Empirical Study of Mathematical Investigation Skill on Graph Theory

Karunia Eka Lestari*, Mokhammad Ridwan Yudhanegara

Departement of Mathematics Education, Universitas Singaperbangsa Karawang, Indonesia

karunia@fkip.unsika.ac.id*, mridwan.yudhanegara@staff.unsika.ac.id

Abstract: *Graph theory allows the student to work on problems that require imagination, intuition, systematic exploration, conjecturing, and reasoning. It implies that mathematical investigation skill is essential to be proficient in Graph Theory. In this study, we conduct empirical research that deals with associational research. There were 97 students selected purposively from sophomore students in the Discrete Mathematics course offered by one of the mathematics education departments in Indonesia. The empirical evidence was analyzed to explain students' thinking behavior on Graph Theory by discovering the association structure between prior knowledge and mathematical investigation skill, then visually depicting its association using the k-means clustering procedure and correspondence analysis. Since in our department, we expect certain prior skills, then this visualization could be used as self-reflection for our department, whether we have gained the results as expected to strengthen certain investigation approaches. Generally, the study concludes some findings that provide novelty and open issues for future research to develop a learning environment supporting mathematical investigation activities.*

Keywords: Correspondence analysis, graph theory, mathematical investigation

INTRODUCTION

Over the last two decades, the mathematics education society has strongly emphasized the essential of investigation-based learning environments (Da Ponte & Pedro, 2007, Leikin, 2014, Yerushalmy, 2009). It suggests that mathematical investigation is an essential skill from an educational point of view—without exception from the Graph Theory point of view as one subject in the Discrete Mathematics course offered by the mathematics education study program in Indonesia. Considering Graph Theory allows the student to work on problems that require imagination, intuition, systematic exploration, conjecturing, and reasoning.

Taylan and Da Ponte (2016) recognize that mathematical investigation as a special form of problem-solving that role in (1) stimulates student engagement for meaningful learning; (2) provides multiple mathematical activities for students at different ability grades; and (3) stimulates holistic thinking that relates a basic condition to many topics for valuable mathematical reasoning.

It implies that a mathematical investigation provides a good context for making the student understand the need to justify their assertions or explain their reasoning.

Barbeau and Taylor (2009) viewed mathematical investigation as a skill required to solve challenging mathematical tasks or problems. It refers to the most inclusive skill that considers different mathematical situations, conjecturing, justifying, explaining the conjectures, proving, inferring, and posing a new question (Leikin, 2007). Particularly, Leikin (2014) argues that conjecturing is the main aspect of mathematical investigation. Mariotti and Pedemonte (2019) state that a conjecture is a statement that is strictly connected to an argument and a set of conceptions where the statement is potentially true because some conceptions allow the construction of an argument that justifies it. However, constructing a conjecture involves a lot of cognition processes, such as organizing and recording data, pattern searching, conjecturing, inferring, justifying, and explaining conjecture (Astawa et al., 2018, Benson et al., 2004, Yeo, 2017). This cognitive process is associated with prior mathematical knowledge.

Some previous studies have been conducted, focusing on developing students' investigation abilities (da Ponte & Pedro, 2007; McCosker & Diezmann, 2009; Quinnell, 2010; Yeo, 2017; Galen & Eerde, 2018). However, no one has studied how mathematical investigation is associated with prior knowledge, even though it is the conceptual foundation for developing mathematical investigation skills. Therefore, this study intends to discover the association structure between prior knowledge and mathematical investigation skills and then visually depict its association using the k-means clustering procedure and correspondence analysis.

In this study, K-means clustering was used to categorize students based on attributes or characteristics of the same prior knowledge level into several groups. Meanwhile, students' mathematical investigation skill is observed and categorized into several aspects based on their measured indicators. This procedure yields two categorical random variables representing the prior knowledge level and mathematical investigation aspects summarized in a two-way contingency table. Furthermore, the association between the two categorical variables was analyzed using correspondence analysis. Correspondence analysis is a powerful statistical tool for the graphical analysis of two categorical random variables that naturally depicts their association structure on a low-dimensional plot, called a correspondence plot (Beh & Lombardo, 2014; Greenacre, 2017; Lestari et al., 2020)

As explained earlier, the main concern in this study is to discover the association structure between prior knowledge level and mathematical investigation aspects. As a limitation, this study led to (1) reveal a significant association between variables using Pearson's chi-squared statistic; (2) visualize their associations through symmetry and asymmetry correspondence plots, and (3) examine the significance of the contribution of each category of variables by constructing an elliptical confidence region. Some findings of this study provide the novelty and open issue for future research to develop a learning environment supporting mathematical investigation activities.

METHODS

This study is empirical research that deals with associational research (also known as correlational research), which investigates the relationships between two variables without any attempt to influence them. In associational research, there is no manipulation of variables, and the existing relationship between variables is described, such that it is also sometimes referred to as descriptive research (Fraenkel et al., 2012; Salkind, 2015). However, the way of describing it is slightly different from other such studies. In this study, the relationship between variables is explained through their dependence. Two variables are said to be unrelated if they are statistically independent. Therefore, the main purpose of this study is to explain students' thinking behavior on Graph Theory by discovering the association structure (dependency) between prior knowledge and mathematical investigation skills.

Participants

The population of this study is sophomore students who are enrolled in a Mathematics Education Study Program at one of the universities in Indonesia. The sample of 97 out of 163 students was selected purposively. The student attends Discrete Mathematics lectures for one semester with three credit points. In other words, students are required to meet a minimum of 136 hours in one semester, which consists of 40 hours for lectures, 48 hours for structured assignments, and 48 hours for private study.

Materials

The study was held during the COVID-19 pandemic, so both lectures and assessments were held online. Online learning is organized using the E-campus platform. This platform provides various learning and teaching resources that allow lecturers and students to interact virtually. Additional features such as the digital library, connection to Google Meet or Zoom, lecture attendance, assignments, quizzes, and exams facilitate teaching and learning activities even in pandemic situations. The assessment and evaluation consist of individual tasks, structured tasks, midterm exams, and final exams.

Since this study was conducted in a specific classroom setting, due to the COVID-19 pandemic, here are practical guidelines for supporting mathematichal investigation skill in other classrooms: (1) review and integrate prior knowledge by providing a conceptual roadmap of the related topics to ensure that students maintain an understanding of concepts and procedures during mathematical investigation activities; (2) incorporate a mix of previously and newly learned problem types during the investigation; (3) propose open problems that students find challenging, which cannot be answered immediately and require them to solve in different ways, arousing their mathematical investigation; (4) expand students' ability to identify relevant information in new contexts by presenting problem information differently; (5) promote discussions that encourage students to offer explanations of their conjecture; and (6) sequence the instructions to allow students' mathematical investigation skills to grow incrementally.

In the Discrete Mathematics course, we discussed discrete objects in mathematics, such as logic, sets, mathematical proofs, and graph theory. Furthermore, we will focus on graph theory as one of the essential materials in this course. The topics include simple graphs, special graphs, isomorphism, invariants, connectivity, coloring, Euler graphs, Hamilton graphs, planar graphs, and trees. The following are definitions of some related concepts in graph theory taken from Ferland (2019).

Definition 1: Simple graph

A graph G consists of a pair of sets: vertex set V and edge set E, denoted by $G = (V, E)$. An edge of G is a function that assigns two vertices, that is $e \in E$ such that $e \mapsto \{u, v\}$ for some $u, v \in V$. The vertices u and v are the endpoints of the edge e. If $e \mapsto \{u\}$ has a single endpoint, then it is called a loop. Two or more edges assigned to the same set of endpoints are called multiple edges. A *simple graph* is a graph $G = (V, E)$ that has no loops and multiple edges. See Figure 1.

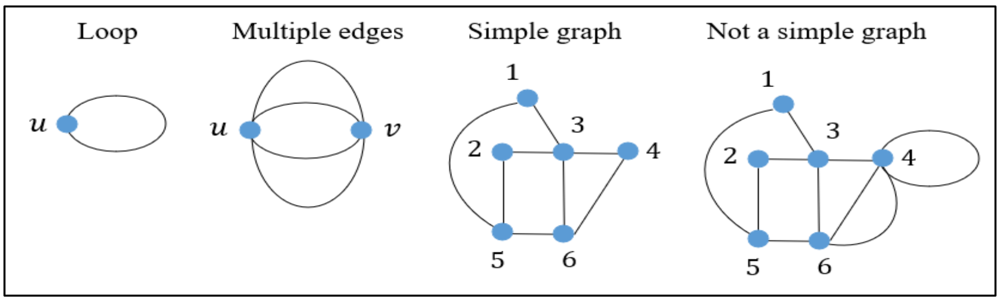


Figure 1. Illustration of the simple graph

Definition 2: Graph isomorphism

Let $G = (V_G, E_G)$ and $H = (V_H, E_H)$ be graphs. A *graph isomorphism* from G to H is a pair of bijections $f_V: V_G \to V_H$ and $f_E: E_G \to E_H$ such that, for $e \in E_G$, the bijection f_V maps endpoints of e to the endpoints of $f_E(e)$. If there exists a graph isomorphism from G to H, then G is isomorphic to H, denoted by $G \cong H$. See Figure 2.

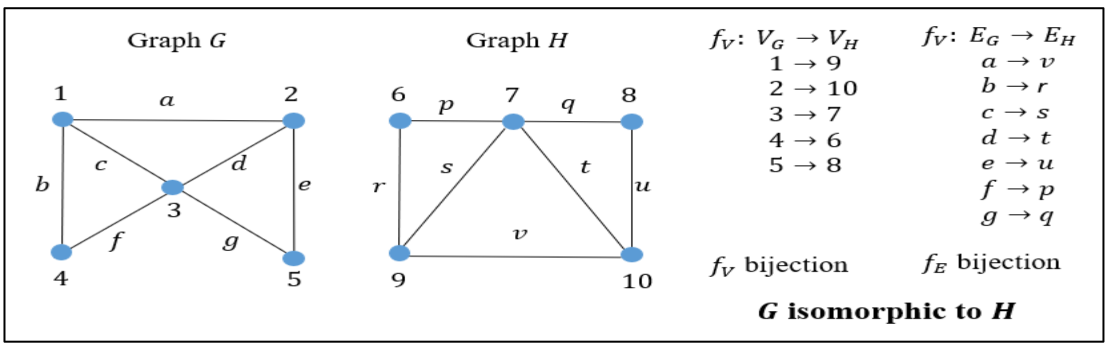


Figure 2. Illustration of a graph isomorphism

Definition 3: Walk, path, circuit, and cycle

A *walk* in graph $G = (V, E)$ is an alternating list of vertices and edges that starts at vertex v_0, end at vertex v_n for $n \geq 0$. A *path* is a walk with no repeated vertices. A *circuit* is a walk of positive length that starts and ends at the same vertex. A *cycle* is a circuit in which the only vertex repetition is $v_n = v_0$.

Definition 4: Hamiltonian graph

Let G be a graph. A *Hamiltonian cycle* in G is a cycle that covers every vertex. A *Hamiltonian path* in G is a path that covers every vertex. Graph G is said to be a *Hamiltonian graph* if it contains a Hamiltonian cycle. See Figure 3.

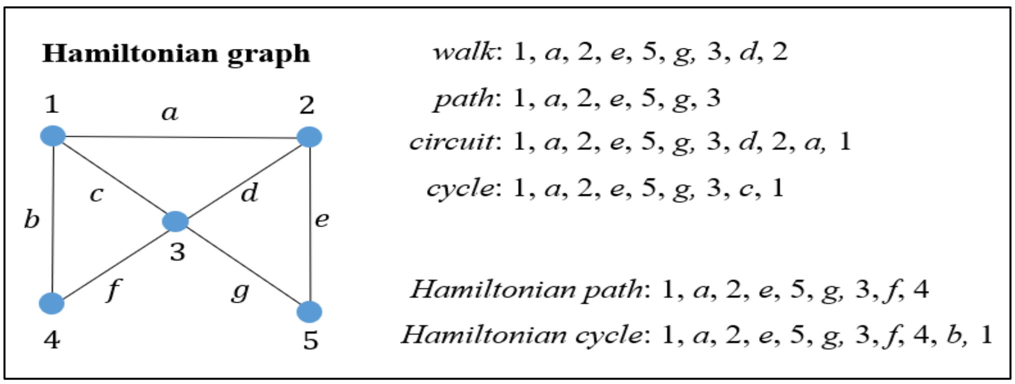


Figure 3. Illustration of definitions 3 and 4.

Definition 5: Subgraph, connected graph and tree

A graph $H = (W, F)$ is a *subgraph* of a graf $G = (V, E)$ if $W \subseteq V$, $F \subseteq E$, and the endpoints of the edges in F all lie in W and the same as in G. A graph G is *connected* if a path exists for any two vertices; otherwise, G is disconnected. A *tree* is a graph that is connected and contains no cycles. See Figure 4.

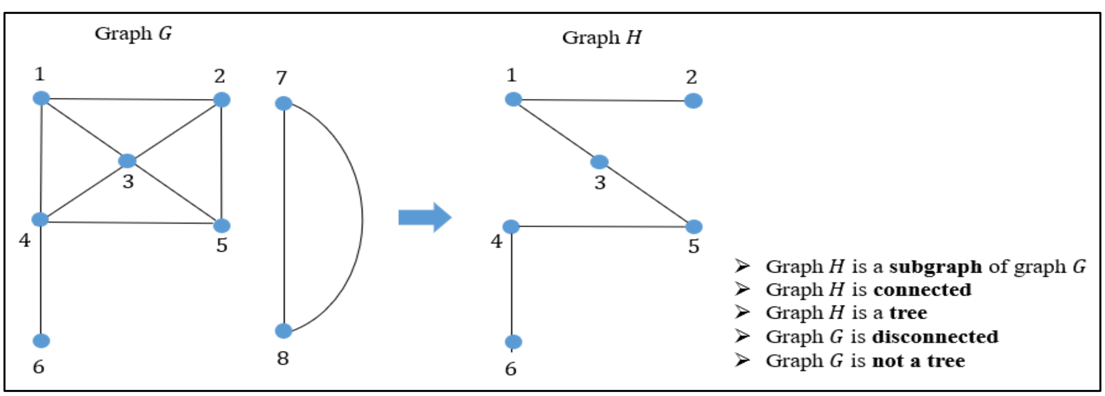


Figure 4. Illustration of a subgraph, a connected graph, and a tree

Definition 6: Weighted graph and minimum spanning tree

A *weighted graph* is a graph $G = (V, E)$ for which the edge has been assigned a positive real number called the weight of the edge. The *weight* of a subgraph is the sum of the weights of the edges in that subgraph. A *minimum spanning tree* for G is a spanning tree with the minimum weight among all spanning trees. See Figure 5.

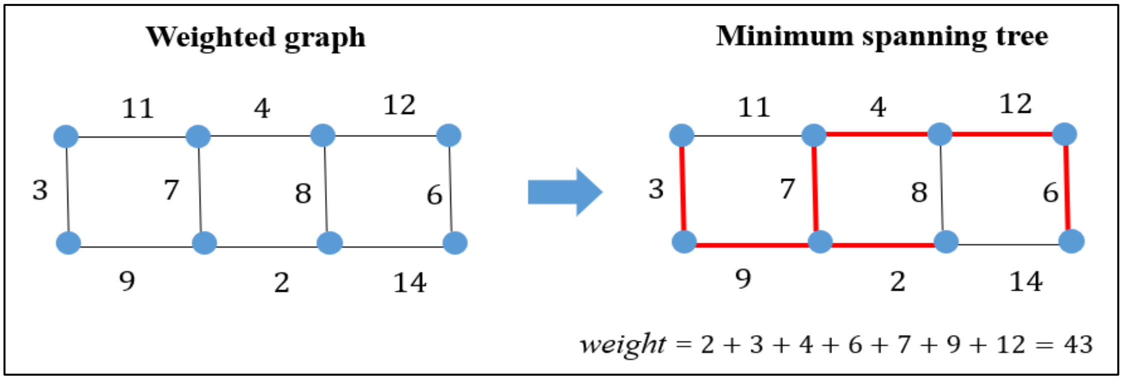


Figure 5. Illustration of a weighted graph and a minimum spanning tree

Measures

Two variables are observed to involve prior knowledge and mathematical investigation skills. Prior knowledge in mathematics is defined as the prerequisite material that students need to know before learning new mathematical concepts. In our study, this knowledge is measured by a preliminary test of basic mathematics such as pre-algebra and number theory. Meanwhile, mathematical investigation skill is measured by final exams. Indicators of mathematical investigation skills that are measured include students' capability in (1) organizing and recording data; (2) pattern searching; (3) conjecturing; (4) inferring, also (5) justifying and explaining conjecture.

Organizing and recording data

Mathematical investigations can begin with organizing and recording data. It involves the ability to integrate several mathematical skills to solve problems. Figure 6 presents the problems given in the final exam that measure students' capability to organize and record data. This problem ordered students to investigate whether the given graph was simple or not by organizing and recording a given set of vertices and a set of edges.

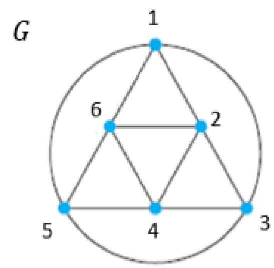
Given graph $G = (V, E)$.
Let $V = \{1, 2, 3, 4, 5, 6, 7\}$
$E = \{\{1,2\}, \{2,3\}, \{1,4\}, \{2,5\}, \{4,5\}, \{4,6\}, \{5,6\}, \{5,7\}\}$
Draw the specified graph and determine whether it is a simple graph.

Figure 6. Mathematical investigation test on organizing and recording data

Pattern searching

Searching for a pattern is often a sensible thing to do at the beginning of an investigation. Finding and describing an observed pattern provides a chance to pursue an investigation beyond the first few minutes (Benson et al., 2004). Figure 7 presents the problems that measure students' capability in pattern searching. In this problem, students must search the pattern for a path or circle that covers every vertex exactly once.

Given the pictured graph $G = (V, E)$.



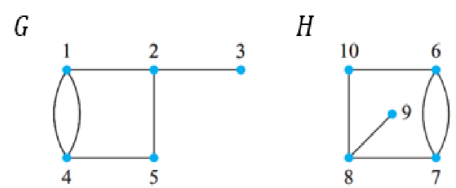
Investigate whether graph G has a Hamiltonian cycle, a Hamiltonian path, or neither. If so, then specify it.

Figure 7. Mathematical investigation test on pattern searching

Conjecturing

In mathematics, one commonly conjectures that a statement follows rule patterns that hold beyond the cases investigated and tries to prove it. It implies that a conjecture bridges someone to investigate the given problem. The following problem measures students' capability for conjecturing. Here, the problem leads students to make conjectures by defining two bijective functions such that graphs G and H are isomorphic. See Figure 8.

Given the two following graphs.



Define a pair of bijection functions $f_v: V_G \to V_H$ and $f_E: E_G \to E_H$ such that G is isomorphic to H.

Figure 8. Mathematical investigation test on conjecturing

Inferring

Mathematical investigations demand students to think through a solution and make inferences (Calleja, 2011). Inference can be viewed as an interpretation or explanation of an investigation through observations involving one's senses. To make an inference, students need to connect what they investigate to prior knowledge and the new information investigated through their senses. The inference can be made from more than one investigation, and it is not just a guess. Therefore, inference can be defined as the process of drawing a conclusion based on the available evidence from an investigation, plus previous knowledge and experience. Figure 9 presents the problems that measure students' capability in inferring.

The graph in the figure below represents an apartment complex for which electric power must be supplied. Each vertex represents an apartment building, each edge represents a potential direct power cable between two buildings, and the weight assigned to an edge reflects the cost of establishing that link.
How to determine the cheapest power grid that connects all of the apartment buildings?

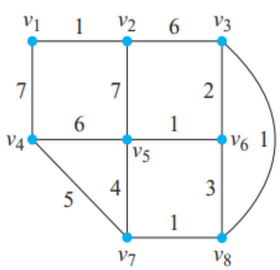
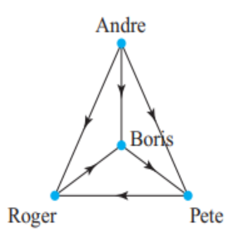


Figure 9. Mathematical investigation test on inferring

Justifying and explaining conjecture

Justification and explanation of a conjecture are a part of the mathematical investigation process that leads the student to give reasons why their conjecture makes sense by investigating a pattern, an algebraic validation, or some other logical methods. The following problem measures students' capability for justifying and explaining conjecture. See Figure 10.

The pictured directed graph is a tournament.



Find a Hamiltonian path through the tournament.
Have the results of this tournament determined a clear ranking of its competitors? Justify your answer.

Figure 10. Mathematical investigation test on justifying and explaining conjecture

Data Analysis

The empirical evidence was analyzed to explain students' thinking behavior on Graph Theory by discovering the association structure between prior knowledge and mathematical investigation skill, then visually depicting its association using the k-means clustering procedure and correspondence analysis. As a first step, the data obtained from the preliminary test that measures students' prior knowledge in mathematics were analyzed using k-means clustering (see Lestari et al., 2022a; Yudhanegara & Lestari, 2019). At this stage, students are classified into five groups (clusters) based on their prior knowledge scores. The level of prior knowledge is defined immediately after the clusters are formed by considering the average in each cluster.

The student's answers in the final test that measured mathematical investigation skills were observed and classified into six categories based on the mathematical investigation aspect, plus one category for “give up”, which represents the students who did not answer the given problem. By doing so, we obtained a two-way contingency table that classified students based on their prior knowledge level and mathematical investigation skill indicator. Both categorical variables were measured on ordinal scales. Finally, a two-way contingency table was analyzed by correspondence analysis to discover the association between prior knowledge and mathematical investigation skills.

Correspondence analysis is a statistical graphical tool to visualize the association between two categorical variables in a two-way contingency table (Lestari et al., 2023). This visualization is displayed in low-dimensional correspondence plots. Each category of two categorical variables is depicted as a coordinate point in the correspondence plot. The association between variables is visually revealed by the relative proximity of the coordinate points of one category to another. The contribution of each category to the association between variables can be determined by constructing its elliptical confidence area. In statistics, an elliptical confidence region is one form of a two-dimensional generalization of a confidence interval. The construction of the elliptical confidence area is determined using Algorithm 1.

Input

Step 1: Read a two-way contingency table

Process

Step 2: Calculate the standardized residual matrix

Step 3: Determine the singular value decomposition of the standardized residual matrix

Step 4: Determine the row and the column principal coordinates

Step 5: Determine the row and the column standard coordinates

Step 6: Determine the major and minor axes of the ellipse confidence area for each row and column principal coordinates

Output

Step 7: Plotting symmetric plot, asymmetric plot, and elliptical confidence regions

RESULTS

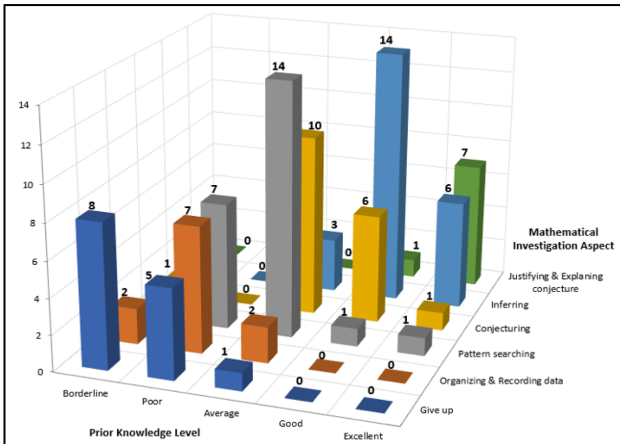
The preliminary test scores reflect students' prior knowledge of mathematics. The data from this test was analyzed using k-means clustering, which yielded five clusters to categorize students' mathematical prior knowledge. By considering the average of each cluster as the centroid, the resulting clusters were used to define five levels of students' prior knowledge, including borderline, poor, average, good, and excellent, with the following descriptions. See Table 1.

Table 1. The prior knowledge level and description based on k-means clustering.

Level	**Class centroid**	**Description**
Borderline	34,27	Not quite up to what is standard or expected borderline knowledge in pre-algebra and number theory, such as set theory, basic proof and logic, relations, functions, simplify and solve algebraic equations.
Poor	53,58	Limited knowledge in pre-algebra and number theory and need for application.
Average	61,50	Know and can apply some concepts and procedures in pre-algebra and number theory but need help to develop them.
Good	68,36	Understand the application of some concepts and procedures in pre-algebra and number theory and can develop a simple idea but needs a more detailed explanation.
Excellent	77,93	Strong understanding of concepts and procedures in pre-algebra and number theory, mostly accurate in applying and developing an advanced concept with detailed explanation

Additionally, students' answers in the final test determine the classification of students based on their achievement of the mathematical investigation aspect. Aspects of mathematical investigation are viewed as categories of ordinal-scaled variables ordered from “organizing and recording data” to “justifying and explaining conjecture”, plus one category of “give up” at the lowest order. Each aspect category was assumed to be mutually independent. It means that each student was classified on only one of their highest achievement aspects of mathematical investigation. As an illustration, a student classified in the "pattern searching" aspect means that the student already has the capability of "organizing and recording data"; a student classified in the "conjecturing" aspect means that such a student already has the capability of "organizing and recording data" and "pattern searching", and so on.

The graphical analysis of association using correspondence analysis needs a contingency table from the cross-classification of two categorical variables. Figure 11 visualize such a table that classifies student by their prior knowledge level and mathematical investigation skill. Each bar in the three-dimensional contingency table reflects the frequency of students who hold characteristics of the joint category represented by the corresponding bar. For example, the height of the bar for the joint category of "excellent-inferring" is 6. It implies that 6 out of 97 students whose excellent prior knowledge could fulfill the inferring aspect of mathematical investigation. Generally, Figure 11 shows that the student who has prior knowledge on average with pattern searching ability and the student who on a good level with inferring ability have the most frequencies.



A two-way contingency table

Prior Knowledge Level	Mathematical Investigation Aspect					
	Give up	Organizing & Recording data	Pattern searching	Conjecturing	Inferring	Justifiying & Explaning conjecture
Borderline	8	2	0	1	0	0
Poor	5	7	7	0	0	0
Average	1	2	14	10	3	0
Good	0	0	1	6	14	1
Excellent	0	0	1	1	6	7

Figure 11. Three-dimensional visualization of the contingency table

Two important aspects of contingency table analysis consider the association within and between variables that are visualized by symmetric and asymmetric correspondence plots. Both plots are obtained by performing correspondence analysis on a two-way contingency table in Figure 11. The correspondence analysis procedure is described in Algorithm 1 and interpreted based on the proximity of a coordinate point from other coordinates' positions and origin. Two categories are strongly associated if the coordinates reflecting those categories are close together. Meanwhile, the category coordinates close to the origin indicate that such a category has a small contribution to the association. Furthermore, the association within the category of variable is displayed on a symmetric plot, as in Figure 12.

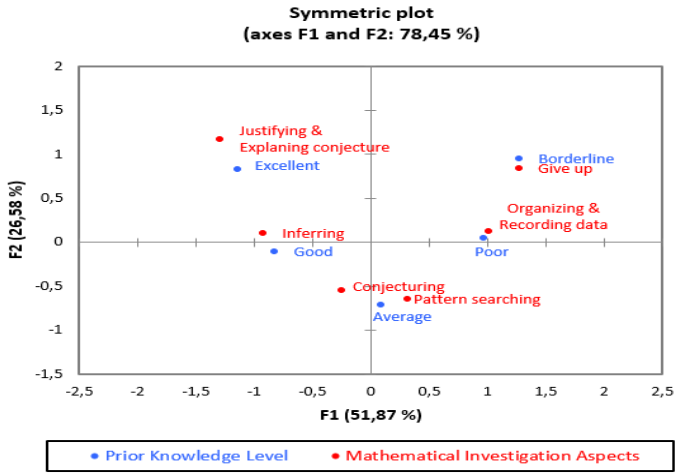


Figure 12. Symmetric correspondence plot

In the symmetric plot, the row-to-row (inter-level of prior knowledge) and column-to-column (inter-aspect of mathematical investigation) distances reflect the approximate chi-squared distance between the respective profiles (Greenacre, 2017). Thus, categories whose frequency is rarely plotted far from the origin, and vice versa (Ginanjar et al., 2016; Lestari et al., 2019a). Moreover, the association between prior knowledge and mathematical investigation skills is displayed on an asymmetric plot, as in Figure 13.

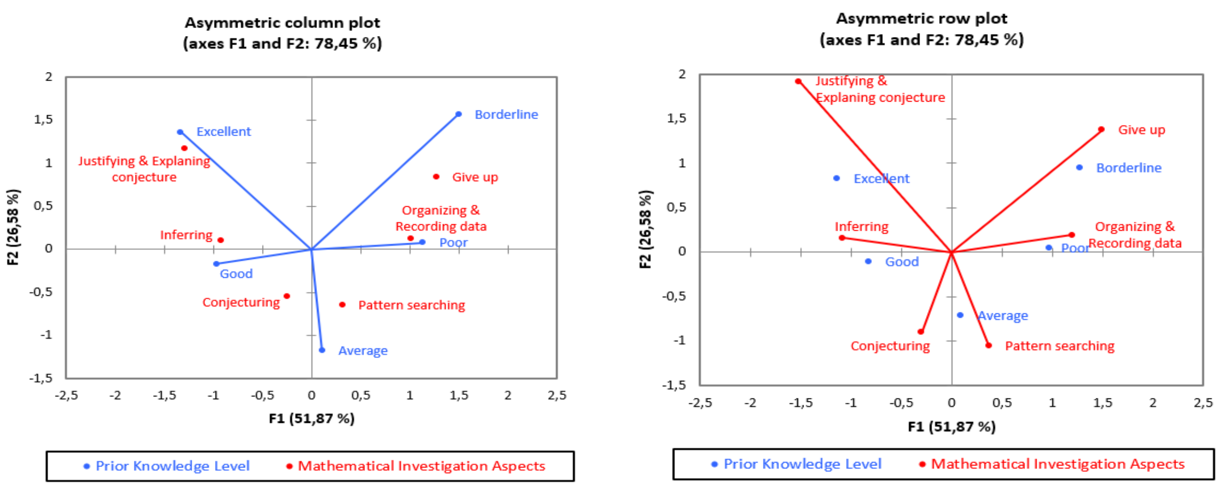


Figure 13. Asymmetric correspondence plot

To seek those categories that make a statistically significant contribution to these association structures can be identified by their proximity of coordinates' positions from the origin. For this reason, we construct an elliptical confidence region for each variable category on the correspondence plot, as shown in Figure 14.

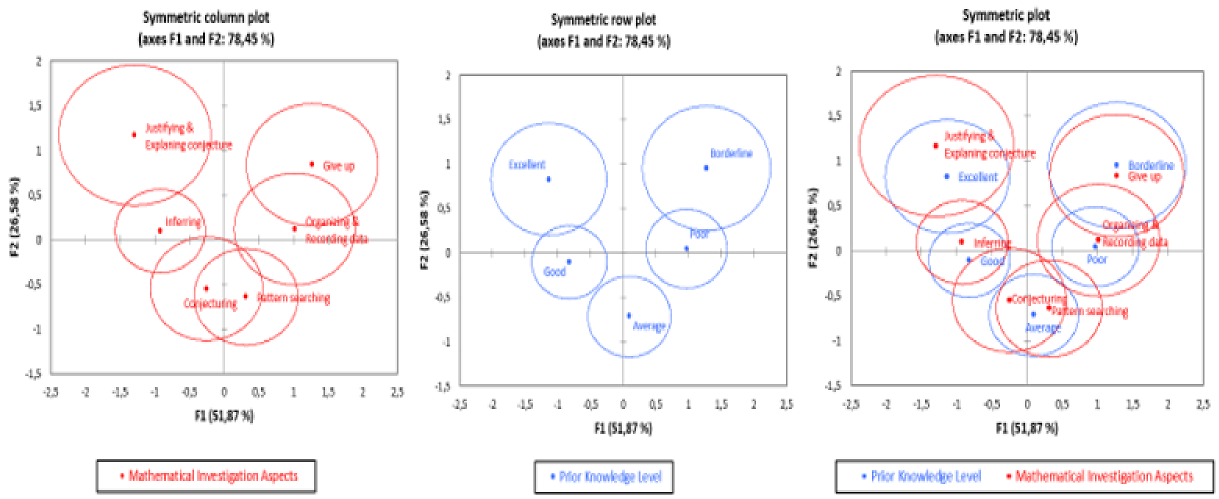


Figure 14. Elliptical confidence regions for each category coordinates

If the origin is included within the ellipse, then the particular category does not contribute to the association structure between the variables (D'Ambra et al., 2020; Lestari et al., 2019b). In other words, an elliptical region that does not include the origin means that, at the specified significance level, the category to which it is related makes a statistically significant contribution to the association structure.

Table 2. Summary statistic of elliptical confidence regions for each category.

Category	**Semi-mayor**	**Semi-minor**	χ^2 **Statistic**	***p*-value**
Borderline	1,2171	0,8712	72,1280	0,0000
Poor	0,9260	0,6629	34,4290	0,0000
Average	0,7370	0,5275	58,4268	0,0000
Good	0,8606	0,6160	29,7193	0,0002
Excellent	1,0422	0,7460	76,1687	0,0000
Give up	1,0788	0,7722	80,4863	0,0000
Organizing & recording data	1,2171	0,8712	22,4563	0,0041
Pattern searching	0,8417	0,6025	39,8883	0,0000
Conjecturing	0,9514	0,6810	22,5812	0,0039

Inferring	0,8417	0,6025	38,6447	0,0000
Justifying & explaning conjecture	1,4271	1,0215	66,8152	0,0000

Taking into the theory underlying the construction of the 100 $(1-\alpha)$% confidence region of a coordinate point in a correspondence plot, one can estimate the p-value of this point concerning its proximity to the origin. The p-value can be used to assess the statistical significance of each category considered on the association between variables (Beh, 2001; Lestari et al., 2022b). The approximation is determined and derived algebraically based on the elliptical region, as summarized in Table 2.

DISCUSSION

Considering the three-dimensional contingency table in Figure 11, the observed value of Pearson's chi-squared is $\chi^2_{stat} = 135.436$. Its value is greater than the critical value $\chi^2_{\alpha,\nu} = 31.410$ with 20 degrees of freedom. It infers that there exists a statistically significant association between prior knowledge and mathematical investigation skills. The graphical representation of this association is depicted by symmetric and asymmetric correspondence plots (see Figures 12 and 13).

The symmetrical plot in Figure 12 shows that the coordinate position for the "pattern searching" is relatively closer to "conjecture" rather than another aspect. It suggests that "pattern searching" and "conjecture" are strongly associated. As Benson et al. (2004) stated that to make conjectures in mathematical investigations, the student is required to search the patterns first. Conversely, category coordinates of prior knowledge tend to be far from each other, hence they have weak associations. It indicates that each level of students' prior knowledge has different characteristics.

The asymmetric plots in Figure 13 naturally depict the association structure between prior knowledge level and mathematical investigation aspect. The figure shows that the coordinate category for the "borderline" is close to the "give up". It suggests that students on the borderline level tend to give up when solving mathematical investigation problems. In addition, the "organizing and recording data" aspect is strongly associated with the "poor" level since their coordinates are close together. Similarly, the student on the "average" level tends to be skillful at "pattern searching" and "conjecturing". The student on the "good" level tends to be capable of "inferring", and the "excellent" student is qualifying in "justifying and explaining conjecture".

The relative position of each category of the variable to the origin reflects its contribution to the association (Lestari et al., 2019c). The closer to the origin suggests a more negligible contribution, such that it will not change the association structure if it is omitted. It means that the category coordinate that is close to the origin indicates that the category represented by such coordinate is considered does not contribute to the association structure between variables. In addition, if its elliptical regions contain the origin, thus those categories are not statistically significant contributions to the association between variables. Figure 14 shows that the origin does lie in any elliptical regions. In addition, summary statistics in Table 2 suggest that all categories of prior knowledge levels and mathematical investigation aspects statistically significantly contribute to the association structure between variables since their p-value is less than the level of significance $\alpha = 0.05$.

Furthermore, we discuss a slightly different approach to explaining how prior knowledge and mathematical investigation skills are associated. Here, we reveal the student's tendencies to solve investigation problems based on their level of prior knowledge. The mathematical investigation steps for each given problem are also described in detail.

Consider the problem in Figure 6, the investigation begins with organizing and recording data about the given set of vertices and edge sets by drawing a specified graph in such a way that (1) the only vertex points hit by a curve are the endpoints of the edge it represents; (2) each curve is one-to-one (that is, it does not intersect itself) with the exception that the ends of a loop edge are assigned to a common point; and (3) the images of curves associated with two distinct edges intersect in at most finitely many points (Ferland, 2019). To determine whether the resulting graph is simple, students should involve their prior knowledge; if it has no loops and multiple edges, it is a simple graph. Based on the assessment and evaluation, students with below-average prior knowledge (borderline, poor, and average levels) tend to reach only the first two steps, where the resulting graph contains a crossing, as illustrated in Figure 15.

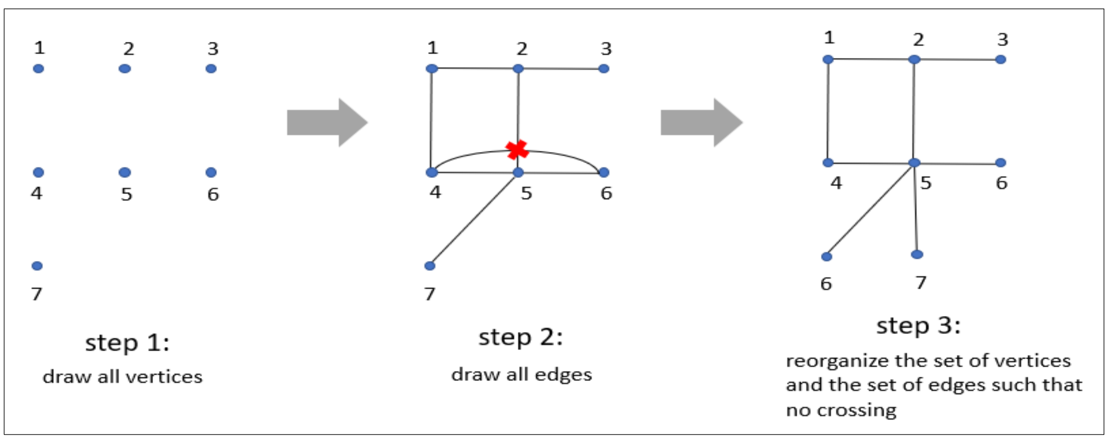


Figure 15. Mathematical investigation steps for the problem in Figure 6

Unfortunately, only a few students with good and excellent levels could complete each step perfectly. Most students who could organize and record data only reached the first two steps to conduct a simple graph by drawing all vertices and all edges without considering whether the resulting graph contains crossing edges or not, as presented in Figures 11(a) and 11(b). On the other hand, Figure 11(c) stands for the answers of students who have been able to organize and record data up to the third step such that it yields a simple graph without crossing edges. However, the student did not explain further whether the graph is simple or not, as asked in the question. It suggests that the student did not fully capture the instructions in the problem. Consequently, the mathematical investigation process was not completed well. From a conceptual point of view, students' answers in Figures 16(a) and 16(b) do not violate the definition of a simple graph. However, in procedural terms, certainly, the student's answer in Figure 16(c) is more appropriate for drawing a graph. Upon tracing and observation, it turns out that Figures 16(a) and 16(b) are answers from students with poor prior knowledge levels, while Figure 16(c) is an answer from a student with good prior knowledge.

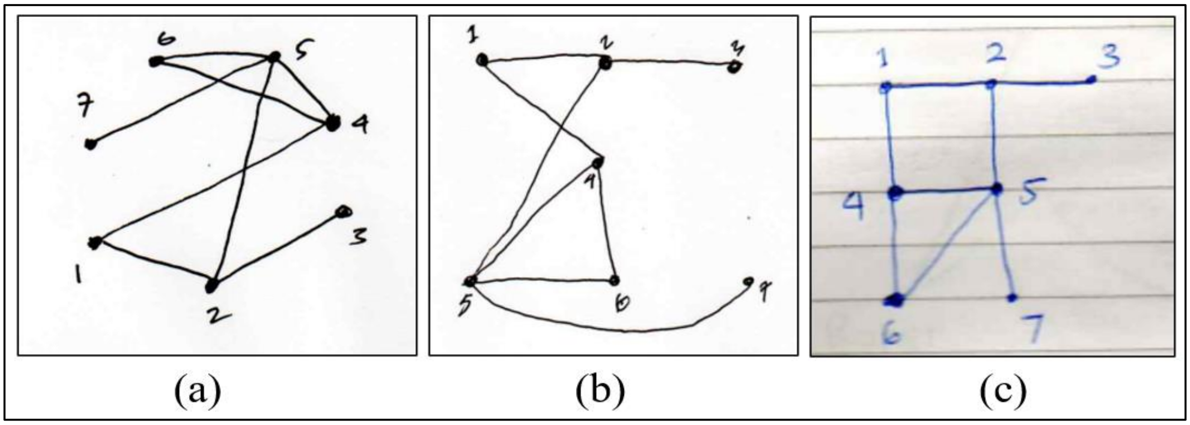


Figure 16. Example of student answers for mathematical investigation test on organizing and recording data

Furthermore, according to the problem in Figure 7, to specify a Hamiltonian graph, the investigation begins by searching for a walk with no repeated vertices or a walk of positive length that starts and ends at the same vertex and covers each vertex. Some possible mathematical investigation steps are illustrated in Figure 17.

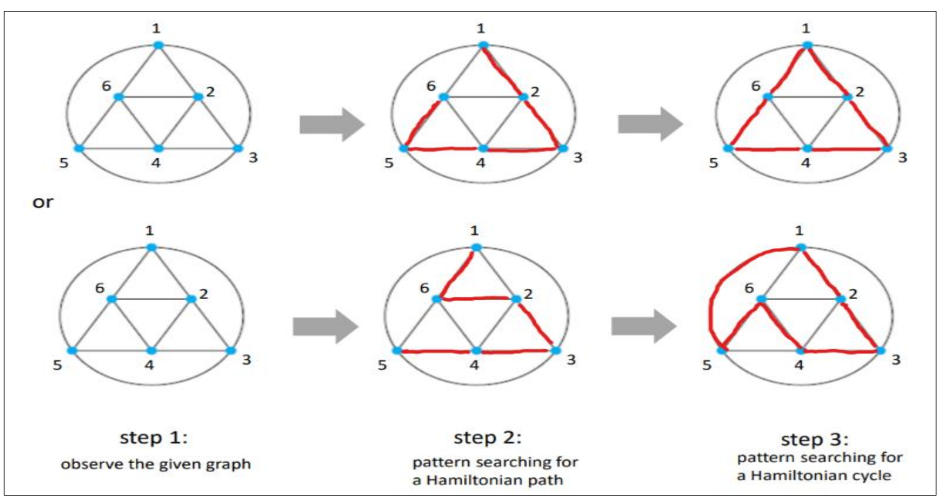


Figure 17. Mathematical investigation steps for the problem in Figure 7

Most of the students performed the mathematical investigation steps in an ordinary. Students tend to search for path or circle patterns with ordered vertices, as the first solving step in Figure 17. Only excellent students can search other patterns randomly to find Hamilton paths or cycles from the graph, as the last answer in Figure 17.

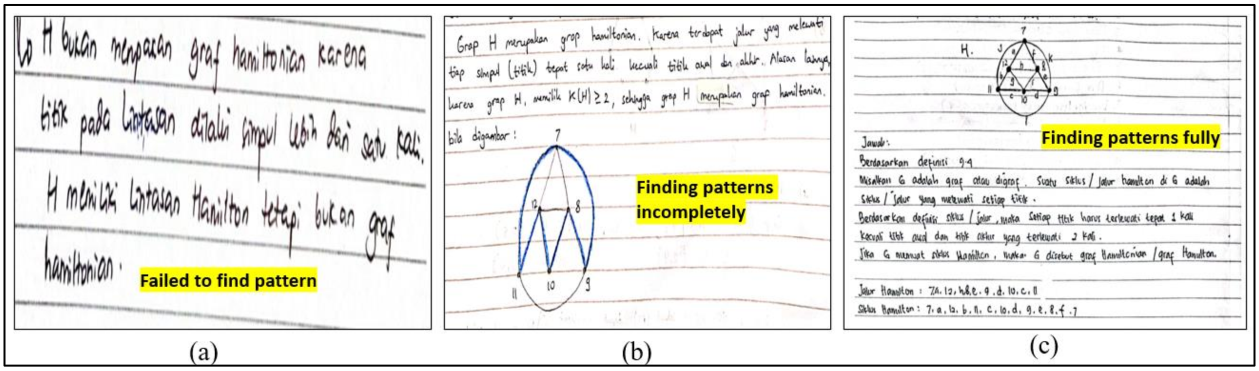


Figure 18. Example of student answers for mathematical investigation test on pattern searching

Figure 18(a) displayed the student answer who failed to find the Hamiltonian cycle or path pattern. The student states that the given graph is not a Hamiltonian graph since the vertex on the path is traversed more than once but does not mention a specific vertex. The student added the argument that the given graph has a Hamiltonian path, but it is not a Hamiltonian graph. Since the student does not specify such a path, it can be identified that the student already knew the definition of the Hamiltonian graph but did not clearly understand how to apply it. Meanwhile, Figure 18(b) presents the student answer who did an incomplete pattern search. The student successfully investigates the given graph to find the Hamiltonian cycle in the graph with an acceptable explanation but is slightly less careful in reading the instruction, hence missing answering the question regarding the Hamiltonian path. Likewise, Figure 18(c) exhibits the student's answer, who finds the pattern of both Hamiltonian path and cycle fully with a proper explanation.

The next problem in Figure 8, given two graphs, G and H. Such a problem asks the student to prove that they are isomorphic. By definition, two graphs are isomorphic if a graph isomorphism exists from one to another. Hence, this problem leads students to make conjectures by defining two bijective functions, f_V and f_E, such that any two vertices of G are adjacent in G if and only if assigned to two adjacent vertices in H.

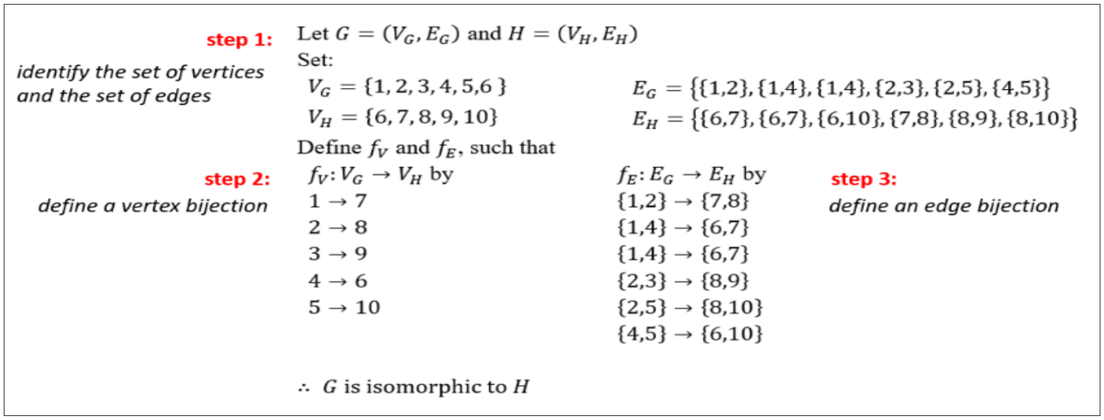


Figure 19. Mathematical investigation steps for the problem in Figure 8

Figure 19 provides a possible mathematical investigation step to solve the problem. The assessment and evaluation result suggests that students with borderline and poor prior knowledge could not make conjectures as asked. Some of them fail to define a vertex bijection and an edge bijection to prove the two graphs are isomorphic.

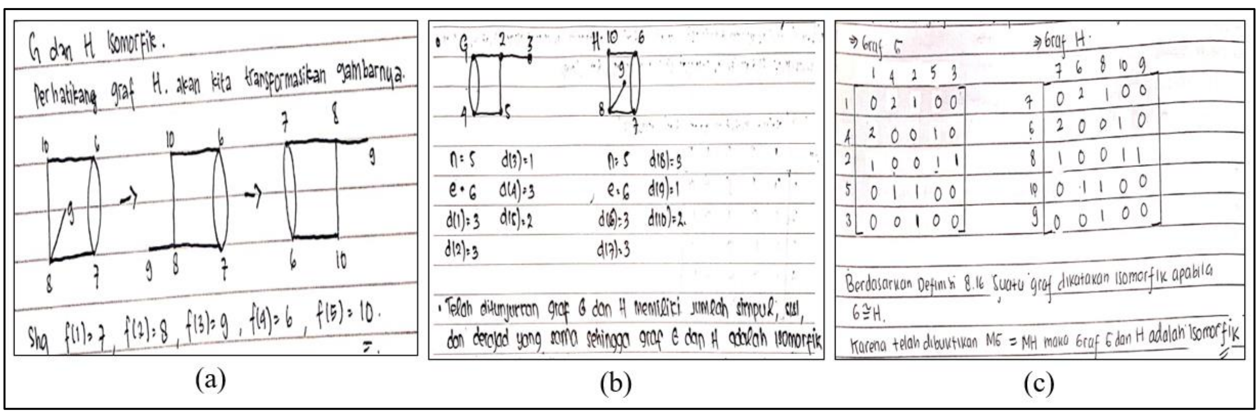


Figure 20. Example of student answers for mathematical investigation test on conjecturing

Another interesting point is that some students make the conjecture in a different way than what is directed by the question. In proving two given graphs are isomorphic, some students made conjectures by performing a transformation process on one of the graphs to obtain the other graph and then defined the mapping vertices, as shown in Figure 20(a). Some others make conjectures by applying graph invariant properties such as degree invariant and having a common adjacency matrix, as presented in Figures 20(b) and 20(c). It suggests that students' constraints are not in making conjectures but restricted to defining a pair of vertex and edge bijections from graph G to H. In the conceptual framework, these ways of constructing conjectures are acceptable.

The next investigation problem is regarding inferring, as in Figure 9. In this problem, students were required to connect their prior knowledge about the weighted tree and investigate the minimum weight. The left hand in Figure 21 shows the step after all the edges of weight 1 or 2 have been chosen. At that point, the edges of weights 3 and 4 cannot be chosen since they yield a cycle. Consequently, the edge of weight 5 is chosen next. After that, note that although there are two edges of weight 6, only one of them is a possible addition to the tree. In this step, we have the minimum spanning tree. Such investigation leads the student to infer that the cheapest power grid connecting all apartment buildings is reflected by a minimum spanning tree (Ferland, 2019).

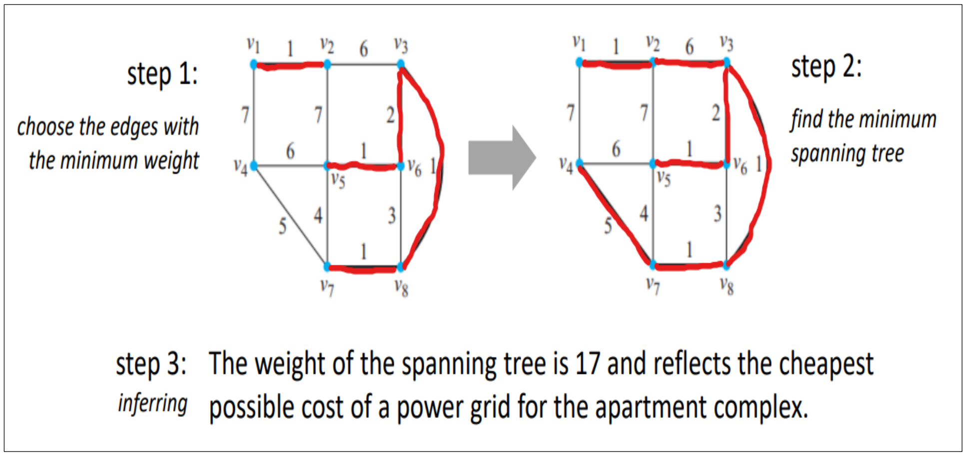


Figure 21. Mathematical investigation steps for the problem in Figure 9

The analysis of student answers shows that only good and excellent students can infer it. Figure 17 shows the answers of students who are misleading in inferring the solution of their mathematical investigation. In Figure 22 (a), the student only focuses on finding a spanning tree of the given graph and ignores choosing the minimum weight. In Figure 2 (b), the student focuses on taking the edge with small weights in the earlier steps and is trapped in executing the final step, hence not considering other possibilities for a minimum spanning tree. An essential point in inferring is validation or cross-checking the solution before drawing a conclusion. However, it is a common misstep for students to omit.

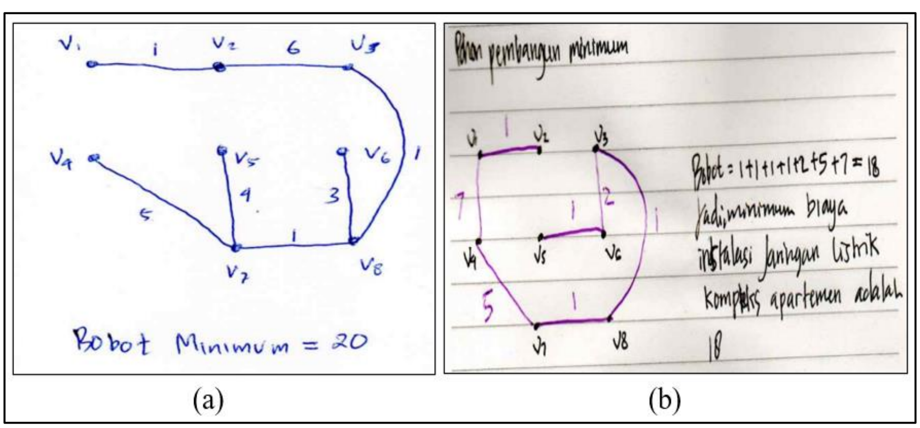


Figure 22. Example of student answers for mathematical investigation test on inferring

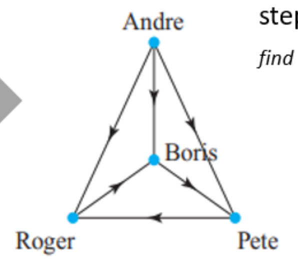
For the last problem, as in Figure 10, students should do an investigation to find a Hamiltonian path through the tournament. A tournament is a directed graph whose underlying graph is complete. In this case, the determination of the tournament is conjecture for justifying whether a ranking of its competitors is clear. After making a justification, the student should give the reason for their justification. See Figure 23.

step 1: *determine a tournament*

Match	Result
1a	Andre defeated Boris
1b	Pete defeated Roger
2a	Andre defeated Pete
2b	Roger defeated Boris
3a	Andre defeated Roger
3b	Boris defeated Pete

(a conjecture)

step 2: *find a Hamiltonian path*

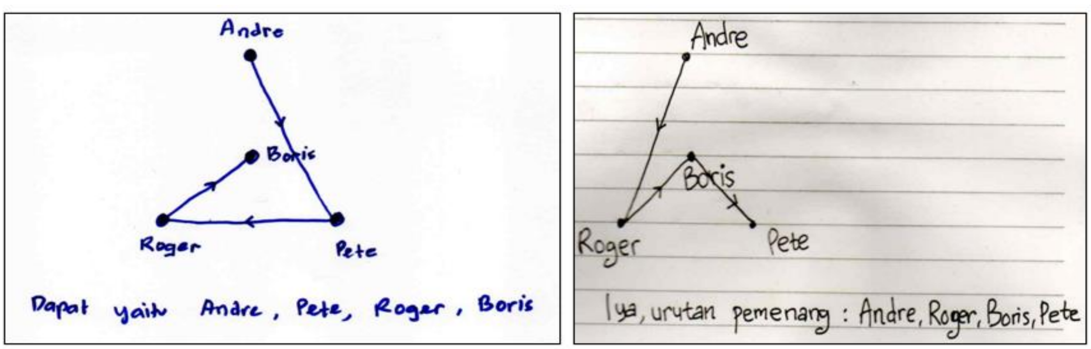


Path: Andre, Boris, Pete, Roger

step 3: *justifying and explaining conjecture* Since every tournament has a Hamiltonian path, thus the Hamiltonian path that obtains can be considered a ranking of its competitors.

Figure 23. Mathematical investigation steps for the problem in Figure 10

The underlying graph is complete graph K_4. In this case, the tournament reflects the results in a table at the first step. For instance, the edge from Andre to Boris represents the victory of Andre over Boris in their match (1a). Here K_4 reflects the results of a tournament in which every possible pair of players competed in a match an edge (u, v) would be present if and only if player u defeated player v. Since every tournament has a Hamiltonian path, at the end of the tournament, a way to rank the players could be provided by a Hamiltonian path (Ferland, 2019).



(a) (b)

Figure 24. Example of student answers for mathematical investigation test on justifying and explaining conjecture

Unfortunately, none of the students with borderline, poor, and average prior knowledge did this investigation. In addition, only a few students with good and excellent levels can justify and explain the conjecture to solve this problem. In summary, this analysis shows that differences in the students' prior knowledge level yield differences in performing the mathematical investigation step for each aspect. Figure 24 shows a common mistake made by students in justifying and explaining conjecture whether the tournament results can be used to determine the winner ranking. Instead of explaining a conjecture, almost all students did not make a conjecture or a roadmap for the tournament. As a result, the ranking of competitors determined by the students was incorrect.

CONCLUSIONS

Some findings lead to the conclusion that (1) there is evidence of a statistically significant association between prior knowledge and mathematical investigation skill; (2) a high level of prior knowledge allows the student to reach the mathematical investigation aspect completely, and vice versa; and (3) all categories of prior knowledge levels and mathematical investigation aspects statistically significantly contribute to the association structure between variables. Furthermore, the general tendency that can be concluded about students' thinking behavior on graph theory in terms of mathematical investigation skills and prior knowledge level is described as follows.

Students at the borderline level are adequate in organizing and recording data but mostly fail to find a pattern and tend to give up on problems that require higher mathematical investigation skills, such as conjecturing, inferring, justifying, and explaining conjecture. Students at the poor level can organize and record the data, sometimes fail to find a pattern, and may struggle to solve problems that require higher mathematical investigation skills, such as conjecturing, inferring, justifying, and explaining conjecture. Students at the average level can organize and record the data, tend to find a pattern incompletely, are limited of conjecturing, and may struggle in inferring, justifying, and explaining conjecture. Students at a good level understand how to organize and record data, tend to find a pattern fully, adequately in conjecturing and inferring, but mostly misleading and inconsistent in justifying and explaining conjecture. Students at the excellent level are experts in organizing and recording data, quickly capturing patterns, appropriately in conjecturing and inferring, and adequately in justifying, but sometimes reluctantly in explaining conjecture.

This study concludes some findings that provide novelty and open issues for future research, our recommendation is to develop a learning environment supporting mathematical investigation activities that involve (1) reinforcement of prior knowledge; (2) teaching multidimensional mathematical investigation problems; (3) teaching and providing a variety of problem-solving strategies including open-ended problems; (4) extending knowledge or applying knowledge in new contexts; (5) promoting a conjecturing atmosphere; and (6) encouraging students to work systematically. Particularly, in our department, we expect certain prior skills; thus, the correspondence plot obtained could be used as self-reflection for our department, whether we have gained the results as expected to strengthen certain investigation approaches.

REFERENCES

[1] Barbeau, E. J., & Taylor, P. J. (2009). *Challenging Mathematics In and Beyond the Classroom: The 16th ICMI Study*.

[2] Beh, Eric J;, & Lombardo, R. (2014). *Correspondence analysis theory, practice, and new strategies*. John Wiley & Sons, Ltd.

[3] Beh, Eric J. (2001). Confidence circles for correspondence analysis using orthogonal polynomials. *Journal of Applied Mathematics and Decision Sciences*, *5*(1), 35–45. https://doi.org/10.1207/s15327612jamds0501_3

[4] Benson, S., Addington, S., Arshavsky, N., Goldenberg, E. P., Cuoco, A., & Karnowski, E. (2004). *Ways to think about Mathematics*.

[5] Calleja, J. (2011). Integrating Investigations in Secondary School Mathematics. *M.Ed. in Mathematics Education*, *May*.

[6] D'Ambra, A., Amenta, P., & Beh, E. J. (2020). Confidence regions and other tools for an extension of correspondence analysis based on cumulative frequencies. *AStA Advances in Statistical Analysis*, 0123456789. https://doi.org/10.1007/s10182-020-00382-5

[7] Da Ponte, P. J. (2007). Investigations and explorations in the mathematics classroom. *ZDM - International Journal on Mathematics Education*, *39*(5–6), 419–430. https://doi.org/10.1007/s11858-007-0054-z

[8] Ferland, K. (2019). *Discrete Mathematics and Applications*. CRC Press.

[9] Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2012). *How to design and evaluate research in education*.

[10] Galen, F. Van, & Eerde, D. van. (2018). *Mathematical Investigations For Primary School.* Utrecht University.

[11] Ginanjar, I., Pasaribu, U. S., & Barra, A. (2016). Simplification of correspondence analysis for more precise calculation which one qualitative variables is two categorical data. *ARPN Journal of Engineering and Applied Sciences*, *11*(3), 1983–1991.

[12] Greenacre, M. (2017). *Correspondence Analysis in Practice*. CRC Press Taylor & Francis Group.

[13] Leikin, R. (2007). Habits of mind associated with advanced mathematical thinking and solution spaces of mathematical tasks. In D. Pitta, P. Phillippou, & G. Phillippou (Eds.), *Proceedings of the Fifth Congress of the European Society for Research in Mathematics Education* (pp. 2330–2339).

[14] Leikin, R. (2014). Challenging Mathematics with Multiple Solution Tasks and Mathematical Investigations in Geometry. In: Li, Y., Silver, E., Li, S. (eds) Transforming Mathematics Instruction. *Advances in Mathematics Education*. Springer, Cham59–80.

https://doi.org/10.1007/978-3-319-04993-9_5

[15] Lestari, K. E., Pasaribu, U. S., & Indratno, S. W. (2019). Graphical depiction of three-way association in contingency table using higher-order singular value decomposition Tucker3. *Journal of Physics: Conference Series, 1280*(2). https://doi.org/10.1088/1742-6596/1280/2/022035

[16] Lestari, K. E., Pasaribu, U. S., Indratno, S. W., & Garminia, H. (2019). The Reliability of Crash Car Protection Level Based on the Circle Confidence Region on the Correspondence Plot. *IOP Conference Series: Materials Science and Engineering, 598*(1). https://doi.org/10.1088/1757-899X/598/1/012061

[17] Lestari, K. E., Pasaribu, U. S., Indratno, S. W., & Garminia, H. (2020). Generating roots of cubic polynomials by Cardano's approach on correspondence analysis. *Heliyon, 6*(6), e03998. https://doi.org/10.1016/j.heliyon.2020.e03998

[18] Lestari, K. E., Utami, M. R., & Yudhanegara, M. R. (2022a). Exploratory analysis on adaptive reasoning of undergraduate student in statistical inference. *International Journal of Intruction, 15*(4), 535–554. https://doi.org/10.4135/9780857020123.n561

[19] Lestari, K. E., Utami, M. R., & Yudhanegara, M. R. (2022b). Simple algorithm to construct circular confidence regions in correspondence analysis using R. *Barekeng: Jurnal Ilmu Matematika dan Terapan, 16*(1), 065–074. https://doi.org/10.30598/barekengvol16iss1pp065-074

[20] Lestari, K. E., Utami, M. R., & Yudhanegara, M. R. (2023). Sequential exploratory design by performing correspondence analysis to investigate procedural fluency of undergraduate student. *In AIP Conference Proceedings*, *2588*(1), 050004. https://doi.org/10.1063/5.0111974

[21] Mariotti, M. A., & Pedemonte, B. (2019). Intuition and proof in the solution of conjecturing problems'. *ZDM - Mathematics Education, 51*(5), 759–777. https://doi.org/10.1007/s11858-019-01059-3

[22] McCosker, N., & Diezmann, C. (2009). Scaffolding students' thinking in mathematical investigations. *Australian Primary Mathematics Classroom*, *12*, 25–27.

[23] Quinnell, L. (2010). Why are mathematical investigations important? *Australian Mathematics Teacher*, *6*(4), 144–147. https://doi.org/10.1093/teamat/6.4.144

[24] Salkind, N. J. (2015). *Encyclopedia of research design. In Sage*. https://doi.org/10.4135/9781412983907.n439

[25] Taylan, R. D., & Da Ponte, J. P. (2016). Investigating Pedagogical Content Knowledge-in-Action. *Journal of Research in Mathematics Education, 5*(3), 212. https://doi.org/10.17583/redimat.2016.2227

[26] Wayan Puja Astawa, I., Ketut Budayasa, I., & Juniati, D. (2018). The process of student cognition in constructing mathematical conjecture. *Journal on Mathematics Education, 9*(1), 15–25. https://doi.org/10.22342/jme.9.1.4278.15-26

[27] Yeo, J. B. W. (2017). Specialising and conjecturing in mathematical investigation. *41st Conference of the International Group for the Psychology of Mathematics Education, 4,* 337–344.

[28] Yudhanegara, M. R., & Lestari, K. E. (2019). Clustering for multi-dimensional data set: A case study on educational data. *Journal of Physics: Conference Series, 1280*(4). https://doi.org/10.1088/1742-6596/1280/4/042025