



HAL
open science

Proof evaluation tasks as tools for teaching?

Kirsten Pfeiffer, Rachel Quinlan

► **To cite this version:**

Kirsten Pfeiffer, Rachel Quinlan. Proof evaluation tasks as tools for teaching?. CERME 9 - Ninth Congress of the European Society for Research in Mathematics Education, Charles University in Prague, Faculty of Education; ERME, Feb 2015, Prague, Czech Republic. pp.178-184. hal-01281065

HAL Id: hal-01281065

<https://hal.science/hal-01281065>

Submitted on 1 Mar 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Proof evaluation tasks as tools for teaching?

Kirsten Pfeiffer and Rachel Quinlan

National University of Ireland, Galway, Ireland, kirsten.pfeiffer@nuigalway.ie

This article reports on our experience, arising from an earlier research study, of incorporating proof evaluation tasks into a university mathematics curriculum. In particular, we discuss a task in which students were asked to evaluate and rank five different proposed proofs of a statement from elementary linear algebra. The students' responses to this task prompted rich learning opportunities on the nature and functions of mathematical proofs, as well as revealing some interesting features of their thinking. We argue that proof evaluation tasks can afford rich learning opportunities as well as enabling novice students to participate in authentic mathematical practice.

Keywords: Proof, proof evaluation, curriculum.

BACKGROUND

A CERME 7 article by Kirsten Pfeiffer (2011a) presents a conceptual schema that provides a frame of reference for consideration of what needs attention in a proof evaluation exercise. In accordance with Hemmi (2008), Pfeiffer regards proof and proofs as *artefacts* of mathematical practice. She adapts ideas of Hilpinen (2004) on evaluation of artefacts in general and specializes them to the case of mathematical proofs. In this context an *artefact* is a (physical or conceptual) object that is designed and made by an author (or authors) in order to fulfil a specific purpose (or purposes). Thus the quality of an artefact can only be judged in terms of its success at achieving its purpose(s). In the case of a proof of a mathematical statement, a primary and non-negotiable purpose is that the argument establishes the truth of the statement. Other purposes might include provision of a satisfying explanation, enhancing understanding of the concepts involved, and so on. Motivated by Hilpinen, Pfeiffer suggests that a proof can be evaluated in relating the three features of an artefact, its *intended character*, its *actual character*, and its *purposes*. Therefore evaluating a proposed proof might involve three considerations:

that the author's intention or "proof design" is appropriately matched to the purpose of the proof, that the author's intention is appropriately realized in the actual written proof, and that the written proof appropriately achieves its purpose. The point of Pfeiffer's schema is to provide some context and terminology for discussion of what proof evaluation entails and for discussion of specific evaluations of particular proofs. It is intended not as a rigid framework but as a helpful theoretical construction.

The outcomes of Pfeiffer's research study (Pfeiffer, 2011b) included strong indications that proof evaluation tasks, including those involving more than one "proof" of the same statement, have the potential to prompt students to consider the mechanism and fitness-for-purpose of a proof in a serious way. Some students in the study even recognised a change in their own thinking stimulated by the task of comparing different proofs of the same statement. These observations encouraged us to include proof evaluation tasks in the curriculum alongside learning activities of other types.

Over the last decade several investigations into students' performances when validating or reading mathematical proofs have shown that students have difficulties in determining whether a proof is valid (Selden & Selden, 2003; Alcock & Weber 2005). Other studies describe the behaviour of experienced mathematicians when validating proofs (Weber & Mejia-Ramos, 2011) or the differences between novice and experienced readers (Inglis & Alcock, 2013). Techniques or teaching methods to improve students' proof comprehension have been suggested, for example *unpacking proofs* or *proof frameworks* (Selden & Selden, 1995), inclusion of instructional sequences in mathematics courses (Stylianides & Stylianides, 2008), *e-proofs* (Alcock & Wilkinson, 2011) or *self-explanation* (Hodds, Alcock, & Inglis, 2014). We consider *proof evaluation exercises* as another possible teaching practice to accomplish proof reading skills. In our experiences

proof validation activities provide a rich teaching and learning tool provoking fruitful discussions and ultimately making a wide range of features and purposes of mathematical proof *visible* to learners. We aim to test the efficiency of proof evaluation exercises incorporated into a University level mathematics course and also to prepare resources for use by teachers.

In this paper we report on one particular proof evaluation exercise performed in a linear algebra course for first year students run by the second author of this article, who is a research mathematician and a lecturer in a university mathematics department. We will show that the students' responses to proposed proofs potentially stimulate a considerable variety of themes to discuss in a teaching/learning environment. As students have engaged with the relevant mathematical context and the suggested proofs in advance, and as they are encouraged to discuss their own feedback rather than experts' proofs and evaluations, students are inclined to participate actively and appreciate these discussions. We will also report on our experiences with the construction of suitable partly flawed 'proofs' and show how Pfeiffer's schema is useful to assure opportunities to highlight various aspects of proof.

EXAMPLE OF A PROOF EVALUATION EXERCISE

The task described below was included in the first written homework assignment in an introductory course on Linear Algebra for first year students. The students were familiar with the concept of a linear transformation of \mathbb{R}^2 as a function that respects addition and multiplication by scalars, and they were familiar with the matrix representation of a linear transformation and with the procedure of using matrix-vector multiplication to evaluate a transformation at a particular point.

The proof evaluation task

Students were presented with the following text.

Alison, Bob, Charlie, Deirdre and Ed are thinking about proving the following statement.

If the function $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is a linear transformation, then T fixes the origin, i.e. $T(0,0)=(0,0)$.

Alison's Proof

Suppose that $T(1,1)=(a,b)$. Then

$$T[(1,1)+(0,0)] = T(1+0,1+0) = T(1,1) = (a,b).$$

But on the other hand since T respects addition

$$T[(1,1)+(0,0)] = T(1,1) + T(0,0) = (a,b) + T(0,0) = (a,b) \text{ from above.}$$

$$\text{So } T(0,0) = (a,b) - (a,b) = (0,0).$$

Bob's Proof

We know that for any element u of \mathbb{R}^2 and for any real number k we have $T(ku) = kT(u)$.

Then applying T to $(0,0)$ and multiplying the result by any real number k must give the same result as multiplying $(0,0)$ by k first and then applying T . But multiplying $(0,0)$ by k always results in $(0,0)$ no matter what the value of k is. So it must be that the image under T of $(0,0)$ is a point in \mathbb{R}^2 which does not change when it is multiplied by a scalar. The only such point is $(0,0)$. So it must be that $T(0,0) = (0,0)$.

Charlie's Proof

Think of T as the function that moves every point one unit to the right. So T moves the point $(0,0)$ to the point $(1,0)$. Then T is a linear transformation but T does not fix the origin. This example proves that the statement is *not* true.

Deirdre's Proof

Suppose that (a,b) is a point in \mathbb{R}^2 for which $T(a,b) = (0,0)$. Then

$$T[2(a,b)] = T(2a,2b) = 2T(a,b) = 2(0,0) = (0,0).$$

Thus $T(2a,2b) = T(a,b)$, so $(2a,2b) = (a,b)$, so $2a = a$ and $2b = b$. Thus $a = 0$, $b = 0$ and $T(0,0) = (0,0)$.

Ed's Proof

Since T is a linear transformation it can be represented by a matrix. Suppose that the matrix of T is

$$\begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

Then the image of $(0,0)$ under T can be calculated as follows:

$$\begin{bmatrix} a & c \\ b & d \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} a0 + c0 \\ b0 + d0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

So $T(0,0) = (0,0)$.

The students were asked the following questions about the text above.

- (a) Does Alison's answer prove that the statement is true? If not, why not?
- (b) Does Bob's answer prove that the statement is true? If not, why not?
- (c) Does Charlie's answer prove that the statement is *not* true? If not, why not?
- (d) Does Deirdre's answer prove that the statement is true? If not, why not?
- (e) Does Ed's answer prove that the statement is true? If not, why not?
- (f) Please rank the five answers in order of your preference (according to your own opinion). Include some comments to explain your ranking.

The five proposed proofs provide a sufficient variety of different approaches to provoke learning and discussion about the nature and features of mathematical proof and about the process of proof evaluation. Alison's proof is sufficient to prove the statement, it actually proves a more general statement. An evaluator might question the unexplained introduction of the point $(1,1)$ and whether there is a reason for this choice. Bob's proof is written in text and also proves a wider statement than required. Charlie mistakenly proposes a counterexample to prove that the statement is incorrect. Deirdre's approach is written in a style which is familiar to students in the context of mathematical proof. However, her proof contains significant logical errors and does not establish the truth of the statement. Using Pfeiffer's terminology, the evaluator may find mismatches between the intended character and purpose of the proof, and between the actual and intended characters. Ed's proof establishes the truth the statement, but other purposes of proof such as enhancing understanding of the content and context of the statement are not met, i.e. the intended character of Ed's proof does not match these wider purposes.

DISCUSSION OF THE RESPONSES AND OPPORTUNITIES FOR DISCUSSION AND LEARNING

The 28 students whose responses are discussed here are those who fully answered all six parts of the question and included comments (many other students answered only some parts or gave "yes/no" answers without explanation). This account is intended to highlight some features of these students' thinking about proof and some opportunities for learning (for both the students and instructor) that arise. We followed the task with a lecture-based discussion session focussing on the themes mentioned below and prompted by the students' work. This session, though conducted with a large group, was notable for the students' interested attention and for an unusually high level of interaction. This may be due to the fact that many of the themes of the discussion arose directly from the students' written comments.

Alison's proof – responses and learning opportunities

Of the 28 respondents, 17 expressed the view that Alison proves that the statement is true. One was non-committal, and the other 10 stated that Alison's answer does not prove that the statement is true.

Five students objected to the introduction of the point $(1,1)$, apparently believing that focussing attention on this chosen point constituted a restriction of the statement to a particular example. Another accepted Alison's proof as correct, but commented:

Student: I would prefer if she used a point (c,d) in \mathbb{R}^2 instead of $(1,1)$.

This last comment prompted a discussion about purposes of proofs. The student approves the proof but suggests altering it to avoid the choice of a particular vector. This alteration may make the argument more accessible for some readers, for example for the five of our students who were misled by the introduction of $(1,1)$ to the extent that they rejected Alison's proof. On the other hand, some readers might see Alison's specification of a particular vector as simplifying the presentation and might prefer this to the alternative of cluttering the text with "general" notation that is not strictly necessary. The comment quoted above gave us the opportunity to highlight the fact that readers may have different mathematical tastes and

that alternative presentations of essentially the same argument may appeal differently to different readers.

Bob's proof – responses and learning opportunities

21 students accepted Bob's proof as correct. One student described it as partially correct, and 6 considered it to be incorrect. It was the second most popular of all the proposed proofs, being ranked first (or joint first) by 8 students.

The six students who rejected Bob's proof stated two reasons for doing so. Two students objected to the assertion in Bob's proof that $(0,0)$ is the only element of \mathbb{R}^2 that "does not change when multiplied by a scalar" pointing out that (for example) $(2,3)$ does not change when multiplied by the scalar 1". This misinterpretation of Bob's intention highlights the importance of precision in mathematical proof.

The other four students who rejected Bob's proof (as well as three who accepted it) complained that it only used part of the definition of a linear transformation, namely the property of respecting scalar multiplication. All seven of these students criticized Alison's proof on the same grounds; the following is a representative comment.

Student: Bob supplies the other half of Alison's proof, he proves the statement for scalar multiplication. He is also half right.

The students who reject or criticize Bob's and Alison's proofs on these grounds appear to recognize the strategy of reasoning from a definition towards a desired conclusion, but their verdict that the argument cannot be complete if it uses only part of the definition seems to be automatic. Their written comments do not indicate attempts to assess the significance to the argument of the "missing" part of the definition; their conclusions appear to be founded purely on an inspection of features of the proof without consideration of its logical structure.

No student cited as a reason to favour the proof of either Bob or Alison that both of these arguments prove a more general statement that they are directly concerned with. Alison's argument proves that every additive function fixes the origin, and Bob's proves that every function that respects scalar multiplication fixes the origin.

In the ensuing discussion, attention was drawn to the logical structure of Bob's proof and to the more general statement that it establishes. Students were reminded that a proof evaluator must consider the full content of what is achieved or omitted in a line of reasoning, and not hastily accept or dismiss an argument on the basis of superficial inspection. From the instructor's point of view, the student responses to Bob's proof highlight the important point that novice students are sometimes more attentive to the internal details of an argument than to its deductive quality.

Charlie's proof – responses and learning opportunities

Charlie's proof was considered incorrect by 25 students, and correct by three. It was ranked last by 21 students.

Of the 25 students who rejected Charlie's proof, only 11 did so on the grounds that the proposed counterexample is not or may not be a linear transformation.

Not all of the remaining 14 students who rejected Charlie's proof gave clear reasons. It is possible that the conflict between Charlie's conclusion and those of the other authors prompted some to object, but only two students cited this as a reason. Six students objected to the restriction of attention to a particular function. There is no sign in the work of these six students of acknowledgement that Charlie's goal is exceptional amongst the five, that he is trying to disprove the statement by exhibiting a counterexample. In the context of Pfeiffer's schema, their evaluations of Charlie's proof do not appear to include consideration of the relationship between the content of the proof and its main purpose.

A possibly surprising feature of the responses to Charlie's argument is that of the three students who considered it to be correct, each also accepted at least one of other four proofs. For example, one commented as follows on Charlie's proof:

Student: As he notes, the linear transformation could possibly move every point one unit to the right. Therefore T does not fix the origin.

The same student accepted (for example) Bob's proof, and recognized the conflict between Bob's and

Charlie's positions, commenting on ranking Charlie's proof 3rd:

Student: Even though Charlie disproves the statement, it's a very valid reason to disprove it.

The three students who approved Charlie's proof did not appear to be troubled by the inconsistency of their own positions, and apparently believed that the statement could simultaneously be validated by a correct proof and contradicted by a counterexample. The intriguing phenomenon of such beliefs is investigated and thoughtfully discussed by Stylianides and Al-Murani (2010).

Our in-class discussion of the responses to Charlie's proof focussed on the exceptional character of his argument among the five, on the roles of examples and counterexamples in mathematical reasoning, and on the inappropriateness of rejecting an argument solely on the grounds that it consists of a single example, without considering what it claims to establish. The opportunity arises also to discuss the question of whether a statement which has a valid proof can ever admit "exceptions", a question whose answer seems to be less clear to inexperienced students than to practising mathematicians.

Deirdre's proof – responses and learning opportunities

Deirdre's proof was considered correct by 19 students, and ranked 1st or 2nd by 11 of these. It was considered incorrect by 8 students, with one reporting no verdict.

Deirdre's argument begins with a linear transformation T and a hypothesized point (a,b) whose image under T is the origin. (There is no *a priori* guarantee that such a point exists.) From the fact that T respects multiplication by scalars, it is established that (a,b) and $(2a,2b)$ have the same image under T . It is then erroneously deduced that these two points must be the same and hence that $a=b=0$. This is a specific error in the line of reasoning documented in Deirdre's argument. The argument also suffers from a structural error in its logic: what it attempts to establish is not that $T(0,0)=(0,0)$ for every linear transformation T , but that if a point is mapped by a linear transformation to $(0,0)$, then that point must be $(0,0)$. In the context of the schema of Pfeiffer, this error corresponds to an opposition between the author's proof strategy

(the *intended character* of her proof) and her purpose (establishing that $T(0,0)=(0,0)$ for a linear transformation T). The "internal" error in Deirdre's proof, (that $T(a,b)=(0,0)=T(2a,2b)$ means $(a,b)=(2a,2b)$) corresponds to a failure in the author's implementation of her strategy, a mismatch between the *intended character* and *actual character* of her proof. That such an error must exist is inevitable in this instance, since the author's intention is to prove an untrue statement.

Obviously notable is the fact that two-thirds of the students accepted an argument that has (at least) two serious flaws, one in its overall logical structure and one in its internal deductions. Many identified similarities between Deirdre's proof and Bob's, which may partly explain their readiness to accept this fundamentally flawed proof.

The 8 students who rejected Deirdre's proof did so for a variety of reasons. Two of them (as well as two who accepted the proof) criticized the use of the scalar 2 instead of a general k . Two of these students suggested that this specialization amounted to restriction to a special case and constituted a reason to reject the proof, the other two only that it compromised the quality of the argument (as opposed to its correctness).

Two students rejected Deirdre's proof on the basis of the erroneous deduction that $T(a,b)=T(2a,2b)$ means $(a,b)=(2a,2b)$. For example,

Student: Her second line contains a mistake, when she states $T(2a,2b)=T(a,b)$ $(2a,2b)=(a,b)$. This is not necessarily true. She is incorrect.

For us, the most remarkable feature of the data on our students' responses to Deirdre's proof is that not one student noted its significant logical flaw. The only possible reference to the unexpected structure of Deirdre's argument is an oblique one from a student who accepted the proof and commented:

Student: she works backwards to reach her conclusion.

From their comments it is not evident that any of the students gave careful critical attention to the question of "fitness-for-purpose" of Deirdre's strategy. In the context of Pfeiffer's schema, none of the students' written comments indicate consideration of the rela-

relationship between the intended or actual characters of this argument and its purpose. A key learning outcome for the instructor here is that the mathematician's practice of constantly testing the connection between the text of an argument and the statement that it purports to prove is not automatically adopted by students. Our discussion on Deirdre's proof focussed on this mental discipline and its essential role in mathematical practice and in the development and validation of mathematical knowledge. The validity of a mathematical argument cannot be assessed without analysis of the deductive process from the hypotheses to the conclusion. To conduct such analysis, a reader of proofs needs to have a measure of confidence in her ability to extract the logical thread from a passage of text, and to assess whether it does what it claims. As students progress through mathematical education at university, we expect their sense of their own reliable mathematical authority to evolve. Alertness to the possibility of logical failures in an argument is a habit of mind whose development may need explicit attention from both teachers and students. It is a key feature of mathematical practice, which might plausibly be encouraged by critical study of proofs, including some that are incorrect or inadequate in different ways.

Ed's proof – responses and learning opportunities

Ed's proof was accepted as correct by all 28 students, and was by far the most popular of the five proofs, being ranked 1st by 18 students and 2nd by 5 students.

Few students commented in detail on Ed's proof. Typical remarks included that it was clear, simple, short and easy to understand. Overall the group demonstrated a clear preference for Ed's translation of the problem into an easy matrix calculation over Alison and Bob's processes of reasoning from the defining properties of a linear transformation. The matrix representation of a linear transformation had been discussed in detail in lectures, and manipulations with matrices featured in several other tasks on the homework assignment that included this proof evaluation exercise.

Our discussion of Ed's proof and the responses to it focussed on the wider purposes of proof and on the reasons that a reader might have for preferring Alison or Bob's proof to Ed's, despite the fact that more effort is required to understand them. Students were invited

to consider whether any of these proofs enhanced their appreciation of the significance of the defining properties of a linear transformation, or their understanding of *why* the statement is true.

CONSTRUCTION OF PROOF EVALUATION TASKS

Composing a suitable collection of "proofs" for a proof evaluation task can be an absorbing but time-consuming challenge for an instructor. It is not essential that such a task involves multiple proposed proofs, but our experience suggests that the invitation to compare different attempts to prove the same statement can stimulate meaningful learning opportunities. A first step in constructing a task of the type described here is to identify a statement is relevant to the disciplinary learning context, and admits different proofs that students have the requisite knowledge to understand. In the preparation of "proofs", there are at least two areas of potential scope for variability. One is the manner in which the proof is presented – whether it primarily consists of text or of algebraic formulation; whether it includes diagrams, either as a support or as the main content; whether the style of text content is formal and technical or more conversational. The presentation style can often be varied independently of considerations of the correctness of the proofs, and we have found it useful to give different styles of writing and presentation to our fictitious authors.

Another important dimension of variability is in the nature of errors or imperfections that appear in the range of proposed proofs. In this context the schema of Pfeiffer provides a useful framework for preparation of variously erroneous proof attempts. A task designer might decide to include one or more "proofs" in which the author's intention is mismatched to the stated aim, for example because of inappropriate logical structure (as in Deirdre's proof), inadvertent restriction to special cases, or unjustifiable deductions. A proof whose intended character is appropriate for the purpose but poorly conveyed in the actual character might also be included. Such a mismatch might be manifested through an insufficiently explained (but justifiable) line of reasoning, through the omission of some routine but necessary ingredient, or through imprecise or unclear statements. Proof evaluation tasks are flexible and adaptable and a number of degrees of freedom exist for their design. Instructors wishing to include such tasks in curricula will find

opportunities to highlight essential points relating both to the nature and purposes of proofs and to relevant disciplinary knowledge.

A shared repository of adaptable proof evaluation tasks relating to different subject areas and levels would be a very useful resource.

CONCLUDING REMARKS

As anticipated by the research study of Pfeiffer (2011b), our incorporation of the activity of proof evaluation into a linear algebra course led to positive learning opportunities for our students as well as giving us some insights into their thinking about proof. We were surprised by the range of discussion opportunities that arose from students' responses to the task. Examples include the importance of precision, the role of counterexamples, and attention in proof-reading to overall structure as well as internal features. Proof evaluation activities with an advanced course in group theory have been similarly encouraging. All of these experiences motivate us to extend our resources for the use of proof evaluation tasks and to conduct comprehensive tests of their effectiveness for the development of proof-reading skills.

As a further argument for the incorporation of tasks of this nature in learning activities, we remark that a great deal of the professional activity of research mathematicians is concerned, directly or indirectly, with the validation and evaluation of proofs. However, in our experience it is rarely the subject of explicit attention in curricula. We propose that proof evaluation tasks, as well as providing meaningful opportunities in teaching and learning, also provide opportunities for students at all stages of expertise to participate in authentic mathematical practice.

Finally, we remark that our classroom experience with proof evaluation tasks demonstrates the potential of collaboration between researchers in mathematics education and academic mathematicians to deliver rich learning opportunities for students. Cooperation and mutual support of this nature is essential if insights arising from research in mathematics education are to have a significant impact on curricula and on the learning of mathematics at university level.

REFERENCES

- Alcock, L.J., & Wilkinson, N. (2011) [e-Proofs: Design of a resource to support proof comprehension in mathematics](#). *Educational Designer*, 1(4).
- Hemmi, K. (2006). *Approaching proof in a community of mathematical practice*. Doctoral Dissertation. Department of Mathematics, Stockholm University.
- Hilpinen, R. (2004). Artifact. *Stanford Encyclopedia of Philosophy*.
- Hodds, M., Alcock, L., & Inglis, M. (2014). Self-explanation training improves proof comprehension. *Journal for Research in Mathematics Education*, 45, 62–101.
- Inglis, M., & Alcock, L. (2012). Expert and novice approaches to reading mathematical proofs. *Journal for Research in Mathematics Education*, 43, 358–390.
- Pfeiffer, K. (2011a). A schema to analyse students' proof evaluations. In *Proceedings of the Seventh Congress of the European Society for Research in Mathematics Education (CERME 7)*. February 9th - 13st 2011, Rzeszów (Poland).
- Pfeiffer, K. (2011b). *Features and purposes of mathematical proofs in the view of novice students: observations from proof validation and evaluation performances*. Doctoral dissertation, National University of Ireland, Galway. NUI Galway. <http://hdl.handle.net/10379/1862>.
- Selden, A., & Selden, J. (1995). Unpacking the logic of mathematical statements. *Educational Studies in Mathematics*, 29(2), 123–151.
- Selden, A., & Selden, J. (2003). Validations of proofs written as texts: Can undergraduates tell whether an argument proves a theorem? *Journal for Research in Mathematics Education*, 34(1), 4–36.
- Stylianides, A., & Al-Murani, T. (2010). Can a proof and a counterexample coexist Students' conceptions about the relationship between proof and refutation. *Research in Mathematics Education*, 12 (1), 21–36 .
- Stylianides A., & Stylianides G. (2008). Enhancing undergraduates students' understanding of proof. In *Proceedings of the 11th Conference for Research in Undergraduate Mathematics Education* ([http://sigmaa.maa.org/rume/crume2008/Proceedings/Stylianides&Stylianides_LONG\(21\).pdf](http://sigmaa.maa.org/rume/crume2008/Proceedings/Stylianides&Stylianides_LONG(21).pdf)), San Diego, CA.
- Weber, K., & Mejia-Ramos, J. (2011). Why and how mathematicians explore proofs: An exploratory study, *Educational Studies in Mathematics*, 76, 329–344.