Komatsu, K., Yamazaki, M., Fujita, T., Jones, K., & Sue, N. (2018). Secondary school students' appraisal of mathematical proofs. *Proceedings of PME42* (Vol. 3, pp. 227-234). Umeå, Sweden: PME.

SECONDARY SCHOOL STUDENTS' APPRAISAL OF MATHEMATICAL PROOFS

Kotaro Komatsu[1], Miho Yamazaki[2], Taro Fujita[3], Keith Jones[4], and Naoki Sue[1]

[1] Shinshu University, Japan; [2] University of Tsukuba, Japan;
[3] University of Exeter, UK; [4] University of Southampton, UK

Research on the reading of proofs is an important area of proof research in mathematics education. As one aspect of the reading of proofs, we focus on 'proof appraisal' by students (that is, students' judgements about given proofs) and explore how students appraise different proofs of an identical statement. Using a simple proof and a generalisable proof of a statement, we analysed the results of a questionnaire completed by 39 Grade 8 secondary school students (13–14 years old). We show aspects of each proof that were appraised by the students, such as simplicity, and the relativeness of their proof appraisals. An implication is a possible 'gap' between the 'mathematical value' appreciated by students and that by researchers and teachers.

INTRODUCTION

Research on the reading of proofs has recently gained more attention in the mathematics education community, and there have been several types of recent studies in this research area (Komatsu et al., 2017). One type relates to students' comprehension of given correct proofs, such as whether students can understand key terms and statements in the proofs and illustrate a sequence of inferences with a specific example (e.g. Mejia-Ramos et al., 2012). Another type of study is of proof validation with students asked to determine the validity/invalidity of the purported deductive proofs (e.g. Inglis & Alcock, 2012). In this paper, we focus on another type of research in the reading of proofs, namely *proof appraisal*. We use this term to refer to students explaining their reasons for preferring a particular given proof.

Some existing studies have examined whether students appreciate certain aspects of proofs, such as verifying that statements are true and explaining why statements are true (e.g. Healy & Hoyles, 2000; Segal, 1999). However, research on the functions of proofs shows that the power of proofs is not only in the verification and explanation of statements (de Villiers, 1990; Hanna & Barbeau, 2008). Recently, Inglis and Aberdein (2014, 2016) have classified mathematicians' appraisals of proofs into four dimensions: aesthetics, intricacy, precision, and utility. Although mathematicians and students are different in various ways, such as mathematical maturity and interest in mathematics, it is anticipated that students' proof appraisals are also likely to be diverse.

To this end, we consider a specific setting where students are given multiple valid proofs of an identical statement and are asked to judge these proofs. This setting is different from those of existing studies that have contrasted valid proofs with insuffi-

2018. In E. Bergqvist, M. Österholm, C. Granberg, & L. Sumpter (Eds.). *Proceedings of the 42nd Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 227-234). Umeå, Sweden: PME.

cient arguments (e.g. empirical arguments) and with purported deductive proofs that actually include errors (e.g. Healy & Hoyles, 2000; Inglis & Alcock, 2012). We use multiple valid proofs of an identical statement because the given multiple proofs have in common the capability of verifying that the statement is true, and thus we expect it to be possible to elicit students' appraisals of various aspects of proofs other than the verification of the statement. Hence, in this paper, we address the following research question: How do students appraise different proofs of an identical statement?

FRAMEWORK: COHERENCE AND GENERALISABILITY OF PROOF

As a framework for classifying proofs that are different in terms of how they draw the conclusion of the statement, we employ the notions of *direct fit* and *familial fit* suggested by Raman-Sundström and Öhman (in press). *Direct fit* here refers to the relationship between a statement and a proof, while *familial fit* refers to the relationship between a proof and a family of proofs. In this paper, we focus on the notion of *coherence*, one aspect of direct fit, and that of *generalisability*, one aspect of familial fit. A proof is regarded as *coherent* if the proof is stated in the same terms as the statement that the proof addresses. A proof is regarded as *generalisable* if the idea of the proof can be used for a larger class of statements.

We explicate these notions by taking a statement as an example: the sum of the interior angles of a star octagon is 720° (this statement was also used in our questionnaire whose data are examined in this paper). Here, a star polygon is defined as a polygon constructed by connecting vertices while skipping the adjacent vertex. Figure 1 shows a star heptagon. Star polygons where the numbers of the vertices are odd (hereafter, star-odd polygons) can be drawn in one stroke, whereas star polygons where the numbers of the vertices are even (star-even polygons) cannot, but can be drawn by a combination of two polygons. Figure 2 shows two different proofs of the statement that the sum of the interior angles of a star octagon is 720°.

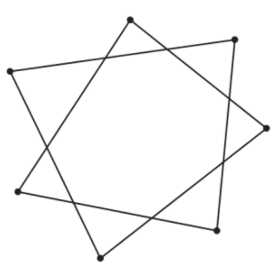
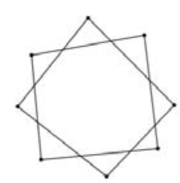


Figure 1: Star heptagon

- Proof A:
The interior angle sum of a quadrilateral is 360°. We multiply it by two because a star octagon consists of two quadrilaterals. Thus, the interior angle sum of a star octagon is 720°.



- Proof B:
There are eight triangles on the outside, and the interior angle sum of these triangles is 1440° (= 180° × 8). The white- and black-marked angles are not the interior angles of the star octagon, and these angles constitute two sets of the exterior angle sum of the 'inside octagon', namely 720° (= 360° × 2). Hence, the interior angle sum of a star octagon is 720° (= 1440 – 720).

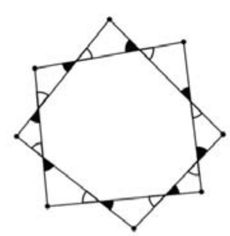


Figure 2: Two proofs of the identical statement

Proof A can be regarded as *coherent* because it uses only interior angles that are stated in the statement, whereas Proof B lacks *coherence* because it introduces the concept of exterior angles, which is not mentioned in the statement. With respect to *generalisability*, Proof B is *generalisable* because the idea of this proof (based on the constancy of the sum of exterior angles in any polygon) can be applied to all star polygons. For example, it is possible to prove that the sum of the interior angles of a star heptagon is 540° by calculating '$180 \times 7 - 360 \times 2$'. On the other hand, Proof A relatively lacks *generalisability* because the idea of this proof is not applicable to the star-odd polygon case. Note that the discussion here is just an example; some proofs may have both coherence and generalisability, and other proofs may have neither.

METHODS

Background

The study reported in this paper consisted of two parts; one part involved designing and implementing a series of proof tasks related to star polygons over four 50-minute lessons in a secondary school classroom (Komatsu et al., in press), and the other part involved conducting a questionnaire after the lessons to investigate how the students appraised two proofs of the aforementioned statement about the star octagon case. This paper presents the results of the second part. In a later section, we briefly describe the implemented lessons because the results of the questionnaire would be influenced by the features of those lessons.

Questionnaire and participants

We produced a questionnaire presenting two proofs almost identical to Proof A and Proof B shown in Figure 2 and then asking, "Which method do you use for finding the sum of the interior angles of a star octagon, Proof A or Proof B? Describe the reason for your choice". We decided to ask such a question because we expected that if students were requested to select either of the two proofs with explaining the reasons for their choices, they would express their appraisal of the advantages of the two proofs more explicitly.

This questionnaire was implemented with 39 eighth-grade students (13–14 years old) in a Japanese lower secondary school affiliated with a national university. The mathematical capabilities of the students were above average for Japan according to their teacher (the fifth author of this paper). The students had covered in class the knowledge necessary to understand the two proofs (e.g. the interior/exterior angle sum theorems of polygons).

Procedure of data analysis

Although our questionnaire was based on a single task, we obtained rich data where the students fully explained the reasons why they preferred one proof to the other proof. Hence, we analysed the students' responses in a qualitative way, coding their responses and then counting the number of students referring to each code in order to investigate what aspects of each proof tended to be appraised by the students. Our coding procedure was as follows. The first author split each of the students' descriptions of their proof appraisals into several segments; there were 97 segments in total for the 39 students (which shows the richness of our data). Temporal codes were then devised to denote these segments. After that, if certain codes were found to be similar, they were unified into a single code. The second author then checked the appropriateness of the coding, and any discrepancies were discussed until the authors reached a consensus (see Tables 2 and 3 for identified codes and their distributions).

Lessons implemented before the questionnaire

As mentioned earlier, our questionnaire was conducted after the implementation of a series of tasks about star polygons over four lessons. Because space here is limited, we give only a brief summary of the implemented lessons, in Table 1 (for more details, see Komatsu et al., in press). The students explored the sums of the interior angles of various star polygons in the lessons. As shown in Table 1, they had had the experience of constructing both Proof A and Proof B before the questionnaire. In particular, in the fourth lesson, they had recognised the generalisability of Proof B, where they found that this proof idea could be applied to all the star polygons. They also invented an algebraic expression for the interior angle sum of a star polygon: $180n - 720 = 180(n - 4)$ (where n is the number of the vertices of the star polygon).

Lesson	Star polygons investigated in each lesson
1st lesson	Star pentagon
2nd lesson	Star-even polygons (e.g. constructing Proof A for a star octagon)
3rd lesson	Star-odd polygons (considering 'outside triangles' and 'inside polygons' like Proof B)
4th lesson	Star-even polygons revisited (e.g. constructing Proof B for a star octagon) and then star polygons in general

Table 1: Summary of the implemented lessons

RESULTS

In the questionnaire, 27 students selected Proof A and 12 students selected Proof B. As a result of coding their proof appraisals, we found that most of the codes could be divided into two types. The first type showed what aspects of each proof the students appraised, while the second type represented the relativeness of the students' proof appraisals (e.g. whether they thought that their proof choices depended on situations). Below, we show the obtained results by type. English translations of the students' responses are rendered from the original Japanese by the authors.

Students' reasons for their proof appraisals

Table 2 shows our classification for the first type of codes, showing what aspects of each proof the students appraised and how many students referred to those aspects. For each proof, the sum of the numbers for all codes is larger than the number of students who selected the proof because there were cases where the proof appraisal by a given student was related to multiple codes.

	Selecting Proof A (n = 27)		Selecting Proof B (n = 12)	
Code	#	% in Proof A	#	% in Proof B
Simple	20	74%	2	17%
Brief	6	22%	1	8%
Understandable	4	15%	2	17%
Free from error	4	15%	0	0%
Immediate	3	11%	0	0%
Generalisable	0	0%	8	67%
Advantage of formula	0	0%	4	33%

Table 2: Students' reasons for their proof appraisals

For the selection of Proof A, the most frequent code is *simple* (indicating that the description of the proof is mathematically simple). This code is also related to other codes, such as *brief* (which means that the proof requires only a single calculation). Below are examples of students' responses for each of these (we use parentheses to show codes assigned to each response):

> S1: Because there are two polygons, if the sums of the interior angles of these polygons can be found, it can be easily solved. (*simple*)

> S2: Proof A does not require complicated calculations, and the answer can be found with a single calculation. (*simple* and *brief*)

As can be seen above, students selecting Proof A considered this proof to be simple and brief because it required only a property well-known by the students (the interior angle sum of a quadrilateral) and a single calculation. These students focused on a specific case mentioned in the questionnaire (the star octagon).

Students choosing Proof B had a different viewpoint, in which they took other star polygons into consideration; the most common reason for the selection of this proof is thus represented by the code *generalisable*. This code is also related to the code *advantage of formula*:

> S3: If we know that the sum of the interior angles of a triangle is 180°, calculate 180 × 8 since 180 × star octagon. Because the sum of exterior angles is always 360°, this method can be used for all cases. (*generalisable*)

> S4: $180(n-4)$ -> $180 \times 4 = 720$. The formula is easy. It has applicability, it can be used for odd cases, and if we use the formula, other problems can be solved as well. Dividing into odd and even cases is bothering. (*advantage of formula, generalisable*, and *simple*)

As shown earlier, the idea represented in Proof B can be generalised to all star polygons, and students choosing this proof appreciated this generalisability. In the lessons implemented before the questionnaire, the students found that the sum of the interior angles of a star polygon can be expressed as $180(n-4)$ (see the methods section). Some students, such as S4, mentioned the advantage of using this algebraic expression.

Relativeness of students' proof appraisal

The second type of codes, which represent the relativeness of the students' appraisals of the two proofs, is summarised in Table 3.

	Selecting Proof A (n = 27)		Selecting Proof B (n = 12)	
Code	#	% in Proof A	#	% in Proof B
Depending on situation	9	33%	1	8%
Appreciation of the other proof	3	11%	1	8%
Limitation of the selected proof	2	7%	0	0%
Criticism of the other proof	2	7%	6	50%

Table 3: Relativeness of students' proof appraisals

One code in this type is *depending on situation*: nine students selecting Proof A stated that they would choose Proof B if the number of the vertices of the star polygon had been different. Other relevant codes are *appreciation of the other proof* and *limitation of the selected proof*: some students choosing Proof A mentioned the value of Proof B as well as the limitations of Proof A:

> S5: When finding the sum of the interior angles of a star polygon in future, I will use Proof A in the case where the number of vertices is even, and Proof B in the odd case. (*depending on situation*)

> S6: When finding the sum of the interior angles of a star polygon, I feel that Proof B is good as it can be used for odd and even cases. For the even case where it is obvious that polygons overlap, I want to use Proof A. (*appreciation of the other proof*)

> S7: Although Proof A can be used only for the case where the number of vertices is even, it is simple, and thinking and calculation are easy. (*limitation of the selected proof* and *simple*)

Students can be regarded as relatively appraising Proof A if they referred to the code *depending on situation*, *appreciation of the other proof*, or *limitation of the selected proof*. This is because these students not only appraised Proof A, but also recognised the limitations of Proof A and the advantage of Proof B. In the questionnaire, 13 students selecting Proof A (48% of that group) appraised it relatively.

On the other hand, half of the students selecting Proof B explicitly criticised Proof A, making *criticism of the other proof* a common code for Proof B but not Proof A:

> S8: The star octagon case can be solved with the method of Proof A, but the star heptagon case etc. cannot be solved with the method of Proof A. The method of Proof B can be commonly used for all of star polygons, so I will use the method of Proof B. (*criticism of the other proof* and *generalisable*)

In relation to this, only one student referred to the code *depending on situation* for Proof B, and the same is the case for *appreciation of the other proof*. Thus, students selecting Proof B tended to appraise this proof absolutely rather than relatively.

DISCUSSION

In this paper, we have examined how secondary school students appraised two different proofs of the same statement. To this end, we employed Raman-Sundström and Öhman's (in press) notions of direct fit and familial fit to prepare two contrasting proofs. In the implemented questionnaire, the most common reasons for selecting Proof A and Proof B were respectively simplicity and generalisablity. Although more students preferred Proof A to Proof B, proof appraisals by almost half of the students selecting Proof A were relative, indicating that they recognised the limitations of Proof A and the value of Proof B in terms of generalisability.

Our findings may raise an issue for mathematics teachers and mathematics education researchers. Proof B can be generalised to all star polygons, and generalisation is much appreciated in the mathematics education community (e.g. Mason, 2002). Generalisable proofs, or proofs that can be used for different purposes, are highly evaluated in mathematicians' practice as well (Hanna & Barbeau, 2008; Weber & Mejia-Ramos, 2011). However, in our study, when asked to select either the simple proof or the generalisable proof, the students tended to prefer the former. This may relate to students' emerging mathematical values (Seah, 2016). It may be that there is a 'gap' between the type of mathematical value appreciated by students aged 13–14 years old and that by teachers and researchers (mathematics education researchers and mathematicians).

That said, a note of caution is that this 'gap' may have arisen from the specificity of our questionnaire where the students were shown a single case (the star octagon) and were asked to select a proof only for this case. In fact, as mentioned above, there were students who preferred the simple proof (Proof A) and, at the same time, appreciated the generalisability of Proof B. Hence, it would be necessary to explore further the gap found in this study by asking different types of questions and adopting different methodologies.

While our questionnaire was based on a single task and implemented with only a relatively small number of students, and, as such, we do not intend to assert the generalisability of all our results, several of the codes that we devised for representing students' proof appraisals may be useful beyond our study. Inglis and Aberdein (2014, 2016) classified mathematicians' proof appraisals, and intricacy (its opposite) and utility in their classification are respectively related to simplicity and generalisability among our codes. Given that simplicity and utility are observed in studies involving different groups (students and mathematicians), these codes, on the one hand, may likely be employed to represent proof appraisals in general. On the other hand, other codes used here (e.g. depending on situation and appreciation of the other proof) are probably best considered as being specific to our study, derived from the specific question in our questionnaire where the students were shown multiple valid proofs of a statement and were asked to show their preferences. Thus, these codes may be useful for capturing students' proof appraisals in similar contexts.

References

De Villiers, M. (1990). The role and function of proof in mathematics. *Pythagoras, 24*, 17–24.

Hanna, G., & Barbeau, E. (2008). Proofs as bearers of mathematical knowledge. *ZDM – The International Journal on Mathematics Education, 40*(3), 345–353.

Healy, L., & Hoyles, C. (2000). A study of proof conception in algebra. *Journal for Research in Mathematics Education, 31*(4), 396–428.

Inglis, M., & Aberdein, A. (2014). Beauty is not simplicity: An analysis of mathematicians' proof appraisals. *Philosophia Mathematica, 23*(1), 87–109.

Inglis, M., & Aberdein, A. (2016). Diversity in proof appraisal. In B. Larvor (Ed.), *Mathematical cultures: The London Meetings 2012–2014* (pp. 163–179). Basel, Switzerland: Birkhäuser.

Inglis, M., & Alcock, L. (2012). Expert and novice approaches to reading mathematical proofs. *Journal for Research in Mathematics Education, 43*(4), 358–390.

Komatsu, K., Fujita, T., Jones, K., & Sue, N. (in press). Explanatory unification by proofs in school mathematics. *For the Learning of Mathematics*.

Komatsu, K., Jones, K., Ikeda, T., & Narazaki, A. (2017). Proof validation and modification in secondary school geometry. *Journal of Mathematical Behavior, 47*, 1–15.

Mason, J. (2002). Generalisation and algebra: Exploiting children's powers. In L. Haggarty (Ed.), *Aspects of teaching secondary mathematics: Perspectives on practice* (pp. 105–120). London, England: Routledge Falmer.

Mejia-Ramos, J. P., Fuller, E., Weber, K., Rhoads, K., & Samkoff, A. (2012). An assessment model for proof comprehension in undergraduate mathematics. *Educational Studies in Mathematics, 79*(1), 3–18.

Raman-Sundström, M., & Öhman, L. D. (in press). Mathematical fit: A case study. *Philosophia Mathematica*.

Seah, W. T. (2016). Values in the mathematics classroom: Supporting cognitive and affective pedagogical ideas. *Pedagogical Research, 1*(2), Article No: 53.

Segal, J. (1999). Learning about mathematical proof: Conviction and validity. *Journal of Mathematical Behavior, 18*(2), 191–210.

Weber, K., & Mejia-Ramos, J. P. (2011). Why and how mathematicians read proofs: An exploratory study. *Educational Studies in Mathematics, 76*(3), 329–344.