that might be missed by a single method.

We first traditionally analysed the data, using a manual approach and applying three steps of thematic analysis: (1) data engagement, coding, and initial theme identification; (2) theme development, refinement, and definition; and (3) thematic reporting and interpretation (Braun & Clarke, 2006). After completing these steps independently, we used a large-language model-assisted analysis (ChatGPT) as a third-party tool, clearly separating the traditional and LLM-assisted phases. We structured prompts using four parts in the framework of Zhang et al. (2025): (1) giving the information on the analysis aim, the questions asked, and the students' answers (task background); (2) explaining the instructions and goal of the task (task description); (3) explaining three thematic analysis steps (methodology); and (4) listing expected output content and format. Throughout this stage, the use of ChatGPT5 was thoroughly documented in the research notes, including the prompt, assigned task, and the manner in which the output was utilised. The added value of using LLM-assisted analysis lies in its time-saving, detailed, and readily implementable outputs. However, the LLM analysis was not intended to replace manual analysis. We maintain and adhere to human verification principles. The LLM helped with initial coding, while we made interpretive decisions. We identified and addressed potential bias through systematic verification and validation procedures to ensure coding and key themes were

consistent, complete, and resolved through manual, collaborative adjudication. Both researchers reviewed the large language model output and, through consensus discussions, established the final codes, theme definitions, and findings.

RESULTS AND DISCUSSION

We report here key findings on how students define equivalence in expressions and equations, how they justify the equivalence and non-equivalence of expressions and equations, and the extent to which their practices align with or differ from their definitions. Following this, 13 students responded to Question 2 (equivalent expressions) and Question 3 (equivalent equations). Subsequently, after manually analysing the data, we compared the codes and key themes with those produced by the LLM. To illustrate this comparison, an example of the analyses conducted by researchers and ChatGPT-5 is shown in Table 1.

Table 1: The example of student excerpts and Human–LLM coding comparison

Question	Student response excerpt	Code	Theme	Comparison
Q2 (Expression)	Two algebraic expressions are equivalent if they have the same value for all values of the variables.	Manual: Value-based/Eval-vars LLM: Universal-value/Identity definition	Manual: value-based equivalence LLM: formal identity	The coding is highly similar; manual work provides numeric frequency counts, whereas ChatGPT organises and interprets the data but does not replace manual numeric labelling.
Q3 (Equation)	Equations are equivalent when they have exactly the same solutions.	Manual: same solution/solutions LLM: same solution set	Manual: solution-based equivalence LLM: solution-set equivalence	

Based on the comparison, we observed that the LLM analyses could strengthen the validity of the researcher's manual coding, e.g., value-based, solution-set, transformational, vague, commutativity conflation, expression–equation conflation, and reliance on single-value checks. However, the LLM analysis could not identify an explicit alignment between definitions and practices. Therefore, researchers conducted a collaborative discussion to compare manual and LLM analyses and then adjudicated the final report and interpretation of the data. Following the discussion, we obtained a

summary of findings on the misalignment between students' definitions and practices of equivalence, as shown in Table 2.

Table 2: The alignment between students' definitions and justifications of expression and equation equivalence

Domain topic	Definition	Justification	Definition-practice alignment
Expression	Predominant: identity (same value for any possible values of the variable). Alternatives: transformational, some confusion with equations.	Equivalent: identity-based or transformational arguments (e.g., expansion/factorisation). Non-equivalent: counterexample; sometimes "cannot be manipulated into the same form."	Tends not to be fully aligned: normative definition but procedural practice; risk of overlooking domain/quantification.
Equation	Predominant: same solution set. Alternative: transformation-based by reversible operations.	Equivalent: demonstrate or preserve the solution set. Non-equivalent: different solution sets.	More aligned: both definition and practice are solution-based.

Based on Table 2 above, at the definitional level of two equivalent expressions, many students provided normative definitions that align with mathematical norms: "has the same value for any possible values of the variable." Some other responses emphasise a transformational perspective, such as "can be simplified/transformed into the same form," or emphasise "interchangeability/commutative" and "same information". A minor misconception arises when "equivalent expressions" are understood as "equivalent equations," and when reliance on a single-value substitution to establish "numerically equal". For equation equivalence, most students define it as "having the same solution set", whereas a minority provide a procedural definition: two equations are equivalent if they can be transformed into one another by valid (reversible) algebraic operations. Some responses confuse expressions and equations; for example, they use the criterion "the same value for all possible values of x" for equation equivalence, which actually belongs to expression equivalence. Based on theories, the definitions "identity for any possible values of variables" (expression equivalence) and "the same solution set" (equation equivalence) are mathematically normative concept definitions (Cook et al., 2022). Meanwhile, the transformational definition emphasises

a concept image of "what is usually done" in algebraic manipulation rather than explicitly articulating universal quantification or solution sets (Sfard, 1991; Tall & Vinner, 1981).

When justifying the equivalence of two expressions, students provided two main justifications: (1) value-based arguments, which assert that two expressions have equal values for any substitution of the variable and (2) transformational arguments, which demonstrate that one expression can be transformed into another through algebra (e.g., $2(x + 4)$ and $2x + 8$). Transformational justifications reflect procedural knowledge or operational conception. Therefore, the justification must be supported by explicit statements about the domain and universal quantification (Sfard, 1991). Some justifications also encounter procedural flaws (e.g., misdistribution, misplaced signs) or present examples that are actually equations rather than expressions. For justifying the non-equivalence of expressions, the most powerful strategy is the counterexample: choosing a single value of the variable that shows that the two expressions produce different values. This strategy can effectively disprove the claim that two expressions are "the same for any substitutions of the variable". However, another argument that an expression "cannot be rewritten to another" is weak unless supported by specifying the domain or providing concrete counterexamples. Students justified two equivalent equations by solving both equations and showing whether their solutions are equal, or by applying an operation to both sides that preserves or changes the solution sets. The practice of emphasising "solution sets" indicates the relational meaning of the equal sign (Knuth et al., 2006), while the use of reversible operations demonstrates a connection between conceptual and procedural knowledge (Baroody et al., 2007).

The alignment between definition and practice differs between expressions and equations. To illustrate this difference more closely, in expressions, most students defined equivalence of expressions as having the same value for any variable; some provided transformational justifications, such as expansion or factorisation. This misalignment highlights the gap between concept definition (having the same value for any substitution of the variable) and concept image (the habit of manipulating symbols procedurally) (Tall & Vinner, 1981). Students often did not explain why the transformations preserve the value for any possible x. This approach is mathematically valid. Nevertheless, without a clear argument that the expressions have the same value for all substitutions and without considering the domain, the reasoning becomes weak and risks procedural errors. Therefore, students require a strong conceptual understanding that connects "knowing what" and "knowing why/how" (Baroody et al., 2007). This finding also indicates that students tend to begin with operational conceptions before forming structural conceptions (Sfard, 1991). In contrast, for equations, there is a closer alignment between students' definitions of equivalence and their actual justifications. Those who describe equation equivalence as "having the same solution set" typically verify the same solutions of the equivalent equations by solving both equations or transforming one into another using solution-set-preserving operations. Hence, the definition–practice misalignment is stronger in Q2 than in Q3,

indicating that students need a more comprehensive understanding to ensure that two expressions are equivalent.

We offer two principal contributions. *First*, we systematically map two topic domains of equivalence (expressions and equations) and an improved typology of misalignment that directly connects students' definitional statements to their practical justifications. *Second*, we integrate manual qualitative analysis with analysis assisted by a large language model (ChatGPT). This approach helps address consistent themes, highlight ambiguous cases, and accelerate the identification of different misconceptions without compromising the researcher's critical judgment. Regarding concept definition and concept image (Tall & Vinner, 1981; Vinner, 1991), structural-operational conceptions (Sfard, 1991) and conceptual-procedural knowledge (Baroody et al., 2007), these findings imply that learning instructions need to cover not only "what is a definition" but also "why and when" transformations preserve equivalence, and "how" quantifications are appropriately stated and verified. Thus, we argue that alignment between definition and practice is a key indicator of readiness for higher-level algebra, particularly when entering university.

CONCLUSION AND IMPLICATIONS

This study found that most participants could state the normative definition of equivalence: "having the same value for all possible substitutions of the variable" for expressions and "the same solution set" for equations. However, students tended to justify equivalent expressions by algebraic transformations and evaluating a specific value that did not satisfy the universal criterion required by the definition. The misalignment between the definition and practice of equation equivalence is less evident than in expressions, suggesting that students need a better understanding of the universal identity in equation equivalence. These results are consistent with the literature on concept definition-concept image, conceptual-procedural understanding, and structural-operational conceptions. Combining manual coding with LLM-assisted analysis proved effective for mapping and investigating this misalignment.

For learning and assessment in mathematics, our study recommends explicitly distinguishing between expression equivalence and equation, and emphasising the connection between definitional knowledge and practical justifications. Using an LLM such as ChatGPT can contribute to thematic analysis, but human verification remains necessary to maintain the integrity of the findings. The primary limitation of this study is its small participant size, which suggests that future research should include larger, more diverse participant groups.

References

Alcock, L., & Simpson, A. (2017). Interactions between defining, explaining and classifying: The case of increasing and decreasing sequences. *Educational Studies in Mathematics, 94*(1), 5–19. https://doi.org/10.1007/s10649-016-9709-4

Baroody, A. J., Feil, Y., & Johnson, A. R. (2007). Research commentary: An alternative reconceptualization of procedural and conceptual knowledge. *Journal for research in mathematics education*, *38*(2), 115–131. https://doi.org/10.2307/30034952

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Cook, J. P., Reed, Z., & Lockwood, E. (2022). An initial framework for analyzing students' reasoning with equivalence across mathematical domains. *The Journal of Mathematical Behavior*, *66*, 100935. https://doi.org/10.1016/j.jmathb.2022.100935

Daro, P., & Burkhard, H. (2012). A population of assessment tasks. *Journal of Mathematics Education at Teachers College*, *3*(1), 19–25. https://doi.org/10.7916/jmetc.v3i1.732

Godfrey, D., & Thomas, M. O. J. (2008). Student perspectives on equation: The transition from school to university. *Mathematics Education Research Journal*, *20*(2), 71–92. https://doi.org/10.1007/BF03217478

Jones, I., Inglis, M., Gilmore, C., & Dowens, M. (2012). Substitution and sameness: Two components of a relational conception of the equals sign. *Journal of experimental child psychology*, *113* (1), 166–176. https://doi.org/10.1016/j.jecp.2012.05.003.

Knuth, E. J., Stephens, A. C., McNeil, N. M., & Alibali, M. W. (2006). Does understanding the equal sign matter? evidence from solving equations. *Journal for research in Mathematics Education*, *37*(4), 297–312. https://doi.org/10.2307/30034852

Melhuish, K. (2018). Three conceptual replication studies in group theory. *Journal for Research in Mathematics Education*, *49*(1), 9–38. https://doi.org/10.5951/jresematheduc.49.1.00095

Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. *Educational Studies in Mathematics*, *22*(1), 1–36. https://doi.org/10.1007/BF00302715

Steinberg, R. M., Sleeman, D., & Ktorza, D. (1991). Algebra students' knowledge of equivalence of equations. *Journal for research in mathematics education*, *22*(2), 112–121. https://doi.org/10.2307/749588

Tall, D. O., & Vinner, S. (1981). Concept image and concept definition in mathematics, with special reference to limits and continuity. *Educational Studies in Mathematics*, *12*(2), 151–169. https://doi.org/10.1007/BF00305619

Vinner, S. (1991). The role of definitions in the teaching and learning of mathematics. In *Advanced mathematical thinking* (pp. 65–81). Dordrecht: Springer.

Wladis, C., Sencindiver, B., Offenholley, K., Jaffe, E., & Taton, J. (2022). Student definitions of equivalence: Structural vs. operational conceptions, and extracted vs. stipulated definition construction. In *Twelfth Congress of the European Society for Research in Mathematics Education (CERME12)* (No. 26).

Zhang, H., Wu, C., Xie, J., Lyu, Y., Cai, J., & Carroll, J. M. (2025). Harnessing the power of AI in qualitative research: Exploring, using and redesigning ChatGPT. *Computers in Human Behavior: Artificial Humans*, *4*, 100144.