Developing an analytical tool for investigating knowledge in a Research Situation for Teacher Education

Mickaël Da Ronch[1], Michèle Gandit[2] and Ismaïl Mili[3]

[1]University of Teacher Education Valais, Saint-Maurice, Switzerland; mickael.daronch@hepvs.ch

[2]Grenoble Alps University, Grenoble, France

[3]University of Fribourg, Fribourg, Switzerland

Research Situations for Teacher Education (RSiTEs) are designed around a discrete mathematics problem to engage prospective teachers in proof-related activity involving not only mathematical knowledge but also didactic and professional knowledge. This article presents the development and subsequent testing of a coding scheme aimed at identifying these forms of situational knowledge in session transcripts. The scheme was applied to two sessions of a RSiTE devoted to mathematical proof. The results show a low level of disagreement regarding segmentation ($WD \approx 2.07\%$) and an almost perfect level of inter-coder agreement for the categorization of sequences across coder pairs ($\kappa_1 \approx 0.88$ et $\kappa_2 \approx 0.94$). The findings also highlight the presence of categories associated with these different forms of knowledge. Overall, the results support the relevance of the proposed model and pave the way for more fine-grained qualitative analyses.

Keywords: Research situation for teacher education, pre-service teacher education, mathematical and didactic knowledge, research methods in mathematics education.

Introduction

A *Research Situation for Teacher Education* (RSiTE), used here as the English translation of the French concept *Situation de Recherche pour la Formation des enseignants (SiRFe*), is a teacher education design (Da Ronch et al., 2025) aimed at engaging prospective or practicing teachers in proof-related activity (Alcock, 2008; Alcock & Weber, 2004) through a problem drawn from discrete mathematics, while fostering the emergence of didactic and professional knowledge. In this paper, we show how such situational knowledge can be identified through the observation of the implementation of a research situation with student teachers, whose primary objective is the acquisition of mathematical knowledge, most of which is not conceptual in nature. We further characterize these forms of mathematical knowledge, which are associated with proof practices and often remain invisible to students and even to teachers. Drawing on the work of Sackur et al. (2005), we refer to them as Type II situational knowledge, in order to distinguish them from Type I situational knowledge, which concerns mathematical objects and their properties. To date, a limited number of studies (e.g., Chanudet et al., 2019) related to formative assessment and focusing on the classroom implementation of problem-solving activities have shown how such mathematical knowledge may emerge from the analysis of actual classroom interactions. However, none of these studies relies on the development of an instrument whose reliability has been established through statistical validation.

Our research question therefore concerns the development of a reliable instrument capable of making such situational knowledge observable within a RSiTE and of extending this analysis to didactic and professional knowledge. Based on the assumption that a RSiTE simultaneously promotes the emergence of mathematical knowledge related to proof processes, didactic knowledge concerning the conditions under which this mathematical knowledge is produced, and professional knowledge, we present the development of an instrument designed to identify these different forms of situational knowledge in the discourse generated during the enactment of a RSiTE. In summary, our research question is as follows: How can a coding protocol be developed and validated in order to identify the mathematical, didactic, and professional knowledge present in transcripts derived from recorded sessions of a Research Situation for Teacher Education (RSiTE)?

Theoretical framework

The notion of a RSiTE is grounded in the Theory of Didactical Situations (Brousseau, 1997) and, more specifically, in the work developed within the *Maths à Modeler* framework (Gravier & Ouvrier-Buffet, 2022) concerning the design of action, formulation, and validation situations based on problems drawn from discrete mathematics. More precisely, a RSiTE consists of an instance of such a problem, designed on the basis of both a mathematical and didactical *a priori* analysis, with the aim of fostering the acquisition of knowledge related to proof process. This process is primarily approached from a semantic perspective (Weber & Alcock, 2004), viewing proof as a process that involves the modes of thinking identified by Alcock (2008), namely instantiation, creative thinking, and critical thinking. The milieu of a RSiTE therefore includes a fundamental situation intended to support the acquisition of specific mathematical knowledge (Schoenfeld, 1992; Deloustal-Jorrand et al., 2025; Da Ronch et al., 2025). Such knowledge emerges through the exploration of particular cases, the formulation of conjectures, their testing, validation or refutation, and, more broadly, through the progressive construction of mathematical argumentation. However, student teachers can engage as genuine authors of such mathematical activity only if the didactical contract (Brousseau, 1997) enables the devolution of scientific responsibility. For this reason, the milieu of a RSiTE is complemented by a didactical contract grounded in scientific debate within teaching situations (Legrand, 1988; 2001). On the one hand, this contract promotes the emergence of students' conceptions by encouraging individual proposals, positions, and arguments, while the teacher refrains from intervening at the mathematical level and instead focuses on structuring the exchanges. On the other hand, it supports the institutionalization of situational knowledge related to heuristic strategies, validation processes, mathematical language, and error recognition, in addition to the situational knowledge embodied in the mathematical results themselves.

To distinguish between these two categories of mathematical knowledge, we define *Type I situational knowledge* as knowledge concerning mathematical objects and their properties, here represented by the mathematical results associated with the problem. In contrast, *Type II situational knowledge* refers to knowledge related to strategies, definition processes, validation practices, and the use of mathematical language. Although such knowledge is essential to mathematical activity, it is only rarely made explicit as an object of instruction. As a result, it often remains transparent not only to students but also to teachers (Deloustal-Jorrand et al., 2025). This distinction is therefore particularly relevant in the context of teacher education and constitutes a central component in the development of the analytical instrument presented later in this paper.

Furthermore, a RSiTE is designed to foster not only mathematical knowledge but also didactic and professional knowledge. These forms of situational knowledge are problematized through the analysis of the fundamental situation described above: student teachers are invited to identify and reflect upon elements of the milieu they encounter and the didactical contract within which they interact. As a result, notions such as milieu, didactical contract, and didactical variables become grounded in and exemplified by their own experience of engaging in the problem-solving process. It can thus be seen that, unlike a simple didactic situation, a RSiTE operates simultaneously within two nested didactic systems (Houdement, 2013): a School system (SS) and a Teacher Education system (TES). The former concerns the mathematical activity developed around the proposed problem, whereas the latter concerns the analysis of this activity from a professional education perspective.

Students are therefore required to act both as researchers when working on the mathematical problem and as prospective teachers when analysing the didactic phenomena that emerge during the process. This dual participation in two intertwined systems makes it possible, *a priori*, to envisage the simultaneous emergence of: (1) Type II mathematical knowledge, but also Type I mathematical knowledge; (2) didactic knowledge related to the functioning of the milieu and the didactic contract, to choices concerning didactic variables, and to the processes of devolution and institutionalisation; and (3) more broadly, professional knowledge, that is, situational knowledge related to the educational system, official curricular guidelines, or psychosocial dimensions of teaching. Identifying these different types of situational knowledge within RSiTE sessions constitutes a major methodological challenge, which we sought to address. To this end, we developed an analytical framework designed to examine the interactions occurring within a RSiTE, with the aim of evaluating the didactic contract as well as the processes of devolution and institutionalisation of mathematical, didactic, and professional knowledge.

Drawing on the theoretical foundations of RSiTEs, we defined a set of *a priori* categories for classifying the discursive actions observable during the sessions. We briefly present here those categories that derive directly from these theoretical considerations. Category C1 concerns the initiation of the devolution process. Category C3 describes a relatively common didactic contract corresponding to a conversation-based lesson, while Category C4 refers to the implementation of scientific debate during its peer-discussion phase. Category C5 concerns the formulation of Type II mathematical knowledge, Category C6 the formulation of didactic knowledge, and Category C7 the formulation of professional knowledge. Each category is further subdivided according to the author of the discursive action: the teacher educator (TEA) or the students (STU). The categories are defined through a set of indicators, most of which were established *a priori*. Owing to space limitations, only two examples are provided here. The first concerns Category C4 and the STU speaker: the student expresses an opinion regarding a proposal or reports uncertainty or even lack of understanding. The second concerns Category C3 and the TEA speaker: the teacher educator intervenes in the discussion by asking the question “Why?”. In addition, two categories (C2 and C8) were introduced *a posteriori* in order to account for all the discursive actions identified in the corpus.

Method

This RSiTE was implemented with approximately twenty first-year Bachelor students enrolled in a Swiss teacher education institution preparing for primary school teaching, who had not yet received any instruction in mathematics education. Two sessions of 1 hour and 45 minutes were conducted. In accordance with the principles of RSiTEs, students were placed in a research situation and engaged in producing, testing, discussing, and validating conjectures within the framework of a collectively organized scientific debate. The proposed situation was based on a domino tiling problem whose potential to foster proof-related activity has already been documented in previous studies, particularly in primary education (Gravier et al., 2008). All discursive actions were audio-recorded and subsequently transcribed verbatim. These transcripts constitute the corpus analyzed in the present study. The underlying assumption is that the various categories presented in our theoretical framework account for the discursive actions, that is, the discourse elements, occurring within the corpus under investigation. Consequently, the corpus was analyzed using a coding scheme developed both deductively, based on the model presented in the theoretical framework, and inductively, from initial analytical observations described below. This approach made it possible to identify the moments at which specific forms of situational knowledge became explicit topics of discussion.

Segmentation and categorization

In a first stage, segmentation and categorization were carried out independently by three researchers (coders A, B, and C) on two random samples extracted from the transcript. In this study, a sequence is defined as a continuous fragment of discourse produced by a single speaker, whose boundaries delimit an internal segment in which the discourse relates entirely, or predominantly, to one of the themes addressed by the categories defined *a priori*. Accordingly, a sequence may consist of one or several consecutive lines and ends when the topic changes. This initial phase of analysis made it possible to refine both our definition of a sequence and the definitions of the categories established *a priori*. An inductive approach subsequently led to the addition of two categories to the analytical framework, complementing the six categories originally derived from the theoretical framework. Category C2 refers to the orchestration of the sessions, independently of any didactic contract, and applies both to the teacher educator (TEA-C2) and to the students (STU-C2). Finally, Category C8 includes all inaudible or unintelligible discourse produced either by the teacher educator (TEA-C8) or by the students (STU-C8), as well as utterances for which the speaker could not be identified (NA-C8). These two categories therefore relate neither directly to the forms of situational knowledge under consideration nor to the didactic contract.

In a second stage, two of the three researchers (coders A and B) independently segmented the entire corpus using the agreed-upon definition of a sequence in order to assess the stability of this segmentation procedure across the corpus as a whole (Figure 1).

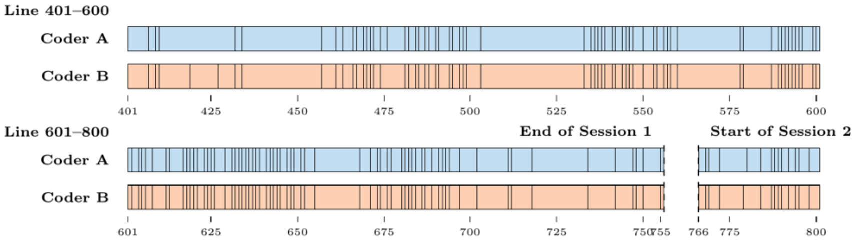

Figure 1: Excerpt from the segmentations produced by coders A and B for lines 401 to 800

The degree of divergence between the segmentation boundaries obtained was assessed using the WindowDiff (WD) coefficient from the field of natural language processing (Pevzner & Hearst, 2002). A consensual version of the segmentation was then established following discussion between the two coders. In a third stage, each sequence was assigned to the category that was predominantly represented within it. Three researchers (coders A, B, and C) performed this task independently before constructing, in a fourth and final stage, a consensual categorization of the entire corpus (Figure 2).

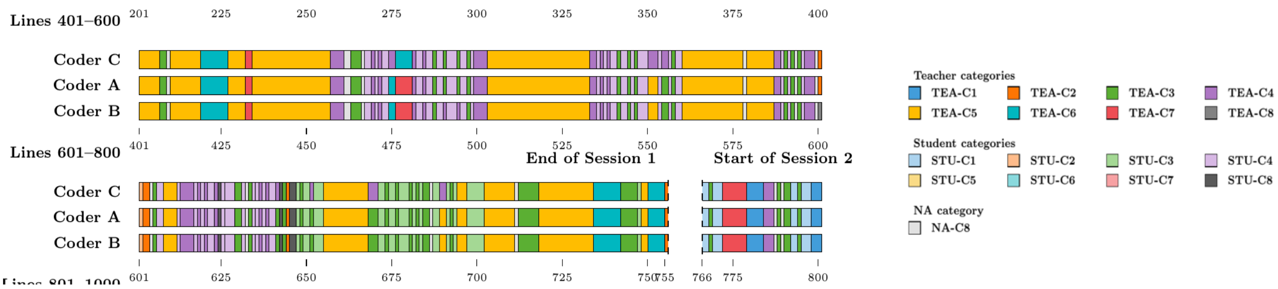

Figure 2: Excerpt from the category coding produced by coders A, B, and C for the sequences corresponding to lines 401 to 800 of the corpus, together with the associated legend

The reliability of the coding scheme was assessed by calculating Cohen's kappa coefficients (Cohen, 1960) for coder pairs (A-C and B-C). Fleiss' kappa was not used because of the potential influence of the coding procedure on the agreement estimates. Specifically, coders A and B had participated in the development of the consensual segmentation, and the segmentation boundaries may consequently have been influenced, at least partially, by the categories assigned to the discourse sequences. The purpose of these analyses was therefore to determine the extent to which the categories defined in the analytical framework could be recognized and applied consistently by independent researchers.

Analysis of the data based on the consensual coding scheme

Once the consensual coding scheme had been established, two levels of quantitative analysis were performed. The first concerned the frequency of the categories within the corpus, measured by the number of sequences assigned to each category (706 sequences in total). The second concerned the relative weight of each category in the discourse, estimated from the number of transcript lines associated with it (1,787 lines in total). This twofold analysis made it possible to characterize not only the frequency of occurrence of the categories identified within a RSiTE, but also the extent to which they contributed to the overall discursive activity observed during the sessions.

Results

Reliability of the coding procedure

Regarding segmentation, the WindowDiff coefficient obtained between coders A and B was approximately 2.07%. This value indicates a low level of disagreement concerning the segmentation boundaries produced independently and suggests that the definition of a sequence was sufficiently explicit to yield convergent and reproducible results. Furthermore, the reliability of the category coding was found to be high. The κ coefficients obtained between the different coders were approximately 0.88 for coders B and C and 0.94 for coders A and C. According to the interpretative thresholds proposed by Landis and Koch (1977), these values correspond to almost perfect agreement. Taken together, these statistical indicators suggest that the categories developed from the RSiTE theoretical framework and the preliminary analyses can be identified with a relatively high degree of stability across different coding processes. They therefore provide initial evidence supporting the reliability of the proposed coding scheme.

Distribution of categories across the corpus sequences

The results (Figure 3) show that all the expected categories were effectively represented in the corpus. The most frequently observed categories corresponded to the didactic organization of the session (TEA-C3: 24.50%), scientific debate among students (STU-C4: 20.11%), scientific debate led by the teacher educator (TEA-C4: 9.49%), and the formulation of mathematical knowledge (TEA-C5: 7.93%). This distribution is consistent with the characteristics of a RSiTE. A substantial proportion of the interactions was devoted to the organization of collective work (TEA-C3) and to mathematical discussions conducted in the form of scientific debates concerning conjectures, arguments, and procedures for validating the solutions obtained to the problem (TEA-C4/STU-C4). On several occasions, these discussions also prompted the teacher educator to make explicit, formulate, or institutionalize mathematical knowledge mobilized during the inquiry process (TEA-C5).

The categories associated with didactic knowledge (C6) and professional knowledge (C7) were less frequent (e.g., TEA-C6: 2.55%; TEA-C7: 1.70%). Nevertheless, their presence in the corpus indicates that the implementation of the RSiTE was not limited to discursive actions related solely to solving a mathematical problem. It also gave rise to moments of making explicite and problematization concerning situational knowledge related to the teaching and learning of mathematics (C6), and more broadly to situational knowledge related to the teaching profession (C7) (Figure 3). An examination of the number of transcript lines associated with each category provides a complementary perspective. Some relatively infrequent categories nevertheless occupied a substantial proportion of the discourse (TEA-C5, TEA-C6, and TEA-C7). For example, TEA-C5 accounted for less than 8% of the sequences but more than one quarter of the corpus lines. This finding suggests that discursive actions belonging to this category tended to be more elaborated and extended, during which the teacher educator explicitly formulated or institutionalized mathematical knowledge emerging from the students' inquiry work (Figure 3).

To a lesser extent, a similar pattern was observed for the categories related to didactic knowledge (TEA-C6) and professional knowledge (TEA-C7) (Figure 3). Although these categories appeared less frequently, they generated relatively lengthy discussions whenever they became objects of explicit reflection (e.g., TEA-C6 represented 2.55% of the sequences but 8.79% of the corpus lines). These results suggest that the frequency with which a category occurs does not fully capture its role within the dynamics of the situation and support the joint examination of both the number of sequences and the amount of discourse associated with each category.

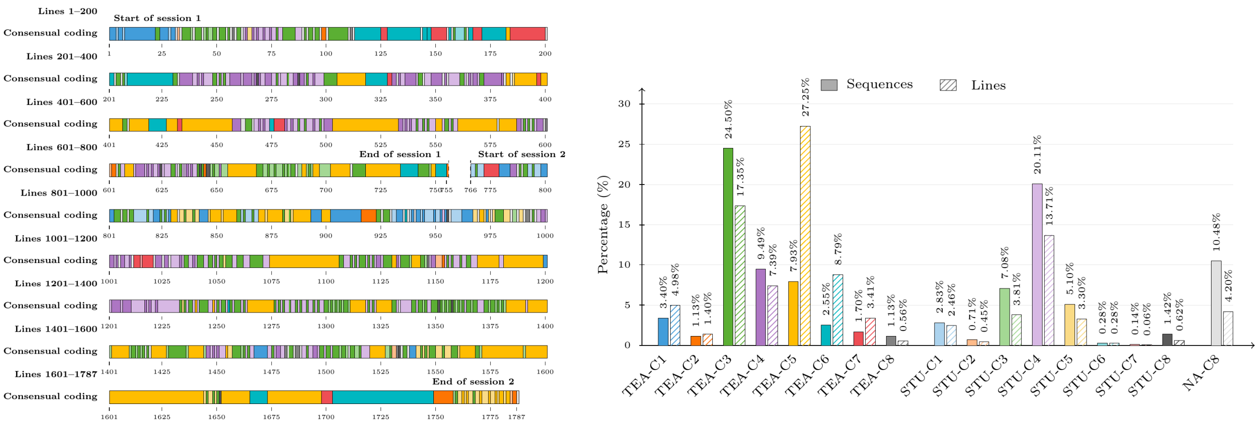

Figure 3: Coding of the dominant category assigned to each sequence of the corpus (left) and distribution of category frequencies based on both the number of sequences and the number of transcript lines (right)

Discussion and conclusion

The aim of this study was to develop and test a coding framework applied to a corpus of transcribed audio data. The purpose of this framework was to identify Type II mathematical knowledge (related to the examination of particular cases, the formulation of conjectures, and similar activities), didactic knowledge (e.g., milieu, didactic variables, and didactic contract), and professional knowledge manifested in the discursive interactions occurring during sessions involving a teacher educator and first-year Bachelor students enrolled in a primary teacher education programme, within the implementation of a Research Situation for Teacher Education (RSiTE).

The results first show that the proposed coding procedure exhibits a high level of reliability, both for segmentation ($WD \approx 2.07\%$) and for categorization ($k_{B-C} \approx .88$ et $k_{A-C} \approx .94$). However, the overrepresentation of certain categories suggests that the reliability of coding reproducibility, assessed through pairwise kappa coefficients, reflects agreement concerning the dominant categories more strongly than agreement concerning categories that are less represented in the analyzed corpus. The findings also reveal the presence of Type II mathematical knowledge, as well as didactic and professional knowledge, which emerge in different ways through the discursive actions of both the teacher educator and the students (Figure 3). For instance, Type II mathematical knowledge is primarily mobilized by students during scientific debate (STU-C4), but occupies a much less prominent place when students are required to make this situational knowledge explicit independently of the problem under investigation (STU-C5). It is precisely at this point that the teacher educator intervenes by making such situational knowledge visible through discursive actions involving its formulation, explicitation, or institutionalization (TEA-C5).

Furthermore, the limited presence of categories STU-C6 and STU-C7 may be explained by the fact that these students had not yet completed any specialized coursework in mathematics education during their first year of study and that their professional knowledge remained relatively limited. It therefore falls to the teacher educator to make this situational knowledge explicit and to problematize it, which may account for the non-negligible proportion of discursive actions associated with categories TEA-C6 and TEA-C7 (Figure 3). These results thus seem to support the assumption that such forms of situational knowledge remain largely transparent to these students. Consequently, it is the responsibility of the teacher educator to render them visible, and RSiTEs appear to constitute a particularly suitable type of situation for achieving this objective.

This study nevertheless remains exploratory. While it establishes the presence of these forms of situational knowledge at a macroscopic level, it does not yet allow for a precise characterization of their content at a microscopic level. An important avenue for future research is therefore to complement this quantitative approach, first, with a systematic qualitative analysis of the targeted categories (e.g., C4, C5, C6, and C7) in order to identify more precisely the situational knowledge that is verbalized through the indicators defined *a priori*; and second, with an examination of the moments at which this situational knowledge becomes salient within the discourse (scientific debate, problem conclusion, or institutionalization processes). Such work would also provide additional evidence regarding the theoretical coherence of the framework and contribute to assessing the construct validity of the proposed analytical tool.

References

Alcock, L. (2008). Mathematicians' perspectives on the teaching and learning of proof. In F. Hitt, D. Holton, & P. W. Thompson (Eds.), *Research in collegiate mathematics education VII* (CBMS Issues in Mathematics Education, Vol. 16, pp. 73–100). American Mathematical Society. https://doi.org/10.1090/cbmath/016/03

Brousseau, G. (1997). *Theory of didactical situations in mathematics: Didactique des mathématiques, 1970–1990* (N. Balacheff, M. Cooper, R. Sutherland, & V. Warfield, Eds. & Trans.). Kluwer Academic Publishers. https://doi.org/10.1007/0-306-47211-2

Chanudet, M., Coppé, S., Gandit, M., & Moulin, M. (2019). Analyse des interactions didactiques dans une perspective d'évaluation formative [Analysis of didactic interactions from a formative assessment perspective]. Dans S. Coppé, É. Roditi, V. Celi, F. Chellougui, F. Tempier, C. Allard, C. Corriveau, M. Haspékian, P. Masselot, S. Rousse, H. Sabra, & M. Kiwan-Zacka (dir.), *Nouvelles perspectives en didactique : Géométrie, évaluation des apprentissages mathématiques* (p. 453–475). La Pensée sauvage.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37–46. https://doi.org/10.1177/001316446002000104

Da Ronch, M., Gandit, M., & Mili, I. (2025a). Modélisation des situations de recherche pour la formation professionnelle des enseignants [Modeling research situations for teachers' professional education]. Dans C. Derouet, V. Durand-Guerrier, C. Lemrich, & A.-C. Mathé (dir.), *Pré-actes du* *colloque international en hommage à l'œuvre de Guy Brousseau* (p. 204–208). Université de Bordeaux. https://hal.science/hal-04975494

Deloustal-Jorrand, V., Gandit, M., & Mesnil, Z. (2025). Spécificités des connaissances en logique et conséquence sur la double discontinuité de Klein [Specific Features of knowledge in logic and their implications for Klein's double discontinuity]. *Recherches en didactique des mathématiques, 45*(3), 239–269. https://doi.org/10.46298/rdm.14819

Gravier, S., & Ouvrier-Buffet, C. (2022). The mathematical background of proving processes in discrete optimization—Exemplification with research situations for the classroom. *ZDM–Mathematics Education, 54*(4), 925–940. https://doi.org/10.1007/s11858-022-01400-3

Gravier, S., Payan, C., & Colliard, M.-N. (2008). Maths à modeler : Pavages par des dominos [Maths à modeler: domino tilings]. *Grand N, 82*, 53–68.

Houdement, C. (2013). *Au milieu du gué : entre formation des enseignants et recherche en didactique des mathématiques* [Midway across the ford: between teacher education and research in mathematics education] (Habilitation à diriger des recherches, Université Paris Diderot - Paris VII, France). https://theses.hal.science/tel-00957166

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174. https://doi.org/10.2307/2529310

Legrand, M. (1988). Rationalité et démonstration mathématiques : Le rapport de la classe à une communauté scientifique [Mathematical rationality and proof: the relationship between the classroom and a scientific community]. *Recherches en didactique des mathématiques, 9*(3), 365–406.

Legrand, M. (2001). Scientific Debate in Mathematics Courses. In: Holton, D., Artigue, M., Kirchgräber, U., Hillel, J., Niss, M., Schoenfeld, A. (eds) The Teaching and Learning of Mathematics at University Level. New ICMI Study Series, vol 7. Springer, Dordrecht. https://doi.org/10.1007/0-306-47231-7_12

Pevzner, L., & Hearst, M. A. (2002). A critique and improvement of an evaluation metric for text segmentation. *Computational Linguistics*, *28*(1), 19-36. https://doi.org/10.1162/089120102317341756

Sackur, C., Assude, T., Maurel, M., Drouhard, J.-P., & Paquelier, Y. (2005). L'expérience de la nécessité épistémique [The experience of epistemic necessity]. *Recherches en didactique des mathématiques, 25*(1), 57–90.

Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense making in mathematics. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning: A project of the National Council of Teachers of Mathematics* (pp. 334–370). Macmillan.

Weber, K., & Alcock, L. (2004). Semantic and syntactic proof productions. *Educational Studies in Mathematics, 56*, 209–234. https://doi.org/10.1023/B:EDUC.0000040410.57253.a1