MODESTUM EURASIA Journal of Mathematics, Science and Technology Education, 2026, 22(12), em2964
ISSN:1305-8223 (online)
OPEN ACCESS Research Paper https://doi.org/10.29333/ejmste/19919

Cognitive and dialogic patterns shaping strategic flexibility in collaborative mathematical problem-solving

Haydeé Ceballos [1,2*], Jeroen Spandaw [3], Theo van den Bogaart [2], Stan van Ginkel [2], Paul Drijvers [1]

[1] Freudenthal Institute, Utrecht University, Utrecht, THE NETHERLANDS
[2] HU University of Applied Sciences Utrecht, Utrecht, THE NETHERLANDS
[3] EWI, Delft University of Technology, Delft, THE NETHERLANDS

Received 17 July 2026 ▪ Accepted 02 November 2026

Abstract

Little is known about strategic flexibility in collaborative mathematical problem-solving. In this study, we investigate how cognitive levels and dialogic moves, together with their collective uptake, relate to strategic flexibility. To do so, we combine the Florida taxonomy of cognitive behavior (FTCB) with the teacher scheme for educational dialogue analysis (T-SEDA) to analyze a 29-minute dialogue among four pre-service mathematics teachers solving a geometry problem. Most dialogue contributions involved lower-order thinking and surface-level dialogic moves. The combined framework distinguished actual strategy shifts, unrealized opportunities for strategic reconsideration, and local checking-and-correction moments. Opportunities to reconsider strategy arose when higher-order contributions coincided with making reasoning explicit or challenging a proposal, but were infrequent, largely steered by one participant, and rarely taken up collectively. From this, we conclude that strategic flexibility may benefit from the collective uptake of cognitively demanding contributions. Methodologically, integrating FTCB and T-SEDA traces how cognitive depth and dialogic function relate to such uptake.

Keywords: cognitive processes, collaborative mathematical problem-solving, dialogic interaction, qualitative interaction analysis, strategic flexibility

INTRODUCTION

Collaborative problem-solving (CPS) involves students in expressing ideas, negotiating meaning, and coordinating decisions on how to proceed. Research on CPS in mathematics shows that such interaction does not automatically lead to productive mathematical thinking: participation may be uneven, authority may remain concentrated, and promising ideas may not be taken up or developed further.

One issue that becomes especially relevant in collaborative settings is how groups respond when their current approach is no longer productive. A group may continue with an unproductive line of reasoning, but it may also revise its approach. A useful construct for understanding this adaptive capacity is *strategic flexibility*, commonly defined as the ability to select, apply, and revise strategies as task demands evolve (Star & Rittle-Johnson, 2008; Verschaffel et al., 2011). Although strategic flexibility has mainly been studied at the individual level, for instance as switching strategies within or across problems (Bolat & Arslan, 2024; Hong et al., 2023; Verschaffel, 2024), little is known about how it becomes visible in collaborative settings, where strategy decisions are negotiated through interaction and shaped by group dynamics such as authority relations and the distribution of participation (Langer-Osuna, 2016; Wu, 2024).

The central knowledge gap addressed in this study, therefore, concerns this interplay between strategic flexibility and group dynamic features. Prior research on CPS in mathematics has shown that productive collaboration depends not only on participation and idea exchange, but also on whether cognitively demanding

© 2026 by the authors; licensee Modestum. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/).
✉ m.h.ceballos@uu.nl **(*Correspondence)** ✉ J.G.Spandaw@tudelft.nl ✉ theo.vandenbogaart@hu.nl ✉ stan.vanginkel@hu.nl ✉ p.drijvers@uu.nl

Contribution to the literature

- This study proposes a conceptualization of strategic flexibility in mathematics, long studied as an individual cognitive capability, as an interactional and epistemic process that relates to the moment-to-moment uptake of cognitively demanding contributions in group dialogue.
- It distinguishes three outcomes of moments of strategic reconsideration – actual strategy shifts, unrealized opportunities, and local checking-and-correction moments – and shows that cognitively demanding contributions create openings for strategy shifts, but that these develop into shared strategic flexibility only when collectively examined and extended; they may remain unrealized when uptake is weak and authority is concentrated in one participant.
- Methodologically, it demonstrates how two established coding frameworks, the Florida taxonomy of cognitive behavior (FTCB) and the teacher scheme for educational dialogue analysis (T-SEDA), can be integrated to trace, turn by turn, how cognitive depth and dialogic function relate to the uptake of contributions in CPS.

contributions are taken up and developed by the group (Ceballos et al., 2026; Hansen & Naalsund, 2026). Studies on authority in small-group mathematical work have further demonstrated that groups often follow the most confident or dominant participant, which can constrain the range of ideas that are collectively examined (Bleiler-Baxter et al., 2023; Ellis & Alzaga Elizondo, 2026; Wu, 2024). Following Hamilton (2026), we understand mathematical authority as a relationship in which one individual views another as a legitimate source of mathematical knowledge or reasoning. As Hamilton (2026) notes, such legitimacy may rest on perceived confidence or status rather than on the soundness of the mathematics.

However, the connection between these interactional dynamics and strategic flexibility has remained largely unexplored. In particular, little is known about how strategy shifts arise in interaction, and what these shifts reveal about the group's capacity for strategic flexibility. Building on individual-level conceptualizations of strategy change (Star & Rittle-Johnson, 2008; Verschaffel et al., 2011), we define a strategy shift here as an observable change in a group's problem-solving approach.

Addressing this gap, the aim of the study presented here is to investigate how cognitive levels and dialogic moves in CPS, together with their collective uptake, relate to strategic flexibility. To do so, we use an integrated analytical lens to analyze a dialogue among four pre-service mathematics teachers solving a geometry problem.

THEORETICAL FRAMEWORK

This section describes four theoretical constructs in relation: (1) strategic flexibility, which we understand as an expression of (2) control that becomes visible in (3) the cognitive level and (4) the dialogic function of students' contributions in group work.

Strategic flexibility refers to the ability to select, apply, and revise strategies as problem demands change (Star & Rittle-Johnson, 2008; Verschaffel et al., 2011). In this study, we distinguish between strategy shifts, understood as observable changes in a group's problem-solving approach, and strategic flexibility, understood as the group's capacity to enable, recognize, and take up such shifts. While this construct is well established at the individual level, less is known about how it becomes visible in CPS, where decisions about strategy use are negotiated through interaction. In this study, strategic flexibility is examined at the group level, not as a fixed trait of individual learners but as an interactional capacity that emerges in the unfolding dialogue as students introduce ideas, respond to one another and create or miss opportunities for strategy shifts. At the group level, selecting, applying, and revising strategies become visible in interaction: selection in whether a proposed approach is taken up or left aside, application in the group's ongoing work within its current approach, and revision in moments in which the group changes its approach or corrects a step without changing its overall approach.

Second, Schoenfeld's (1985) model of control provides a useful lens here. Control refers to the processes through which problem solvers coordinate resources, monitor progress, and decide on strategic actions. Strategic flexibility is understood as one expression of control, concerned specifically with decisions about selecting, adapting, or abandoning strategies. At the group level, strategy shifts are the observable moments in which such decisions result in an actual change of approach. In collaborative settings, these strategic decisions become visible in students' talk as ideas are introduced, taken up, challenged, or overlooked across turns. Interaction therefore offers a vantage point for examining not only when a group changes its strategy, but also the preceding conditions under which such shifts occur or fail to occur.

Because strategic flexibility at the group level depends on the interplay between what is said and how it is taken up, its analysis requires attention to both cognitive and dialogic dimensions of interaction. By

collective uptake we mean the extent to which a contribution is picked up by the group in the turns that follow, that is, examined, built on, or tested by several participants, rather than briefly acknowledged by one participant or left aside (cf. Hansen & Naalsund, 2026, on a peer's input being assimilated, refined, or neglected). Contributions differ in cognitive complexity, from recall to analysis and evaluation (e.g., Bloom, 1956; Webb, 1970), and dialogic processes shape whether ideas are shared, developed, or left unexplored (e.g., Hennessy et al., 2016; Littleton & Mercer, 2013). Together, these dimensions provide insight into how a group's current strategy is maintained, questioned, or refined, and under which conditions interaction leads to an actual strategy shift.

Third, the FTCB (Brown et al., 1966) enables analysis of the cognitive level of students' contributions by operationalizing cognitive processes at the level of observable verbal turns. It is based on Bloom's (1956) taxonomy and distinguishes seven hierarchically ordered levels (knowledge, translation, interpretation, application, analysis, synthesis, and evaluation), each described in terms of observable behaviors, such as detecting an error in thinking or pointing out an unstated assumption at the level of analysis (Webb, 1970). In the present study, *analysis*, *synthesis*, and *evaluation* were retained as separate codes, while the lower levels were grouped as *lower order level*.

Fourth, the T-SEDA (Kershner et al., 2020) complements this by capturing the dialogic functions of those same turns, such as proposing ideas, making reasoning explicit, or challenging peers. It was derived from the scheme for educational dialogue analysis (Hennessy et al., 2016), which is grounded in a sociocultural perspective and codes contributions according to their communicative function, and condenses it into a smaller set of categories focused on dialogue features considered productive for learning (Kershner et al., 2020). The frameworks are commensurable in that both treat publicly observable talk as the primary access point to cognitive and interactional processes. They nevertheless stem from different traditions: FTCB classifies levels of thinking, whereas T-SEDA locates the meaning of a contribution in its function within the dialogue. We therefore treat FTCB codes not as indicators of individual mental states but as characteristics of what is expressed in a turn, interpreted in relation to the surrounding interaction. Used together, the frameworks make visible what neither captures on its own: whether cognitively demanding contributions are taken up in the dialogue, since a turn may reach the level of *analysis* without being built upon, and a dialogically active move may involve little cognitive depth.

As strategic flexibility involves both cognitive processing and interactional negotiation, neither a purely cognitive nor a purely dialogic framework would be sufficient to capture how strategies are jointly maintained, questioned, or revised in interaction. For this study, we therefore developed an innovative integration of these two complementary frameworks. We integrate the two frameworks by assigning each coding unit both an FTCB and a T-SEDA code, and by analyzing the co-occurrence and sequential patterning of these codes. This integrated approach makes it possible to examine how cognitive depth and dialogic function jointly shape opportunities for strategy shifts, and what these moments reveal about a group's strategic flexibility. Accordingly, the question guiding the study is:

How do cognitive levels and dialogic moves in students' interaction, together with their collective uptake, relate to strategic flexibility in collaborative mathematical problem-solving?

This main question is specified into two research questions (RQs).

RQ1. What cognitive and dialogic features, as captured through the FTCB and T-SEDA frameworks, characterize the interaction during collaborative mathematical problem-solving?

RQ2. How do cognitive levels and dialogic moves, together with their collective uptake, characterize actual strategy shifts and unrealized opportunities for such shifts?

METHODS

Research Design

In line with Flyvbjerg's (2011) and Bakker's (2018) conception of analytical generalization, this study adopts an instrumental single-case study design (Stake, 1995): the case of a single group's collaborative work is examined not for its own sake, but as a means of understanding a broader phenomenon, namely how strategic flexibility is enabled or constrained in interaction. This case was information-rich for this purpose: the students verbalized their reasoning extensively, and during the session the group attempted several approaches, made and detected errors, and carried out an actual strategy shift. This made it possible to examine how opportunities for strategy shifts arose, were taken up, or remained unrealized in interaction.

Analytically, the case was examined through interaction analysis (Jordan & Henderson, 1995), using a microanalysis that traces cognitive and dialogic contributions turn by turn. The aim is accordingly not to generalize from a sample, but to identify a mechanism that may operate across comparable collaborative settings. Such a design is well suited to examining complex learning processes in mathematics education (Ghic, 2012).

To address the RQs, the analysis proceeded in two stages. First, the entire interaction was examined to characterize the general cognitive and dialogic patterns across the session. Second, selected excerpts were analyzed more closely to explore how combinations and sequences of cognitive and dialogic contributions coincided with moments in which the group reconsidered, refined, or shifted its strategy. The study draws on two data sources: a 29-minute video recording of the group work, together with a verbatim transcript, and the students' written work produced during the session: their sketches, diagrams, and calculations. The written work was analyzed alongside the talk, since the group's strategic moves were frequently enacted on paper as well as in speech.

Participants and context

Four first-year pre-service mathematics teachers (pseudonyms P, Q, R, and S) from the part-time bachelor of mathematics teacher education program at a university of applied sciences in the Netherlands participated in this study. They were adult students, many of whom had completed another degree before deciding to become a teacher. They volunteered to take part in a recorded problem-solving session. The students knew each other well, had frequently worked together in the program's mathematics courses, and had previously completed a course on mathematical problem-solving in which heuristic approaches were practiced. For the study, the group was asked to work collaboratively on a geometry problem and to make their thinking explicit while working. This instruction was intended to elicit rich verbal interaction so that their reasoning and strategic decisions could be analyzed in depth. The session took place in a quiet room at the university, without the teacher or researcher present, to minimize external influence. All participants provided informed consent for video recording and for the use of anonymized excerpts in research dissemination.

Based on the first author's impressions from being their lecturer, the students were hard-working and critical, and they appeared to feel safe voicing their ideas to one another. Participant P was regarded as the strongest student in mathematics in the group and was respected by the other group members; the other three students differed in mathematical achievement level. P thus already held a recognized position in the group before the session, which is relevant for interpreting the asymmetries in uptake described in the Results.

Task and setting

The group worked on a geometry problem from the national mathematics Olympiad training. Geometry was chosen because it lends itself particularly well to examining strategic flexibility: problems in this domain can often be solved in multiple ways and require evaluation of alternative approaches. Solving this task requires only secondary school mathematics, such as Pythagoras' theorem and properties of circles, with which the students, as first-year pre-service mathematics teachers, were already familiar. In the meantime, the task is non-routine, as no solution method is suggested and students must devise and evaluate an approach themselves. This made the task challenging but accessible for first-year pre-service teachers who had completed a course on problem-solving. To illustrate the range of possible approaches, we present three different solution strategies below, based on Pythagoras' theorem, Thales' theorem and trigonometry.

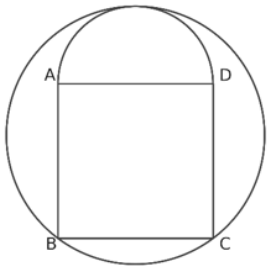


Figure 1. Geometry problem (the authors' own elaboration, redrawn from Stichting Nederlandse Wiskunde Olympiade, 2022)

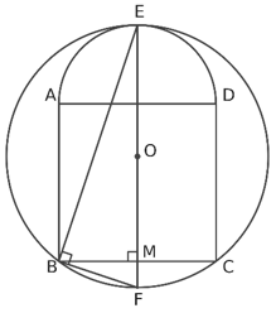


Figure 2. Auxiliary lines illustrating the Thales approach (the authors' own elaboration)

The problem (Stichting Nederlandse Wiskunde Olympiade, 2022): ***Figure 1*** *consists of a square ABCD and a semicircle with diameter AD outside the square. The sides of the square have length 1. What is the radius of the figure's circumscribed circle?*

Possible solution strategies

Perhaps the most obvious solution idea is to apply Pythagoras' theorem: once the position of the center of the circumscribed circle has been determined, right triangles can be formed with the center and the points where this circle meets the figure, in which the radius is related to the known lengths. Alternatively, the problem can be solved using Thales' theorem. Let O be the center of the circumscribed circle, E the point of tangency with the semicircle, and M the midpoint of BC. Extend EM beyond M until it meets the circumscribed circle again at F (**Figure 2**).

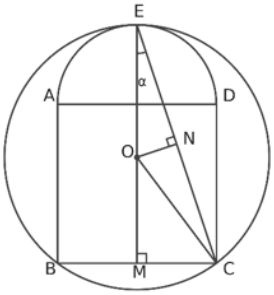


Figure 3. Auxiliary lines illustrating the trigonometry approach (the authors' own elaboration)

Since EF passes through the center O, it is a diameter, so Thales' theorem implies that $\angle \mathrm{EBF} = 90°$. In the right triangle EBF, BM is the altitude to the hypotenuse, so triangles EMB and BMF are similar. It follows that $|BM|^2 = |EM| \cdot |MF|$. Substituting $|BM| = \frac{1}{2}$ and $|EM| = \frac{3}{2}$ gives $\frac{1}{4} = \frac{3}{2}\,|MF|$, so $|MF| = \frac{1}{6}$. Therefore, the diameter is $|EF| = \frac{3}{2} + \frac{1}{6} = \frac{5}{3}$, which yields the radius $\frac{5}{6}$. Unlike the first solution, this approach does not require an equation in the unknown radius.

The problem can also be solved using trigonometry (**Figure 3**).

Let $\alpha = \angle MEC$. In the right triangle EMC, $|EM| = \frac{3}{2}$ and $|MC| = \frac{1}{2}$, so Pythagoras' theorem gives $|EC| = \frac{\sqrt{10}}{2}$ and hence $\cos\alpha = \frac{|EM|}{|EC|} = \frac{3}{\sqrt{10}}$. Since $|OE| = |OC|$, triangle OEC is isosceles, so the perpendicular from O to EC meets EC at its midpoint N. In the right triangle ENO, $\cos\alpha = \frac{|EN|}{|OE|}$, so $|OE| = \frac{|EC|}{2} \div \cos\alpha = \frac{\sqrt{10}}{4} \div \frac{3}{\sqrt{10}} = \frac{5}{6}$.

Key decision points in the different approaches and possible misconceptions

Each approach involves several decisions and may invite various misconceptions. Below, we describe those we consider most relevant to the task and note where they surfaced in the episodes analyzed. All three approaches rely on recognizing that, by symmetry, the center *O* lies on *EM* and that the circumscribed circle passes through *B, C* and *E*. A likely misconception at this stage is to assume that the center of the circle coincides with the center of the square, or to misjudge how the unknown segments on the axis relate to the known lengths (Episode 5). In the Pythagorean approach, a key decision point is to express the unknown segments in terms of the radius, so that a single unknown remains; this route did not surface in the episodes analyzed. In the Thales approach, a key decision point is to extend *EM* to a diameter; a likely misconception is to apply the theorem to a triangle whose side is a radius rather than a diameter, or to place the right angle at the wrong vertex (Episode 1). In the trigonometric approach, a key decision point is to recognize that the required angle can be derived from the given lengths; a likely misconception is that trigonometry cannot be used because no angles are given (Episode 6). These decision points mark where a change of strategy could open up a more productive route, and they serve as a benchmark for interpreting the episodes in the results.

Data and Data Analysis

Data consist of a 29-minute video recording of the group work. The recording was segmented into 172 coding units. A coding unit consists of talk by a single speaker with a distinguishable cognitive and/or dialogic function. A coding unit typically corresponds to one sentence or a short utterance. A single speaker's turn can contain multiple coding units when it involves more than one mathematically or interactionally distinct step. This level of segmentation was chosen because it allowed us to capture distinct cognitive or dialogic functions within a single speaker turn, for instance, when a student first reported a result and then evaluated it, while avoiding the loss of context that would result from coding at the level of individual words or clauses. Minimal acknowledgements or simple feedback, such as "mhm," "mmm," or basic "yes/no" responses, were not segmented as separate coding units when they served only to signal listening, attention, or continuation of the talk. They were segmented separately only when they carried a clear dialogic function, for example by expressing agreement, disagreement, or inviting clarification. Because a single speaker turn could contain more than one coding unit, the subsequent unit in the sequential analysis could be produced either by the same student or by a different student.

Coding the entire dialogue (all 172 coding units) served to address **RQ1** by providing an overview of the general cognitive and dialogic patterns across the session. To address **RQ2**, selected excerpts were analyzed qualitatively. These excerpts focused on two types of episodes:

(1) moments in which the group reconsidered, refined, or shifted its problem-solving approach and

(2) moments in which potential or actual errors prompted students to check or adjust their strategy.

These episodes were identified by the first and second author through a joint close listening of the entire recorded dialogue. Such collaborative viewing contributes to challenging the individual analysts' possible biases and to grounding interpretations in the recorded interaction (Jordan & Henderson, 1995). Candidate moments were selected when the interaction showed an observable change in the group's problem-solving approach, a proposal for such a change, or the identification of an error or inconsistency that prompted the group to reconsider its current line of work. Each candidate moment was then classified as an actual

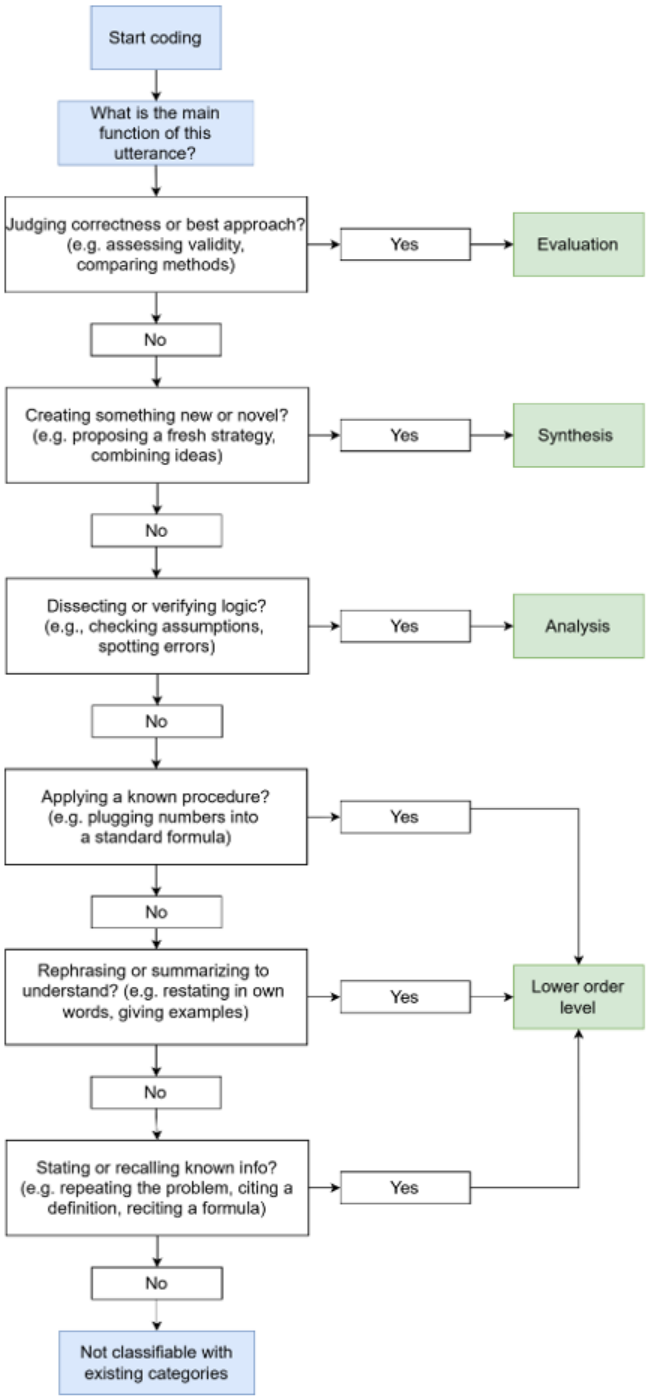


Figure 4. FTCB decision tree (the authors' own elaboration based on Webb, 1970)

strategy shift (the group adopted a different approach), an unrealized opportunity for strategic reconsideration (an opening arose that was not collectively taken up), or a local checking-and-correction moment (an error or inconsistency was identified and corrected without a broader change of approach). Classification was discussed until agreement was reached.

Examining these episodes in greater depth made it possible to trace how particular cognitive and dialogic contributions preceded, accompanied, or followed strategic adjustments during the problem-solving process.

Development of the coding trees and coding process

Coding was conducted in ATLAS.ti 23 (ATLAS.ti Scientific Software Development GmbH, 2023). Each coding unit received a T-SEDA code; an FTCB code was added when a mathematical contribution was present. Based on Webb's (1970) taxonomy, the first author drafted an FTCB decision tree, which was refined with the second author after independently coding the same 10% of the data and discussing discrepancies in two sessions until agreement was reached. The first three authors independently coded 10% of the corpus using the original T-SEDA table and discussed their coding in three sessions, which led to the development of a T-SEDA decision tree to sharpen distinctions between dialogic moves. These choices clarified boundaries between codes and reduced subjectivity by making selection criteria explicit. The first and second author then recoded another 10% using both trees, resolving remaining differences through discussion in two further sessions. Although some overlap between dialogic and cognitive categories was unavoidable, explicit decision rules supported consistent and transparent coding across the two frameworks. Finally, the first author applied the refined coding trees to the entire dataset. Reliability was thus established through the documented resolution of coding disagreements during calibration, which resulted in the decision rules incorporated in the trees, rather than through a reliability coefficient. This follows a consensus-based approach to intercoder reliability, in which coders arrive at a shared understanding of the codes through dialogue and the resulting coding instrument is then applied to the remaining data (Cofie et al., 2022). In addition, the coded episodes are presented with their transcripts, so that readers can inspect the coding decisions directly. The code frequencies reported in the results are used descriptively, to situate the episode analysis rather than to support statistical inference.

FTCB coding

The FTCB decision tree (**Figure 4**) guides coders through yes/no questions to determine whether a coding unit should be coded as *evaluation, synthesis, analysis,* or *lower order level*. The three higher FTCB levels were applied using the following criteria. *Evaluation* refers to judging the adequacy or correctness of a result or approach. A coding unit such as "Then it should work, right?" was coded as *evaluation* when it expressed a substantive judgment. Brief affirmations such as "Yes, that's right" were coded as *evaluation* only when they conveyed a meaningful judgment; otherwise, they were treated as *lower order level* or not coded. *Synthesis* was coded only when a student constructed a new, coherent plan by combining elements into a new approach. A high threshold was applied: tentative suggestions were not coded as *synthesis* unless they involved a structured,

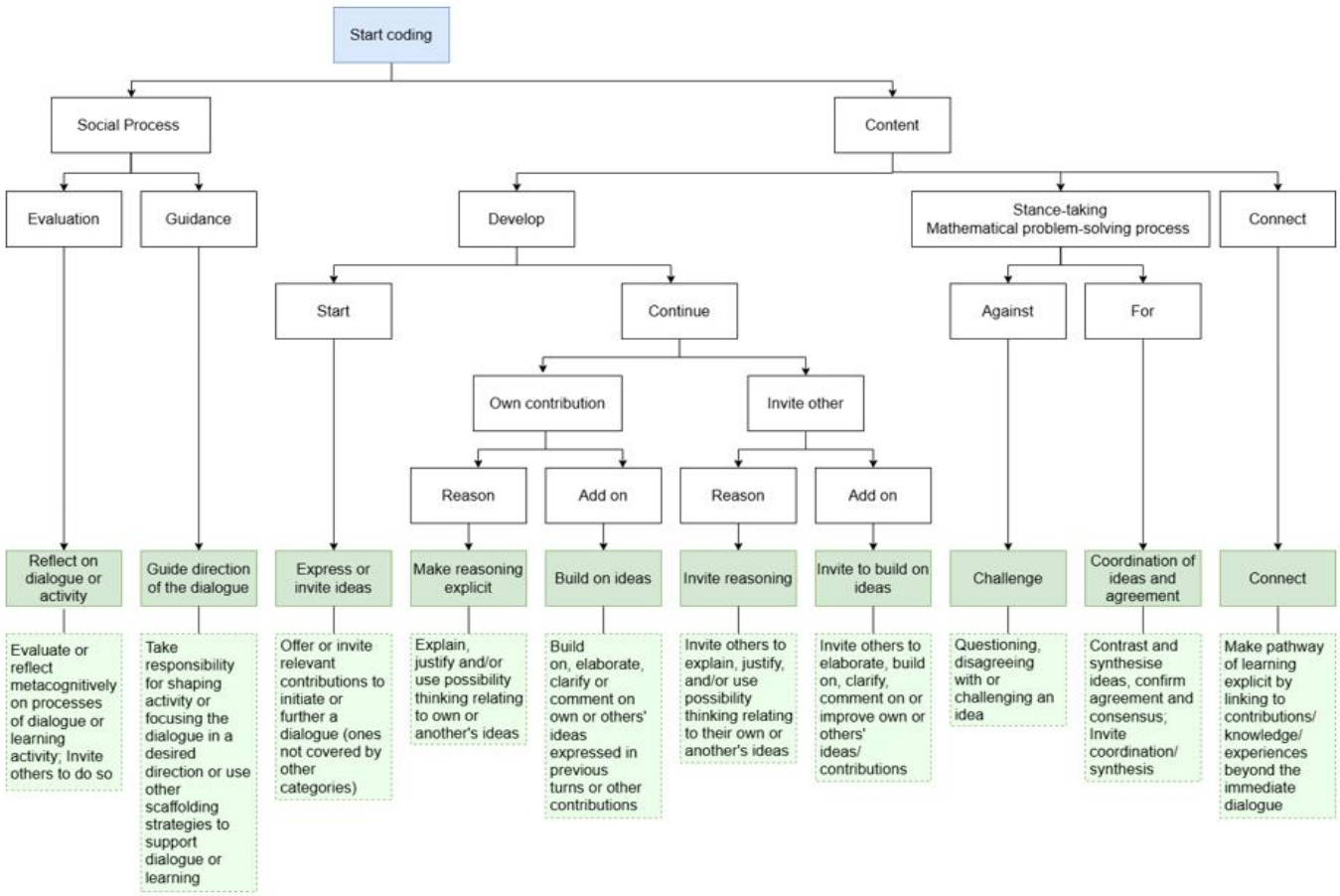


Figure 5. T-SEDA decision tree (the authors' own elaboration, adapted from Kershner et al., 2020)

multi-step strategy. For example, "Maybe it's an isosceles triangle?" was coded as *analysis*, as it offers a possible interpretation but does not, within the same coding unit, combine ideas into a broader strategy, which is required for *synthesis*. *Analysis* refers to interpreting, organizing, or restructuring mathematical information, for instance, identifying relationships or distinguishing structural differences ("The center of the circle is not the same as the center of the square"). Error spotting was also coded as *analysis,* as it involves examining relationships and identifying inconsistencies. *Lower order level* coding units included applying formulas without justification, reporting results, or referring to notational choices. Although the lower-order part of the decision tree contains three subtypes, these were grouped under a single code because they share a focus on familiar procedures or information. The subtypes include stating or recalling known information, paraphrasing to support understanding, and applying known procedures without considering alternatives. These contributions involve engagement with mathematical content but do not reorganize or restructure it. Clarification questions (e.g., "What do you mean by half plus x?") were coded as *lower order level* unless followed by restructuring of the problem. Making these lower-order subtypes explicit helps coders recognize when a coding unit truly involves *analysis*, instead of mistakenly assigning a higher level simply because it is not evaluative or synthetic. Across all categories, coders prioritized the highest cognitive level expressed in a coding unit.

T-SEDA coding

The T-SEDA decision tree (**Figure 5**) distinguishes whether a coding unit addresses the 'social process' of the dialogue or the (mathematical) 'content'. For social-process contributions, coders chose between *reflect on dialogue or activity* (commenting on the collaboration or inviting reflection) and *guide direction of the dialogue* (steering the discussion, proposing actions, or providing feedback).

When a coding unit focuses on 'content', coders selected between 'develop', 'stance-taking', and 'connect'. Within 'develop', coders determined whether a contribution 'starts' or 'continues' an idea. The 'start' branch corresponds to *express or invite ideas*, involving introducing a new idea, suggestion, or open question without explicitly referring to previous contributions. The 'continue' branch applies when a student builds on an existing idea. Two codes apply here: *make reasoning explicit* (explaining or justifying an idea, providing reasons, evidence, analogies, or exploring possibilities) and *build on ideas* (extending or reformulating an idea without offering reasons). Coding units were coded as *make reasoning explicit* when explanations or justifications

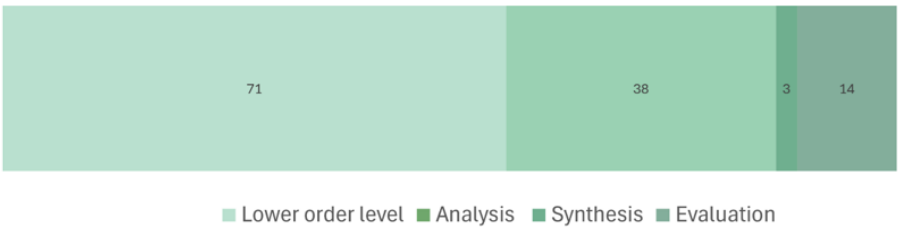


Figure 6. Total frequency of code occurrences FTCB (the authors' own elaboration)

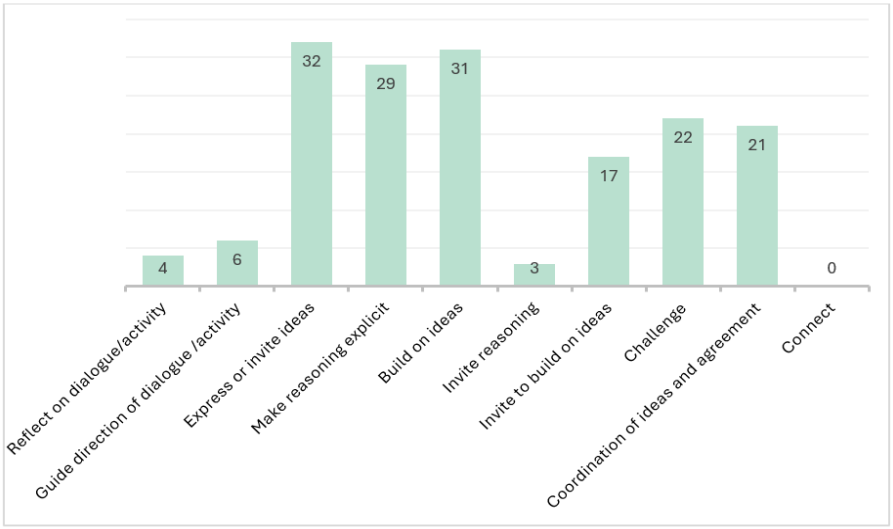


Figure 7. Total frequency of code occurrences T-SEDA (the authors' own elaboration)

were provided, and as *build on ideas* when additional information or examples were added without explanation. A second distinction within 'continue' is whether the speaker develops the idea themselves (own contribution) or invites others to do so. Inviting others to extend an idea was coded as *invite to build on ideas*, whereas inviting reasoning was coded as *invite reasoning*. We separated these functions to distinguish between inviting reasoning and inviting elaboration.

'Stance-taking' refers to expressing agreement or disagreement with an idea. This leads to two codes: *challenge* (disagreement) and *coordination of ideas and agreement* (agreement combined with links to prior contributions). We treated these codes separately because they represent different forms of engagement. Challenges may target a line of reasoning (questioning an idea) or factual content ("Yes, but an isosceles triangle doesn't have to have all its angles equal"), and this distinction proved relevant in our analysis. *Connect* captures contributions that link to earlier ideas or extend the discussion to wider contexts, such as referring back to previous statements, making the learning trajectory explicit, or encouraging further inquiry.

RESULTS

This section presents the findings in two parts, corresponding to the two sub-questions. First, in response to **RQ1**, we provide an overview of the cognitive and dialogic characteristics of the group's interaction across the 29-minute problem-solving session. Second, in response to **RQ2**, we examine selected episodes in which strategic flexibility became visible, focusing on moments where the group reconsidered, refined, or adjusted its problem-solving strategy, including instances triggered by potential or actual errors.

General Patterns of Dialogic and Cognitive Activity

Overall distributions and co-occurrences

This section addresses **RQ1** by outlining the group's overall cognitive and dialogic profile across the 29-minute problem-solving session. We first present descriptive patterns in the distribution and co-occurrence of codes to characterize the interaction of this particular group across the session. We then briefly report two sequential analyses. These analyses serve to bridge **RQ1** and **RQ2** by showing which dialogic moves tended to sustain ongoing activity and which more often preceded openings for strategy shifts.

Figure 6 shows that most contributions were coded as *lower order level*, with fewer instances of *analysis*, and only occasional *evaluation* or *synthesis*.

Figure 7 shows that the most frequently used T-SEDA codes were *express or invite ideas*, *build on ideas*, and *make reasoning explicit*. The code *challenge* also occurred relatively often, while process-management moves like *reflect on dialogue or activity* and *invite reasoning* were rare.

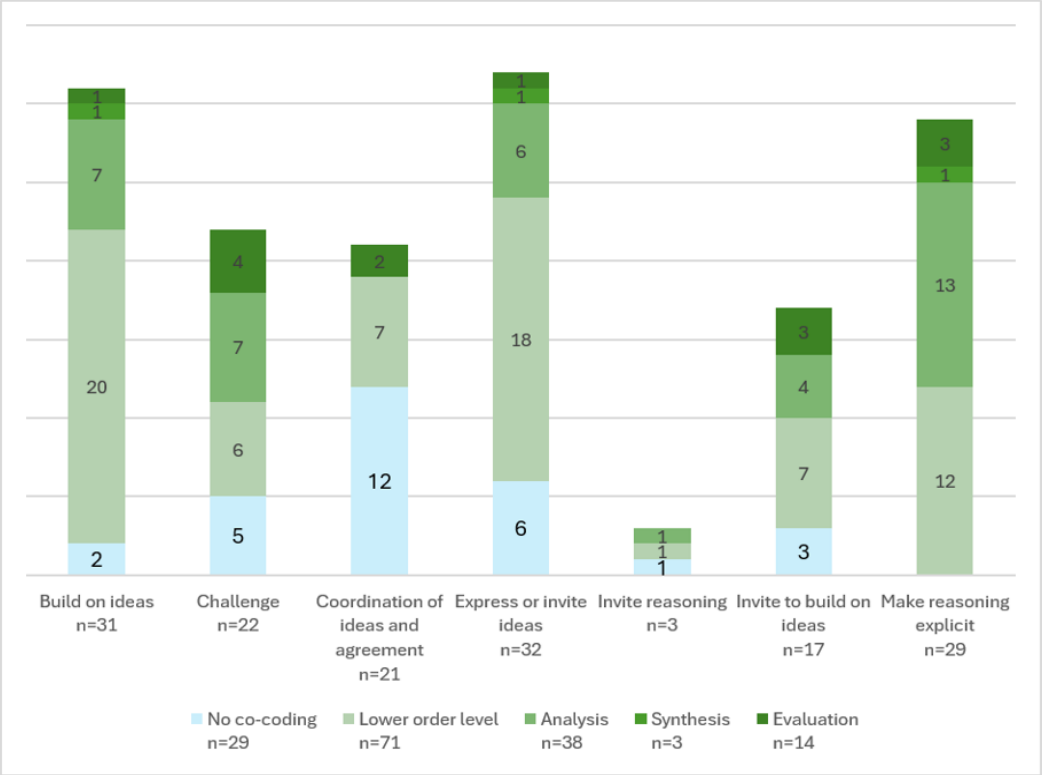


Figure 8. Code co-occurrence (the authors' own elaboration)

Notably, the code *connect* did not occur at all. These frequencies provide a first indication of the group's dialogic tendencies.

Most coding units received two codes, one from FTCB and one from T-SEDA. **Figure 8** displays the co-occurrence of codes from the two frameworks, indicating how frequently each T-SEDA code appeared alongside specific FTCB codes. For example, the code *build on ideas* co-occurred with an FTCB code in 29 out of 31 instances. Of these, 20 were categorized as *lower order level*, seven as *analysis*, one as *synthesis*, and one as *evaluation*. *Express or invite ideas* and *build on ideas* frequently coincided with *lower order level*. In contrast, *make reasoning explicit* and challenge co-occurred relatively more often with higher-order codes (analysis, evaluation, and synthesis).

Sequential background to the episode analysis

To complement the descriptive overview and to contextualize the episode analysis in **RQ2**, we conducted a sequential analysis of immediate code-to-code transitions between coding units. We focused on two types of transitions that were most relevant to the later episode analysis: transitions between dialogic moves and transitions from dialogic moves to subsequent cognitive levels. For each unit, we examined which codes it received and which codes appeared in the immediately following unit. Because a single speaker turn could contain more than one coding unit, the following unit could be produced either by the same student or by another student. The final unit in the video was excluded, as no subsequent unit followed it.

Figure 9 presents a sequential analysis of T-SEDA codes. The vertical axis represents the codes assigned to initial coding units, while the horizontal axis shows the codes assigned to the immediately following unit. The area of each dot is proportional to the frequency it represents. For example, *make reasoning explicit* was followed by *challenge* six times.

Build on ideas and *express or invite ideas* were often followed by the same code, indicating that idea-sharing tended to be sustained across successive units. In contrast, *challenge* was often followed by *coordination of ideas and agreement*, suggesting that some disagreements were resolved quickly rather than explored further. The most notable departures from these stable patterns involved *make reasoning explicit*, which was often followed either by another instance of the same move or by *challenge*. **Figure 9** thus suggests that idea-sharing was typically self-sustaining, whereas only a limited subset of dialogic moves opened space for strategic reconsideration.

Figure 10 presents a sequential analysis of T-SEDA-to-FTCB code transitions. The vertical axis shows the T-SEDA codes assigned to initial coding units, while the horizontal axis displays the FTCB codes assigned to the immediately following units. Each dot represents the frequency with which a specific T-SEDA code was followed by a particular FTCB code. For example, *challenge* was followed by *analysis* five times.

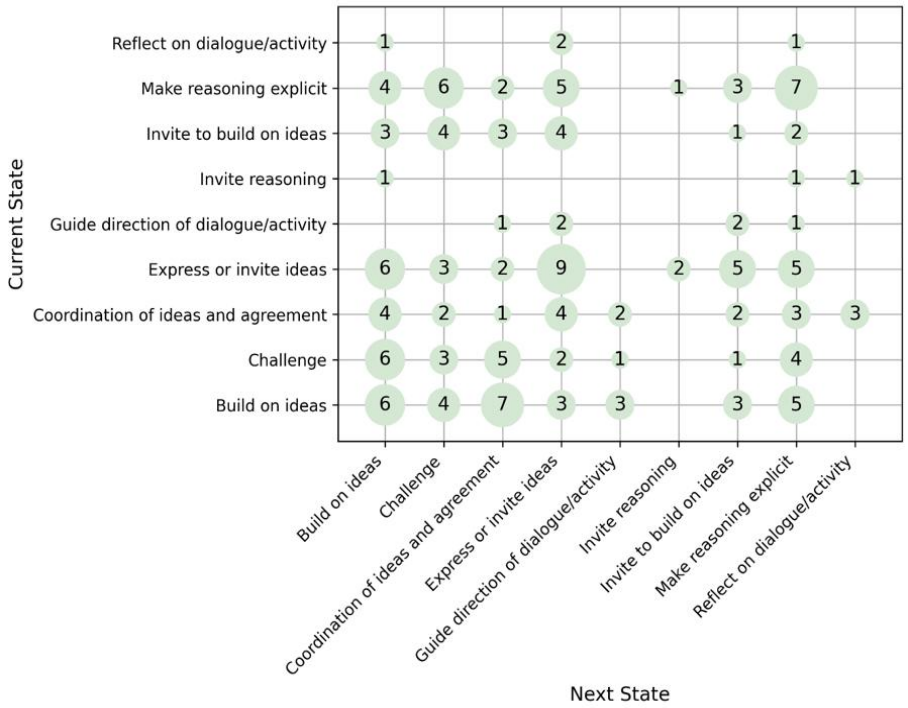


Figure 9. Sequence of T-SEDA codes across contributions (Álvaro Schipper based on the authors' analysis)

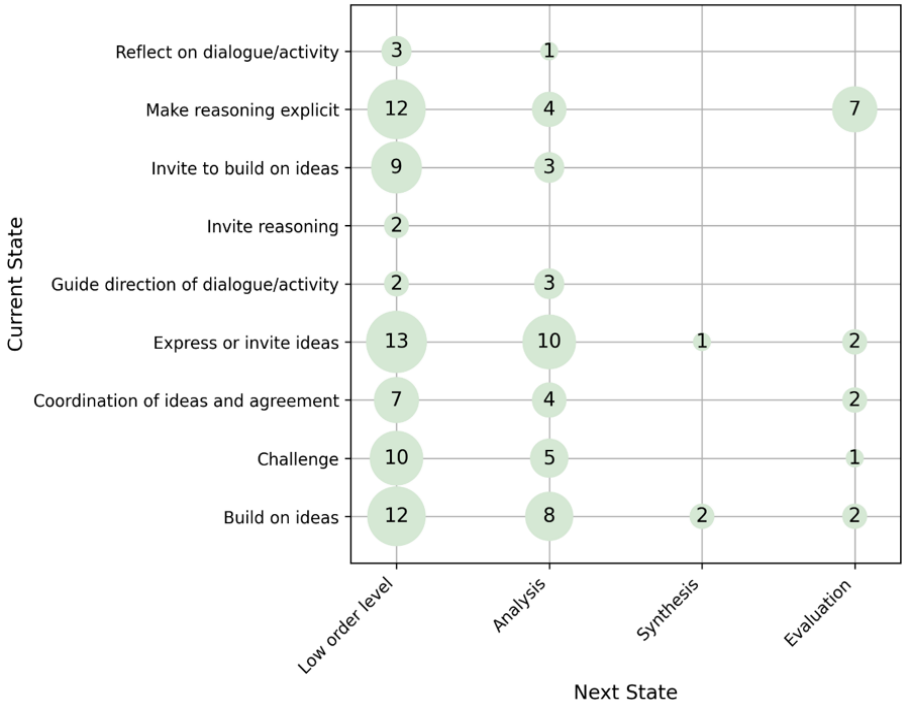


Figure 10. Distribution of subsequent FTCB codes per T-SEDA code (Álvaro Schipper based on the authors' analysis)

Lower-order responses occurred most frequently across all dialogic categories, consistent with the overall prevalence of lower-order thinking in the group. However, some dialogic moves were more often followed than others by higher-order responses. *Analysis* followed *express or invite ideas* ten times, *build on ideas* eight times, and *challenge* five times. The strongest pattern involved *make reasoning explicit*, which was followed seven times by *evaluation* and four times by *analysis*. Reported here as contextual background for **RQ2**, these patterns suggest that some dialogic moves more often than others were followed by analytically or evaluatively richer contributions, with *make reasoning explicit* standing out most clearly.

Overall, the descriptive analyses indicate active participation but limited cognitive depth across the session. This baseline provides context for the following section, which examines specific episodes in which the group reconsidered, refined, or shifted its approach.

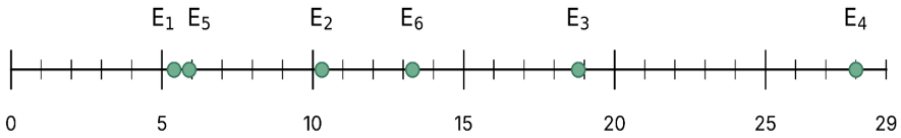


Figure 11. Timeline of the 29-minute CPS-session showing the approximate starting points of the six episodes (E1-E6). The timeline is schematic: episodes occasionally overlap as the dialogue alternated between productive and unproductive reasoning paths (the authors' own elaboration)

Table 1. Episode 1

Turns	Transcript excerpt	Coding	
		FTCB	T-SEDA
38	P: Suppose this is the center. And then it's umm ... how did it go again ... not that ... If it goes to here, then this is a right angle.	*Analysis*	*Express or invite ideas*
39	R: Is that 90 degrees?	*Analysis*	*Invite reasoning*
40	P: This, yeah. I haven't drawn it very nicely, but then it's the 90 degrees.	*Lower order level*	*Make reasoning explicit*

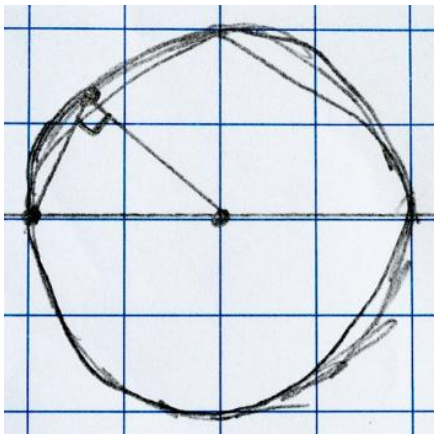


Figure 12. Student P's initial sketch referring to Thales' theorem (student's written work, reproduced with consent)

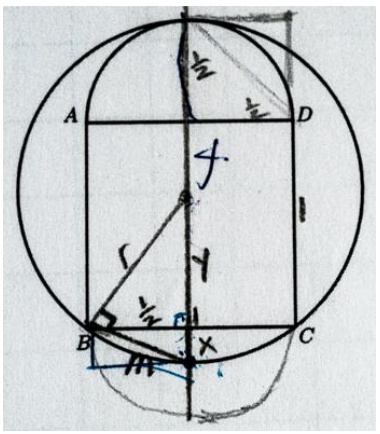


Figure 13. Student P's revised diagram with added right angle (student's written work, reproduced with consent)

Episodes Focusing on Strategic Flexibility

To address **RQ2**, this section zooms in on the episodes that are relevant to the analysis of strategic flexibility. These episodes represent moments where the group reconsidered, refined, or shifted its approach, including instances in which potential or actual errors prompted a check or adjustment of the strategy. By examining these sequences in context, we can see how particular combinations of cognitive and dialogic contributions were associated with moments of strategic adjustment during the task.

Each episode in **Figure 11** marks a new phase in the group's problem-solving activity, points where the group reconsidered, refined, or adjusted its strategy. Time is displayed in minutes from the beginning of the session.

Strategy shifts

Drawing on the integrated FTCB and T-SEDA analysis, this section focuses on sequences in which the group considered, initiated, or enacted changes in its problem-solving approach, showing how different types of contributions related to actual or potential strategy shifts during the task. During the problem exploration phase, P introduced an idea based on what we suspect was a misapplication of Thales' theorem, sketching an initial diagram and then adding a right angle to the problem figure.

In the episode in **Table 1**, P's contribution was coded as *analysis* (FTCB) combined with *express or invite ideas* (T-SEDA). This episode is analytically relevant as an unrealized opportunity for strategic reconsideration rather than as an actual strategy shift. It marks the moment when P introduces a new way of structuring the figure, first in the sketch shown in **Figure 12** and then in the original problem diagram in **Figure 13**. However, the assumption underlying this construction was mathematically unjustified, and the group did not examine it collectively. When R questioned whether the marked angle was indeed 90 degrees, this created a potential opportunity to recognize the problem and reconsider the strategy. However, R was quickly satisfied with P's response and did not probe further. The group continued with an unproductive approach for a substantial part of the session. P later developed this line of reasoning into a system of equations with four equations and three unknowns.

Table 2. Episode 2

Turns	Transcript excerpt	Coding	
		FTCB	T-SEDA
69	R: I was thinking if you have the area …	*Analysis*	*Express or invite ideas*
70	R: That's radius times radius times pi, the radius is one and a half plus x divided by 2, times one and a half plus x divided by 2 times pi …	*Analysis*	*Build on ideas*
71	R: … then you only have x, only one unknown.	*Synthesis*	*Build on ideas*
72	P: But you don't have an equals sign, so you now have an equation with two unknowns: area and x are still unknown.	*Analysis*	*Challenge*

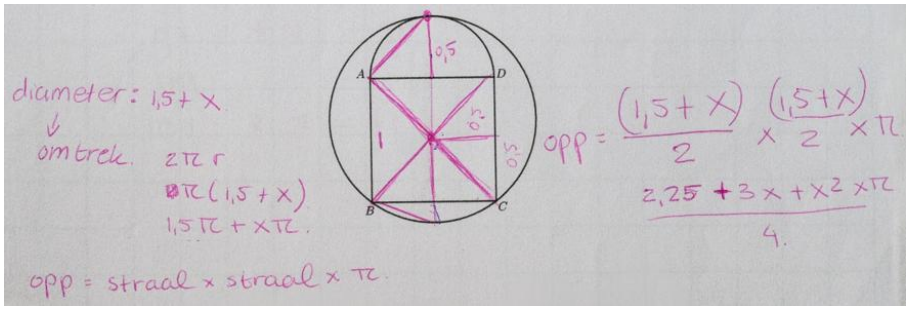


Figure 14. R's proposed alternative approach presented to the group during turns 69-71 (student's written work, reproduced with consent)

Table 3. Episode 3

Turns	Transcript excerpt	Coding	
		FTCB	T-SEDA
86	P: … so y is ¼ over x ...	*Lower order level*	*Make reasoning explicit*
87	P: Only when I look at it like this ... I don't really agree ... somewhere I did something weird ...	*Evaluation*	*Challenge*
89	R: [unintelligible] that y equals $4x$ because y is longer than x.	*Analysis*	*Make reasoning explicit*
90	P: If I do it the other way around it turns out exactly the same and that would make more sense, right? Yeah, but how come?	*Analysis*	*Make reasoning explicit*
93	P: Oh I think I made a pretty big mistake somewhere. I made a mistake, that's why we're not getting anywhere, look. And you didn't notice either? Nope. I drew from here to there, this is then the center of the circle ... that's the radius of the circle. This distance to B should obviously be the same as this one. If I've got the center of the circle here ... then the r has to be the same length as this … So what I drew here is just completely wrong. So I must've made a false assumption somewhere …	*Evaluation*	*Challenge*

At this point, R proposed an alternative approach (**Table 2**) that she had worked out on paper (see **Figure 14**). In **Figure 14**, 'omtrek' means 'circumference', 'opp' is an abbreviation for 'area', and 'straal' means 'radius'. In transcript excerpt 69 to 71, R explains this new idea to the rest of the group.

R's suggestion (**Table 2**) here was also coded as *analysis* (FTCB) and *express or invite ideas* (T-SEDA), making this episode analytically relevant as a proposed alternative strategy that was recognized and evaluated, but not developed further. P's subsequent identification of a flaw in this idea, which the others accept without further discussion, suggests that strategy decisions continued to be guided mainly by P, with limited collective examination of alternative approaches.

P succeeds in identifying an error in R's reasoning, and the rest of the group agrees. P then continues working with the system of equations and, using the method of substitution, arrives at a solution. **Table 3** shows the moment (turns 86 to 93) when P evaluates the obtained solution by examining the ratio between x and y in the diagram (**Figure 15**) and shares her doubts with the group. R joins in the reasoning. This episode is analytically relevant as an unrealized opportunity for strategic reconsideration. At this point, the group stops working on the problem. Although P's doubt and R's brief participation created an opportunity to question the viability of the current approach, this opening was not taken up collectively or developed into a revised strategy. Later, P revisits her steps and discovers an error in her own reasoning (excerpt 93).

In this sequence, P's talk (**Table 3**) was coded as *lower order level* (FTCB) combined with *make reasoning explicit* (T-SEDA), as she set out her computations and as *evaluation* and *challenge* when she questioned whether the obtained solution was plausible. This episode is

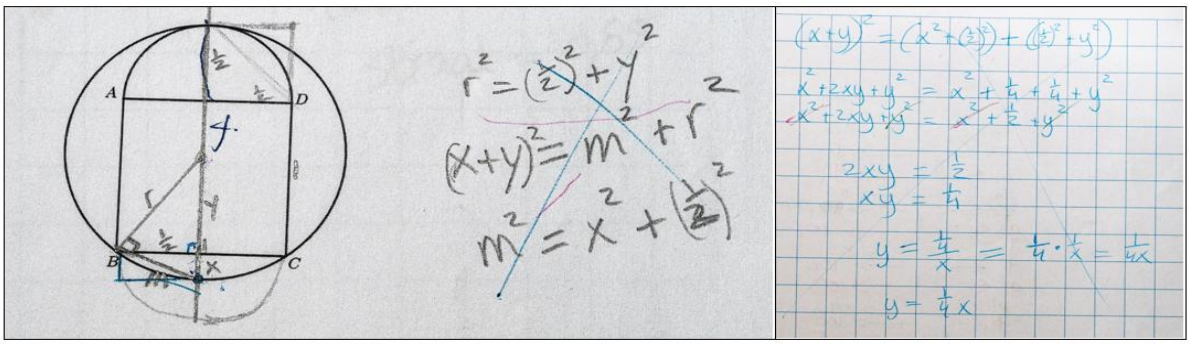


Figure 15. P's cancelled calculations before exploring an alternative approach (student's written work, reproduced with consent)

Table 4. Episode 4

Turns	Transcript excerpt	Coding	
		FTCB	T-SEDA
99	P: I'll just draw it again then ...	*Lower order level*	*Express or invite ideas*
100	P: And then just measure it, right? (laughs).	*Lower order level*	*Express or invite ideas*
108	P: Maybe I can just make a ratio table here ...	*Synthesis*	*Express or invite ideas*
109	P: Because I've redrawn it exactly to scale. I've drawn it all nicely to scale ... and then it's [unintelligible] ... and the full circle is 6.8. So then I can calculate: 6.8/4 … that's 1.7. And that's the diameter …	*Analysis*	*Make reasoning explicit*
113	Q: But then ... does that work out nicely with the measurement? That's the question. Doesn't it give a weird root or something?	*Evaluation*	*Challenge*
114	P: Now I just measured it with a ruler ... or redrew it to scale (unintelligible).	*Lower order level*	*Build on ideas*

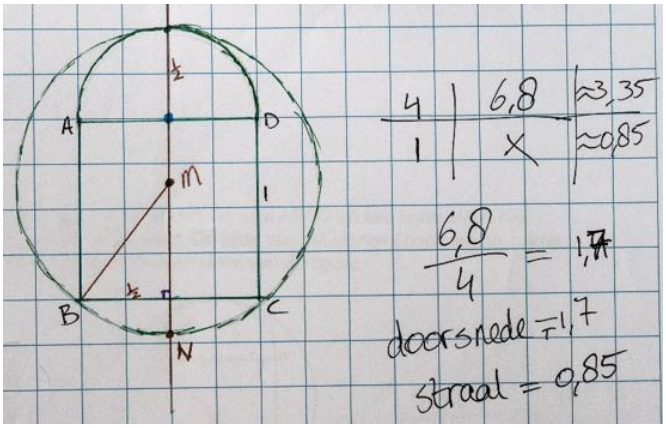


Figure 16. P's scale drawing used to estimate the value of r (student's written work, reproduced with consent)

analytically relevant as an unrealized opportunity for strategic reconsideration. R's brief involvement suggests some shared attention, but the issue was not taken up collectively, and the group moved on without resolving the inconsistency, indicating that this moment did not develop into a shared reconsideration or strategy shift.

The transition from Episode 3 to Episode 4 illustrates the interactional conditions preceding the only actual strategy shift in the session. P's evaluative recognition of her own error (turns 87, 93), coded as *evaluation* combined with *challenge*, created the opening that made a change of approach possible. This sequence was preceded by P making her reasoning explicit and then subjecting her own result to scrutiny, a pattern consistent with the sequential finding that *make reasoning explicit* was more often followed by higher-order contributions than were other dialogic moves. P then crosses out the earlier calculations (**Figure 15**) and initiates a new approach by redrawing the figure to scale (**Figure 16**). Notably, while the conditions for this shift involved cognitively rich contributions, the shift itself (turn 99) was introduced procedurally and without collective deliberation, suggesting that the group's capacity to create openings for strategic reconsideration did not extend to a capacity for shared enactment of a new strategy.

The episode in **Table 4** is analytically relevant as an actual strategy shift. Episode 4 marks an observable shift in strategy: P abandons her previous calculations and initiates a new approach by redrawing the figure to scale. This coding unit was coded as *lower order level* (FTCB) and *express or invite ideas* (T-SEDA), indicating that the strategic shift was introduced procedurally and with little explicit justification, and that the others followed her lead rather than co-constructing the new plan.

Despite Q's expressed doubts, P maintains this solution, so the new strategy is pursued without further collective examination.

Local checking and correction

This section examines episodes in which potential or actual errors prompted the group to check or reject

Table 5. Episode 5

Turns	Transcript excerpt	Coding	
		FTCB	T-SEDA
43	Q: This piece is a half plus x.	*Lower order level*	*Build on ideas*
44	P: No, as in, if it had been a half, then it would've been exactly in the middle, and it's not ...	*Analysis*	*Challenge*

Table 6. Episode 6

Turns	Transcript excerpt	Coding	
		FTCB	T-SEDA
79	S: Can't you ... because erm ... because this angle is 90 degrees, you know it's a half ... Can't you calculate it with sine and cosine?	*Analysis*	*Invite to build on ideas*
80	P: We don't know the angles ... for sine and cosine you need angles.	*Lower order level*	*Challenge*
81	S: No, you can [unintelligible] opposite angle [unintelligible].	*Lower order level*	*Build on ideas*
82	P: But with sine and cosine ... you've got a 90 degrees ... and you assume: another angle, an angle that isn't 90 degrees.	*Lower order level*	*Challenge*

locally proposed relations, procedures, or suggestions. Several such moments occurred during the session, often identified by P. In these cases, noticing an inconsistency or a misapplied relation helped the group rule out a non-viable line of reasoning before it developed further. These instances can be viewed as local forms of checking and correction rather than as actual strategy shifts. Although these moments were primarily identified by P, they played an important role in maintaining progress by encouraging the group to reconsider aspects of its approach. The following excerpts illustrate how such checks unfolded in dialogue and how students' cognitive and dialogic contributions related to the resulting strategic adjustments. Most interactions in these episodes involved relatively low cognitive demand, as they focused on verifying known procedures or relationships rather than developing new ideas.

Table 5 presents one such episode. While the group was following a particular strategy, Q proposed a relationship between two line segments, one known and one unknown, which initially appeared plausible. P immediately recognized that the relationship could not hold and explained why, prompting the group to reject this local suggestion and continue without developing it further.

In this episode, Q's contribution was coded as *lower order level* (FTCB) and *build on ideas* (T-SEDA), introducing an incorrect relationship between two line segments. This coding unit was coded as *lower order level* rather than *analysis* because it stated an assumed relationship without examining or justifying it; according to the FTCB decision tree, *analysis* requires interpreting or restructuring mathematical information, whereas stating a relationship as given falls within the *lower order level* category of applying known information without considering alternatives. This episode is analytically relevant as a local checking-and-correction moment rather than as a strategy shift. P immediately challenges the suggestion and explains why it cannot hold. The others accept her correction without further discussion, which redirects the group but does not lead to deeper collective examination.

Table 6 presents another episode in which an error is identified, but this time it occurs during the search for an appropriate strategy. S points to a triangle in the diagram and suggests applying a particular formula to it. P corrects S by pointing out the necessary conditions for using that formula in a triangle.

In Episode 6, S's suggestion to apply trigonometric rules was coded as *analysis* (FTCB) and *invite to build on ideas* (T-SEDA), making this episode analytically relevant as another local checking-and-correction moment rather than as a strategy shift. This unit was coded as *analysis* rather than *lower order level* because S related a feature of the figure, the right angle, to the possibility of using sine and cosine, thereby identifying a relationship between elements of the problem. P's subsequent clarification that the necessary conditions for using sine or cosine were not satisfied shows how a potential error was identified early and filtered out quickly, with limited elaboration from the rest of the group.

Overall, P played a central role in steering the group's work across the episodes analyzed. She was involved in all six: she initiated three of them, including the only actual strategy shift, and in each of the other three she identified the flaw in a peer's proposal. Across these episodes, she contributed six of the seven coding units coded as *challenge*. Although her corrections often helped the group avoid unproductive paths, they may also have reduced opportunities for others to contribute to the consideration of alternative strategies.

CONCLUSION, LIMITATIONS, DISCUSSION AND IMPLICATIONS

This final section draws conclusions with respect to the overarching question, namely how cognitive levels and dialogic moves, together with their collective uptake, relate to strategic flexibility in collaborative

mathematical problem-solving. It first presents the conclusions, followed by the study's limitations, a discussion of the conclusions, and their implications for teaching and future research.

Conclusions

In addressing **RQ1**, the analysis showed frequent idea sharing and elaboration, creating active participation across the session. Contributions involving clarification or challenge were less common, but when they occurred they more often prompted analytical or evaluative responses. Consistent with the sequential analysis, idea-sharing moves tended to be followed by similar contributions, whereas clarifications and challenges were more often followed by higher-order moves. These moments briefly opened space for reconsidering the group's approach, but were rarely taken up or sustained, reinforcing the descriptive picture that most interaction concerned lower-order contributions and the sharing and extension of ideas rather than their collective examination. As a conclusion on **RQ1**, therefore we provide a descriptive interactional profile that helps situate the **RQ2** analyses.

In addressing **RQ2**, the analysis identified three types of episodes: actual strategy shifts, unrealized opportunities for strategic reconsideration, and local checking-and-correction moments. These episodes arose when analytically or evaluatively richer contributions were introduced and taken up, or left undeveloped, by other group members. When students made their thinking explicit or challenged a proposal, these moves were often accompanied by analytical or evaluative thinking and opened space for examining whether the current line of work should be maintained, revised, or abandoned. This was shown not only in the episode analyses themselves, but also in the sequential patterns: idea-sharing moves tended to sustain themselves, whereas moves that made reasoning explicit or put ideas under scrutiny more often preceded analytically richer follow-up contributions. In several instances, simple or inaccurate statements triggered such higher-order responses, but these did not automatically develop into shared strategy shifts. Rather, the episodes showed different outcomes: some openings remained unrealized, some led only to local checking or rejection of a non-viable step, and only one developed into an actual change of strategy. These moments reflect what Schoenfeld (1985) describes as the continuous monitoring needed to steer problem-solving activity in productive directions. At the same time, the episode analyses showed that such openings were rarely taken up and developed collectively. They were often identified by P, whose contributions played a disproportionate role in introducing, evaluating, and filtering possible next steps.

As a conclusion, what characterized the group's strategic work was not only the cognitive depth and dialogic function of contributions, but also whether these contributions were taken up within an interactional structure in which authority was unevenly distributed, a concern that resonates with broader research on collaborative mathematical problem-solving and small-group authority relations (Bleiler-Baxter et al., 2023; Wu, 2024).

The study's overarching question was: *How do cognitive levels and dialogic moves in students' interaction, together with their collective uptake, relate to strategic flexibility in collaborative mathematical problem-solving?*

From the findings we conclude that strategic flexibility emerged in episodes where cognitive depth aligned with interactional moves that made reasoning explicit or placed ideas under scrutiny. However, the openings created by such moves were rarely taken up by the group. Most were followed by lower-order or idea-sustaining contributions, which kept the conversation moving but did not build on the opportunity for deeper examination. Because higher-order contributions were not collectively taken up or extended, openings for strategy shifts often remained local, brief, or unrealized rather than developing into shared strategic flexibility. Strategic adjustment therefore occurred in brief, individual moments rather than as a sustained, shared process within the group.

Limitations

Several limitations should be considered when interpreting the findings of this study. First, the sequential analysis was based on one-step transitions between coded contributions. In practice, a single contribution can influence several later turns, meaning that longer-term or more diffuse effects of cognitive or dialogic moves may not have been fully captured. This analytic focus was intentional, prioritizing local interaction patterns that make moments of strategic checking or adjustment visible on a turn-by-turn basis.

The study was conducted under research conditions: students were aware that their interaction was being video recorded and were encouraged to articulate their thinking. This setting may have increased verbalization compared to ordinary classroom work. However, because all participants operated under the same conditions, these effects are unlikely to have influenced the relative distribution of dialogic and cognitive contributions.

Another limitation concerns the influence of group dynamics. Because participation was uneven, the findings reflect strategic adjustment within an asymmetrical interactional structure. This configuration may have influenced how freely students challenged ideas or pursued alternative strategies, meaning that the observations represent this particular group rather than a general model of CPS. Some of the observed interactional asymmetry may stem from task difficulty,

which limited how confidently some students could develop or take up mathematical ideas. Relatedly, our interpretation of uneven authority is based on observable participation, uptake, and the distribution of mathematically substantive contributions, rather than on direct access to students' intentions, confidence, or internal knowledge states.

The coding frameworks themselves also come with certain constraints. While FTCB and T-SEDA provide clear and reliable categories, they may not capture subtle or implicit forms of thinking, such as silent planning, hesitation, or internal evaluation. Although both frameworks provide explicit decision rules, applying a cognitive taxonomy to short coding units inevitably involves interpretative judgment, as some contributions can plausibly serve multiple functions. The analysis focused on the group's talk and written work during a single session; other data, such as interviews, were not collected. In addition, after the calibration rounds the remaining data were coded by the first author alone, which means that the consistency of the coding rests on the decision trees and the calibration process rather than on continued double coding.

Moreover, because FTCB and T-SEDA operate at the level of individual turns, their combined use can only provide a partial view of group-level strategic regulation. Even though the sequential analysis allowed us to trace how contributions influenced one another across turns, broader forms of shared monitoring and collective decision-making extend beyond what can be inferred from isolated coding units. This applies in particular to synthesis, which was coded only when a structured, multi-step plan was expressed within a single coding unit. This high threshold, reflected in the small number of synthesis codes, may have missed planning that emerged gradually or was distributed across several turns and speakers. Accordingly, our claims about strategic flexibility are limited to moments in which opportunities for strategy shifts became visible in the interaction, and do not capture all possible forms of strategic regulation that may have remained implicit or unspoken.

Finally, the findings are not statistically generalizable, but they do offer analytical generalization (Bakker, 2018; Flyvbjerg, 2011) by identifying mechanisms that may operate across CPS settings. This case provides insight into how dialogic and cognitive behavior shape strategic flexibility in small-group mathematics work. The study focused on a single group solving one mathematical problem within a specific educational context. Different task structures or subject domains may elicit other forms of interaction and strategic adjustment. Quantitative counts were used descriptively to contextualize qualitative patterns rather than to establish statistical relations.

Discussion

This study examined how cognitive and dialogic dimensions of students' interaction related to moments in which the group reconsidered or shifted its strategy. Although dialogic participation was high, much of the interaction remained cognitively shallow, and openings for strategy shifts were rarely taken up by the group as a whole. Below, we discuss three contributions of these findings.

First, conceptually, the findings suggest that strategic flexibility at the group level cannot be inferred from interactional activity alone. Whereas strategic flexibility has mainly been conceptualized as a capacity of individual solvers (Hong et al., 2023; Verschaffel, 2024), in this group it became visible in whether cognitively demanding contributions were taken up by others. What characterized these episodes was not simply a lack of higher-order contributions, but the limited collective uptake of such contributions once they had been voiced. This points to a conceptualization of strategic flexibility not only as an individual cognitive capability but also as an interactional and epistemic process that unfolds through sequences of talk.

Second, empirically, the analysis distinguished three outcomes of moments of strategic reconsideration: actual strategy shifts, unrealized opportunities, and local checking-and-correction moments. Only one episode developed into an actual shift; the others faded without collective examination or led to the correction of a single step. Prior research has established that authority relations shape whose ideas are pursued in small groups: Wu (2024) showed that group-level authority forms dynamically as individual and group authority compete and disperse across the interaction. Bleiler-Baxter et al. (2023) found that, when granted authority to co-construct proofs, undergraduates often accepted a peer's argument because that peer was confident and "owned" it, rather than through collective mathematical evaluation. Similarly, Ellis and Alzaga Elizondo (2026) showed how the right to validate a group's work was tied to an expert position, and Hansen and Naalsund (2026) described one-directional interaction patterns in which a peer's input was assimilated, refined, or neglected, with little joint evaluation. The present study relates such asymmetries to strategic flexibility: in the episodes analyzed, proposals by P's peers were evaluated and set aside by P, whereas the few questions raised about P's own proposals were brief and quickly settled. This suggests that, in this group, authority relations were associated not only with who contributed, but also with which strategic openings were examined and which remained unrealized.

Third, methodologically, the combined use of FTCB and T-SEDA made visible what neither framework captures on its own. The only actual strategy shift was introduced through a contribution coded as *lower order*

level and *express or invite ideas*, which, taken on its own, would suggest little cognitive depth. It became recognizable as a shift in relation to the preceding contributions coded as *evaluation* and *challenge*, in which P questioned her own result. Conversely, contributions coded as *analysis,* such as P's proposal in Episode 1 and S's suggestion in Episode 6, did not lead to a change of approach: the former was not examined by the group, the latter was checked and set aside. Coding each unit on both dimensions and examining their sequence thus made it possible to distinguish cognitively demanding contributions that were taken up from those that remained unexamined.

The analytic generalizability of this study does not lie in the frequency of particular codes, which serve to describe the interactional context of this specific group, but in a mechanism observed in the interaction: openings for strategy shifts arose when cognitively richer contributions were made visible in the dialogue, but these openings developed into shared strategic flexibility only when they were collectively examined and extended. Where uptake was weak and authority remained concentrated, such openings stayed brief, local, or unrealized. By highlighting the fragility of strategic adjustments in the absence of reciprocal uptake, the study contributes a more process-oriented understanding of how strategic decisions are jointly negotiated in collaborative mathematical problem-solving.

Implications For Practice and Research

The findings of this study suggest that effective CPS mathematics education depends not only on encouraging participation and open dialogue, but also on fostering interactional norms that support deeper engagement with mathematical strategy use. Although students frequently shared and elaborated on one another's ideas, opportunities for strategic adjustment were limited, particularly when contributions from more dominant participants were not questioned or taken up for further examination. These findings point to the importance of instructional practices that support students in making their reasoning explicit, questioning and testing emerging ideas, asking clarifying questions, and examining alternative approaches as part of collaborative mathematical work, regardless of status or perceived expertise. More specifically, teaching for CPS should not only promote participation, but also support forms of uptake through which mathematically substantive contributions are collectively examined rather than briefly acknowledged and dropped.

Beyond their pedagogical relevance, the findings also suggest directions for future research in mathematics education. Applying the combined FTCB and T-SEDA frameworks across a broader range of mathematical tasks, group compositions, and educational contexts could help identify how cognitive-dialogic patterns support or constrain opportunities for strategic adjustment. In particular, future research might examine more systematically how uptake, authority relations, and task characteristics shape whether openings for strategy shifts are collectively developed, locally corrected, or left unrealized. Comparative analyses may further clarify how such patterns unfold in related domains that share similar epistemic and didactic challenges. Taken together, these implications highlight both practical and theoretical pathways for advancing understanding of how collaborative mathematical dialogue shapes students' opportunities to engage flexibly with strategy use during problem-solving.

Author contributions: HC: data curation, formal analysis, writing – original draft; **JS & TB:** validation; **SG & PD:** conceptualization, supervision, writing – review & editing. All authors agreed with the results and conclusions.

Funding: No funding source is reported for this study.

Acknowledgments: The authors would like to thank the software engineer Álvaro Schipper who wrote the Python code for this analysis and created **Figure 9** and **Figure 10**. The authors would also like to thank the students who participated in this study.

Ethical statement: This study was approved by the Ethics Committee for Research in the Social Domain at HU University of Applied Sciences Utrecht on 20 June 2024 with approval number 2024-42. Written informed consent was obtained from all participating students.

AI statement: During the preparation of this work, the first author used Claude (Anthropic, Opus 5.5) to assist in improving English language clarity and readability. Perplexity (web version, 2026) was used to search for literature. The authors reviewed and edited the content as needed and took full responsibility for the final version of the manuscript.

Declaration of interest: No conflict of interest is declared by the authors.

Data sharing statement: Data supporting the findings and conclusions are available upon request from the corresponding author.

REFERENCES

ATLAS.ti Scientific Software Development GmbH. (2023). ATLAS.ti (version 23.2.1). *ATLAS.ti Scientific Software Development GmbH*. https://atlasti.com/

Bakker, A. (2018). *Design research in education: A practical guide for early career researchers*. Routledge. https://doi.org/10.4324/9780203701010

Bleiler-Baxter, S. K., Kirby, J. E., & Reed, S. D. (2023). Understanding authority in small-group co-constructions of mathematical proof. *The Journal of Mathematical Behavior, 70*, Article 101015. https://doi.org/10.1016/j.jmathb.2022.101015

Bloom, B. S. (1956). *Taxonomy of educational objectives: The classification of educational goals*. Longmans.

Bolat, R. C., & Arslan, Ç. (2024). Examination of mathematics teachers' strategic flexibility in solving mathematical problems. *Thinking Skills and Creativity, 54*, Article 101679. https://doi.org/10.1016/j.tsc.2024.101679

Brown, B. B., Ober, R. L., Soar, R., & Webb, J. N. (1966). *Florida taxonomy of cognitive behaviour* [Unpublished internal report]. University of Florida.

Ceballos, H., van den Bogaart, T., van Ginkel, S., Spandaw, J., & Drijvers, P. (2026). How collaborative problem solving promotes higher-order thinking skills: A systematic review of design features and processes. *Thinking Skills and Creativity, 59*, Article 102001. https://doi.org/10.1016/j.tsc.2025.102001

Cofie, N., Braund, H., & Dalgarno, N. (2022). Eight ways to get a grip on intercoder reliability using qualitative-based measures. *Canadian Medical Education Journal, 13*(2), 73-76. https://doi.org/10.36834/cmej.72504

Ellis, B. M., & Alzaga Elizondo, T. (2026). Emergent storylines that influence positions and mathematical status in collaborative small-group proof activity. *Educational Studies in Mathematics, 123*, 275-292. https://doi.org/10.1007/s10649-026-10504-6

Flyvbjerg, B. (2011). Case study. In N. K. Denzin, & Y. S. Lincoln (Eds.), *The SAGE handbook of qualitative research* (4th ed., pp. 301-316). SAGE.

Ghic, G. (2012). Case studies – The role of them in mathematics teaching. *Procedia – Social and Behavioral Sciences, 46*, 2681-2685. https://doi.org/10.1016/j.sbspro.2012.05.546

Hamilton, M. W. (2026). What is mathematical authority? A Weberian approach to defining the construct and exploring its affordances. *Educational Studies in Mathematics*. https://doi.org/10.1007/s10649-026-10554-w

Hansen, E. K. S., & Naalsund, M. (2026). Toward an analytical model for noticing interaction patterns in creative collaborative mathematical reasoning. *Mathematical Thinking and Learning, 28*(3), 295-313. https://doi.org/10.1080/10986065.2025.2452222

Hennessy, S., Rojas-Drummond, S., Higham, R., Márquez, A. M., Maine, F., Ríos, R. M., García-Carrión, R., Torreblanca, O., & Barrera, M. J. (2016). Developing a coding scheme for analysing classroom dialogue across educational contexts. *Learning, Culture and Social Interaction, 9*, 16-44. https://doi.org/10.1016/j.lcsi.2015.12.001

Hong, W., Star, J. R., Liu, R.-D., Jiang, R., & Fu, X. (2023). A systematic review of mathematical flexibility: Concepts, measurements, and related research. *Educational Psychology Review, 35*, Article 104. https://doi.org/10.1007/s10648-023-09825-2

Jordan, B., & Henderson, A. (1995). Interaction analysis: Foundations and practice. *The Journal of the Learning Sciences, 4*(1), 39-103. https://doi.org/10.1207/s15327809jls0401_2

Kershner, R., Dowdall, K., Hennessy, S., Owen, H., & Calcagni, M. E. (2020). Teachers as 'natural experimenters': Using T-SEDA to develop classroom dialogue. In E. Hargreaves, & L. Rolls (Eds.), *Reimagining professional development in schools* (pp. 102-122). Routledge. https://doi.org/10.4324/9780429293337-8

Langer-Osuna, J. (2016). The social construction of authority among peers and its implications for collaborative mathematics problem solving. *Mathematical Thinking and Learning, 18*(2), 107-124. https://doi.org/10.1080/10986065.2016.1148529

Littleton, K., & Mercer, N. (2013). *Interthinking: Putting talk to work*. Routledge. https://doi.org/10.4324/9780203809433

Schoenfeld, A. H. (1985). *Mathematical problem solving*. Academic Press.

Stake, R. E. (1995). *The art of case study research*. SAGE.

Star, J. R., & Rittle-Johnson, B. (2008). Flexibility in problem solving: The case of equation solving. *Learning and Instruction, 18*(6), 565-579. https://doi.org/10.1016/j.learninstruc.2007.09.018

Stichting Nederlandse Wiskunde Olympiade. (2022). Puzzelspel eerste ronde [Puzzle game first round]. *Nederlandse Wiskunde Olympiade*. https://www.wiskundeolympiade.nl/files/opgaven/1e-ronde/2022/opgaven.pdf

Verschaffel, L. (2024). Strategy flexibility in mathematics. *ZDM – Mathematics Education, 56*, 115-126. https://doi.org/10.1007/s11858-023-01491-6

Verschaffel, L., Luwel, K., Torbeyns, J., & Van Dooren, W. (2011). Analyzing and developing strategy flexibility in mathematics education. In J. Elen, E. Stahl, R. Bromme, & G. Clarebout (Eds.), *Links between beliefs and cognitive flexibility: Lessons learned* (pp. 175-197). Springer. https://doi.org/10.1007/978-94-007-1793-0_10

Webb, J. N. (1970). The Florida taxonomy of cognitive behavior. In A. Simon, & E. G. Boyer (Eds.), *Mirrors for behavior II: An anthology of observation instruments*. Research for Better Schools.

Wu, J. (2024). Research on individual authority and group authority relations in collaborative problem solving in middle school mathematics. In Y. Cao (Ed.), *Students' collaborative problem solving in mathematics classrooms* (pp. 75-99). Springer. https://doi.org/10.1007/978-981-99-7386-6_4