AKSIOMA: Jurnal Program Studi Pendidikan Matematika
Volume 15, No. 2, 2026, 775-788

ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v15i2.13792

STATISTICAL LITERACY KNOWLEDGE DIFFUSION: DIDACTICAL TRANSPOSITION FROM SCHOLARLY KNOWLEDGE TO KNOWLEDGE TO BE TAUGHT ON LINEAR REGRESSION

Palupi Sri Wijayanti[1*], Didi Suryadi[2], Dadan Dasari[3], Al Jupri[4]

[1,2,3,4] Universitas Pendidikan Indonesia, Bandung, Jawa Barat, Indonesia
*Corresponding autho: palupi@upi.edu

Received 24 July 2025; Revised 20 January 2026; Accepted 22 May 2026

Abstract

This study examines the process of external didactic transposition in linear regression material, namely the transformation of *scholarly knowledge* into *knowledge to be taught*. This transposition is important to ensure that knowledge that is complex and tends to be abstract for students can be accessed and easily understood, especially in the context of linear regression so that it can develop students' statistical literacy. The research method was conducted with a document study by analysing primary books on statistics, reference guidelines for the preparation of mathematics textbooks, and educational curriculum documents in Indonesia. The results of the study showed that the knowledge adaptations that have been identified are the meaning of dependent and independent variables, the concept of linear regression, the concept of best fit line, linear regression equation model, and random error. Although there are knowledge adaptations that facilitate understanding for students, there are also limitations in the presentation of random error distribution assumptions in the knowledge to be taught. The contribution of this research is an understanding of the construction of statistical knowledge presented in the educational context and the strengthening of materials that can develop students' statistical literacy.

Keywords: Didactic transposition; linear regression; knowledge to be taught; scholarly knowledge; statistical literacy.

Abstrak

Penelitian ini mengkaji proses transposisi didaktis eksternal pada materi regresi linear, yaitu transformasi pengetahuan ilmiah (scholarly knowledge) menjadi pengetahuan yang akan diajarkan (knowledge to be taught). Transposisi ini penting untuk memastikan pengetahuan yang kompleks dan cenderung abstrak bagi siswa dapat diakses dan mudah dipahami, khususnya dalam konteks regresi linear sehinga dapat mengembangkan literasi statistik siswa. Metode penelitian dilakukan dengan studi dokumen melalui analisis buku primer materi statistika, pedoman acuan penyusunan buku teks matematika, dan dokumen kurikulum pendidikan di Indonesia. Hasil studi menunjukkan adaptasi pengetahuan yang telah teridentifikasi yaitu pemaknaan variabel dependen dan independen, konsep regresi linear, konsep garis best fit, model persamaan regresi linear, serta random error. Walaupun terdapat adaptasi pengetahuan yang memfasilitasi pemahaman untuk siswa, tetapi ditemukan pula adanya keterbatasan dalam penyajian asumsi distribusi random eror dalam knowledge to be taught. Kontribusi penelitian ini yaitu pemahaman terhadap konstruksi pengetahuan statistika yang disajikan pada konteks pendidikan serta penguatan materi yang dapat mengembangkan literasi statistik siswa.

***Kata kunci**: Transposisi didaktis; regresi linear; pengetahuan yang akan diajarkan; pengetahuan ilmiah; literasi statistik.*

INTRODUCTION

The dynamic world that cannot be separated from the flow of information dominated by data requires the ability to understand, translate, interpret, and criticise statistical-based information. This competency is crucial for every individual because it is not only data

ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v15i2.13792

scientists who must master statistical literacy, which is considered a specialised skill. However, statistical literacy is a foundation skill for every citizen in making informational decisions in various fields of life such as education (Callingham & Watson, 2017), health, economy, or other fields (Tran et al., 2023). Therefore, through education in learning will be the first step in equipping the younger generation to master adequate statistical literacy.

Learning mathematics that teaches statistics both in secondary school and in higher education is a medium for developing statistical literacy (Droogers & Drijvers, 2017; Weiland & Sundrani, 2022). However, often the learning of statistics focuses on calculation and procedural aspects so that it tends to ignore the conceptual understanding that is deep and essential for literacy (Droogers & Drijvers, 2017; Tran et al., 2023). The impact is the existence of learning difficulties for students in applying their statistical knowledge to everyday problems and being able to critically evaluate statistically based information.

Statistics materials are taught in secondary schools and higher education, and have high relevance in various disciplines one of which is regression analysis (Aziz & Rosli, 2021). The use of linear regression to perform powerful inference statistical tests in finding relationships between variables (Aityan, 2022), making predictions, and explaining phenomena contextually. Regression analysis can basically be done when some mathematically prerequisite assumptions are met (Anandhi & Nathiya, 2023). This presentation of understanding is still seen as an introduction only yet to lead to the construction of understanding and knowledge part of linear regression analysis. Furthermore, basic calculations involving statistical data will help in determining the regression model equation so that understanding in interpreting coefficients, validating models, interpretting regression models (Anandhi & Nathiya, 2023) obtained becomes a challenge in the learning process. This tends to cause students' understanding is still not deep and more mechanistic (Makkawy, 2023)

The complexity of linear regression knowledge derived from *scholarly knowledge* and *knowledge to be taught* to students is often a gap as described earlier. This is a fundamental problem in the didactic process. This phenomenon is referred to as didactic transposition which involves a series of knowledge transformations, from decontextualisation (El Fadel, 2024) and depersonalisation (Mezaini et al., 2022) to restructuring (Do & Nguyen, 2020; Son et al., 2020; Yin et al., 2025) and recontextualisation (Amna Saleem et al., 2021; Krogh et al., 2021), so that it fits the context and learning objectives.

The transformed knowledge needs to be diffused so that it can conceptually spread linear regression knowledge from a group of scientists or experts in mathematics and statistics to curriculum holders or textbook writers and teachers as transposition agents. Furthermore, how linear regression knowledge is transposed from complex statistical concepts is able to facilitate in-depth statistical literacy.

Previous research highlights the process of didactic transposition on various mathematical concepts such as functions and straight line equations. Andriatna & Sujadi (2024) conducted a study related to the transposition of function concepts from *scholarly mathematics* to *school mathematics*, which shows the process of adapting

ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v15i2.13792

knowledge to the context and presentation of material by taking into account the cognitive development of students (Andriatna & Sujadi, 2024). Another study was also conducted by Sulastri (2023) who specifically explored didactic transposition on the knowledge of the concept of limit function. The results of his study identified that the introduction of concepts intuitively in the presentation of material in textbooks sometimes does not provide adequate formal reinforcement and has the potential to cause learning obstacles (Sulastri, 2023). In addition to the didactic transposition study on function and limit function material above, Suarsana and his research team examined the process of didactic transposition on straight line equation material from *scholarly knowledge* to *knowledge to be taught*. The results of the study stated that there were potential learning obstacles from the presentation in the textbook of the general form of the line gradient equation and its definition (Suarsana et al., 2024). Puspita and Kustiawan (2024) also examined different material from before, namely about more complex calculus in lectures in higher education. The results of the study found that in the didactic transposition process that had been passed, learning barriers and knowledge acquisition processes of prospective mathematics teachers were identified in the fold integral material. El Fadel examined the practices and pedagogical issues in didactic transposition as a process of knowledge transformation in order to achieve targeted learning objectives (El Fadel, 2024)

Previous studies that have been conducted by several researchers above, still show a gap in the didactic transposition process that discusses scholarly knowledge to be taught related to linear regression material. In addition, integration and holistic analysis of the diffusion process of linear regression knowledge specifically to develop students' statistical literacy fill the gap. This research explicitly examines how scientific knowledge about linear regression is transformed didactically to foster understanding and develop statistical literacy (Do & Nguyen, 2020; Gök et al., 2019; Rudi et al., 2022). Practically, the research findings provide empirical guidance in the preparation of curricula, textbooks, and for educators to design linear regression teaching strategies that are expected to be more effective. This will be stimulant in overcoming the common challenges in teaching more complex statistical concepts and useful to ensure that students can master the concepts not only limited to the computation but more deeply understand and criticise based on statistical data. Theoretically, the contribution of this research is to provide an in-depth understanding of how linear regression knowledge is adapted from scientific knowledge from mathematicians into knowledge that is easy to understand for learners. Therefore, this research will answer how the diffusion process of linear regression knowledge seen from the perspective of external didactic transposition through the transformation of scientific knowledge to taught knowledge and facilitate statistical literacy.

METHODS

Research Design

This study used a qualitative research approach with a hermeneutic phenomenological design to investigate the meaning and transformation of linear regression knowledge during external didactic transposition. The project was coordinated by the

ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v15i2.13792

Didactical Design Research (DDR) paradigm, which provided a methodical lens for examining how scholarly knowledge is transformed into knowledge (Phuong & Quan, 2023; Wang et al., 2023) for senior high school mathematics instruction. The research design involved sequential analytical stages, which included identifying essential linear regression concepts in scholarly sources, examining curriculum-related mediating documents, analyzing official mathematics textbooks, and comparing conceptual representations across these didactic levels.

Subject, Location, and Time of Research

Instead of human volunteers, the study used document-based sources. The study was conducted within the framework of the Indonesian education system, as evidenced by the curriculum materials and government-issued textbooks analyzed. Data collection and analysis were carried out between January and June 2024. The materials were chosen using a purposive sample technique that took into account authority, relevance, and official use in mathematics education, ensuring that the texts appropriately represented scholarly knowledge, mediating norms, and knowledge to be taught.

Research Materials

The research material focused entirely on linear regression, a key issue in statistics teaching. The analysis addressed key conceptual components such as the meaning of dependent and independent variables, relationships between dependent and independent variables, regression models/equations, interpreting regression equations and coefficients, best-fit lines, the role of random error and residuals, interpolation and extrapolation, and the use of linear regression in problem-solving contexts. Particular emphasis was placed on to how these concepts are positioned to help students acquire their statistical literacy.

Data Sources

The data sources represented several stages of the didactic transposition process, from scholarly knowledge to knowledge to teach. Scholarly knowledge was represented by seven authoritative university-level statistics textbooks written by mathematics and statistics professionals and widely used in higher education and teacher education programmes. Table 1 contains references to intellectual knowledge. Mediating papers included official guidelines for the creation of school mathematics textbooks provided by the Indonesian Ministry of Education and Culture via BSKAP, which served as normative guides for content adaptation. The knowledge to be taught was represented by government-published mathematics textbooks for Grade XI senior high school, which are used nationwide and serve as the approved instructional resources for students.

Table 1. List of primary statistics books that become data sources for scholarly knowledge

No.	Author & Year	Title
1	Prem S. Mann (2010)	Introductory Statistics
2	Neil A. Weiss (2012)	Introductory Statistics (9Ed)
3	Roxy Peck, Chris Olsen, Jay L. Devore (2012)	Introduction to Statistics and Data Analysis (4Ed)

ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v15i2.13792

No.	Author & Year	Title
4	Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye (2013)	Essentials of Probability & Statistics for Engineers & Scientists
5	Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye (2016)	Probability & Statistics for Engineers & Scientists. Ninth Edition, global edition.
6	Joseph F. Healey & Steven G. Prus (2016)	*Statistics:*A Tool For Social Research
7	Wahyudin (2023)	Book of Mathematics Subject Matter to Improve Mathematical Proficiency of Professional Mathematics Teachers

Research Instruments and Techniques for Data Collection

Data were obtained utilizing a document analysis tool that took the shape of an analytical rubric. The rubric was developed using didactic transposition theory, DDR principles, and recognized statistical literacy frameworks. It included metrics that assessed conceptual accuracy, representational forms, interpretative depth, and pedagogical emphasis of linear regression principles across various document kinds. Data was collected by reading, classifying, and categorizing key text parts about linear regression in each publication.

Validation of the Instrument

The validity of the study instrument was ensured by expert assessment. Two experts in mathematics education and statistics assessed the analytical rubric to assess the indicators' relevance, clarity, and appropriateness in expressing the desired concepts. Revisions were performed based on expert feedback to improve content validity and assure alignment with study objectives and analytical focus.

Techniques for Data Analysis

Data analysis involved qualitative content analysis and hermeneutic interpretation. Each document was classified using the established indications in the analytical rubric, followed by a cross-source comparison of scholarly information, mediating texts, and knowledge to be taught. The investigation aimed to discover patterns in conceptual alignment, simplification, omission, and reinterpretation of linear regression principles. Hermeneutic interpretation was then used to explain the didactic reasoning behind these modifications (Morales-López et al., 2023). Iterative coding and triangulation across data sources helped to preserve analytical rigor. The appearance of consistent patterns explaining how linear regression knowledge is adapted to support students' statistical literacy demonstrated the analysis's adequacy.

RESULTS AND DISCUSSION

The research results obtained from the didactic transposition process from scholarly knowledge to knowledge to be taught on linear regression material are presented in the following description.

1. Interpretation of Dependent and Independent Variables

In this material, the delivery of understanding related to the difference between dependent and independent variables must be clearly emphasised because it will be the basis for analysis and interpretation into the problems that are being discussed. In the linear regression material taught in class, exposure related to the definition of independent variables that affect the dependent variable needs to be emphasised through the relationship between variables (Jacobs, 2023; Prasad, 2024).

ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v15i2.13792

In scholarly references, the understanding of this variable is presented with examples that are expected to be accepted by readers. For example, an economist investigates the relationship between expenditure on food consumption and household income. What factors or variables do households consider when deciding how much money to spend on food each week or each month? Of course, household income is one factor. But there are many other variables that influence spending on food consumption. For example, the assets/furniture that the household already owns, the size of the household, the preferences and tastes of household members, and the special dietary needs of household members are some of the variables that influence household decisions on food expenditure. These variables are called independent or explanatory variables because they vary independently, and explain the variation in food expenditure among different households. In other words, these variables explain why different households spend different amounts of money on food. Food expenditure is called the dependent variable because it depends on the independent variables. Studying the effect of the independent variable on the dependent variable uses a method called regression analysis. However, choosing only one (usually the most important) independent variable and studying the effect of that single variable on the dependent variable is called simple regression (Mann, 2011).

The submission of the above example shows that the existence of dependent and independent variables is also related to the existence of a relationship between the two either causal or association. The concept of the meaning of the dependent and independent variables is explained in detail how their role in regression analysis affects each other between variables.

Furthermore, the adaptation of knowledge of dependent and independent variables at the stage of knowledge to be taught can be seen from the exposure in the mathematics textbooks used by students in the classroom. Although conceptually not deviating, the definitions presented tend to be still abstract, circular, less explicit, not yet showing formal operations in explaining the nature of causal or predictive relationships in depth, when compared to the exposure in primary books that represent *scholarly knowledge.* The following is a presentation of knowledge to be taught represented in mathematics textbooks in Figure 1.

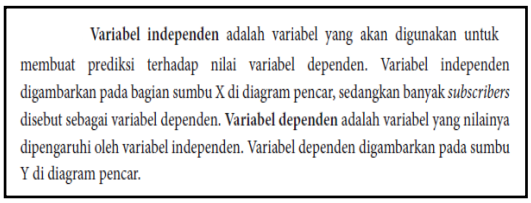

Variabel independen adalah variabel yang akan digunakan untuk membuat prediksi terhadap nilai variabel dependen. Variabel independen digambarkan pada bagian sumbu X di diagram pencar, sedangkan banyak *subscribers* disebut sebagai variabel dependen. **Variabel dependen** adalah variabel yang nilainya dipengaruhi oleh variabel independen. Variabel dependen digambarkan pada sumbu Y di diagram pencar.

Figure 1. Defining the Independent and Dependent Variables in the Learners' Mathematics Textbooks

2. Concept of Linear Regression through *Best-Fit* Line

To continue the understanding of linear regression material, it is necessary to emphasise the relationship between the dependent and independent variables that can be visualised through a line. This linear regression material is the first time it appears in mathematics learning in Indonesia for high school students so using a line representation that shows the relationship between the two associations of the two variables is something that makes it easier for students to construct their knowledge.

ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v15i2.13792

The research findings show that linear regression knowledge is introduced in scholarly knowledge and knowledge to be taught by modelling the relationship between two variables through a line drawing which is then called the best fit line. The consistency that occurs in defining linear regression using a straight line representation indicates an effective transposition process at the basic level.

The scholarly knowledge reference states that a group of points on the Cartesian plane that represent the position of the dependent and independent variables can be seen the relationship pattern between the two. The clarity of the relationship pattern can be improved by drawing a straight line through the group of points in such a way that the line touches each point or is as close as possible to it. This line is then called the regression line (Healey & Prus, 2015). Furthermore, the knowledge to be taught is represented in the presentation of material in the mathematics textbook by providing an explanation as in Figure 2.

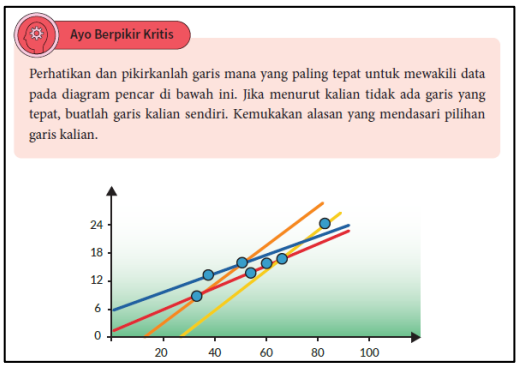


Figure 2. Pieces of Student Activities in Mathematics Textbooks for Students page 96 (Susanto et al., 2021)

In Figure 2. it can be seen that learners are directed to build understanding related to regression lines. This knowledge construction will become knowledge when there is a validation situation manifested in discussions with fellow students in the classroom. Furthermore, there is reinforcement from the teacher based on the validation process related to the explanation of the regression line. This has also been conveyed in the mathematics textbook for students (can be seen in Figure 3).

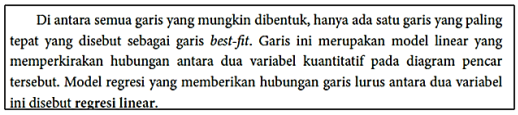

Di antara semua garis yang mungkin dibentuk, hanya ada satu garis yang paling tepat yang disebut sebagai garis *best-fit*. Garis ini merupakan model linear yang memperkirakan hubungan antara dua variabel kuantitatif pada diagram pencar tersebut. Model regresi yang memberikan hubungan garis lurus antara dua variabel ini disebut regresi linear.

Figure 3. Explanation of the regression line in the student mathematics textbook(Susanto et al., 2021)

In Figure 3, learners are introduced to the term "*best-fit*" line, which is first encountered in mathematics lessons. In the student's mathematics textbook, it is stated that the *best-fit* line is the line that best represents the data in the scatter diagram (Susanto et al., 2021). In theory in *scholarly knowledge* that a group of data points can be drawn many different lines through which the "*best*" line that represents can be chosen (Weiss, 2012). To introduce the concept of *best-fit*, it is necessary to have a model that shows *best* or hypothetical and *non-best-fit* lines. This will facilitate participants to institutionalise knowledge related to linear regression through a different context to the creation of a *best-fit* line on a scatter diagram. One way to institutionalise the understanding of the best-fit line, learners can be given a context in the form of a method of determining the best-fit line, namely through the least squares criterion (Weiss, 2012) by finding the regression coefficient (Walpole et al., 2012).

The presentation of this material uses a scatter diagram visualisation tool. The scholarly knowledge explains how to analyse the relationship of the

ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v15i2.13792

variables being observed through the help of a scatter diagram. Meanwhile, the knowledge to be taught asks students to identify any patterns seen in the plot on the scatter diagram. This visible pattern can be expressed by drawing a line so that it shows the best fit that represents the data by exploring various possible suitable lines. This approach is intended to provide deep and long-lasting understanding. This is different from the presentation of scholarly knowledge, which tends to directly use the theoretical best fit line method. This adaptation shows innovation in the knowledge diffusion process aimed at making it easier for students to understand through visualisation, theoretical and conceptual.

The use of the concept of best fit used in scholarly knowledge is a start in introducing the concept of the least squares method as a criterion in determining the best fit line and calculating the regression coefficient. Meanwhile, the introduction of best fit in the knowledge to be taught through mathematics textbooks has not presented a detailed mathematical method at the initial stage of why to use best fit and justify the correctness of this best fit to represent data on variables.

3. Linear Regression Equation Model and the Concept of Random Error

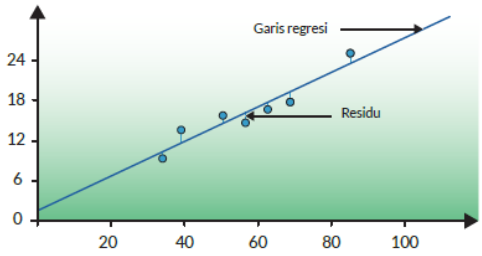
The next essential linear regression knowledge is the linear regression equation. In discussing this understanding, it is necessary to know the transition process so that it can be approached through the calculation of a linear equation model. In scholarly knowledge, the presentation of a linear regression model that is preceded by an understanding of the best fit line, assumes that the dependent variable is fully influenced by the independent variable, intercept parameter, and regression coefficient. This regression model is called the deterministic model which can be expressed as$y = a + bx$ where y is the dependent variable, x is the independent variable, a is the intercept, and b is the regression coefficient (Peck et al., 2016). The deterministic model is a conceptual starting point in understanding the basis of linear relationships before introducing data to real and complex contextual problems (Wu & Ye, 2015). In the presentation of knowledge to be taught also describes this deterministic model but there is a transition process that is still not clearly described. This lack of clarity is presented in the explanation of the relationship between variables that is not perfect because there are other factors that influence the dependent variable and cannot be fully explained by the independent variable. These other influencing factors that are included in the model are represented as random errors. The deterministic regression model with this condition becomes a probabilistic regression model represented in the form of the equation $y = A + Bx + \varepsilon$ (Peck et al., 2016), where y is the dependent variable, x is the independent variable, a is the intercept, b is the regression coefficient, and ε is random error. ε explains the variation in y that cannot be explained by x due to unobserved variables or not included in the regression model obtained, measurement error, or lack of human accuracy (Montgomery et al., 2021; Yan, 2022). Therefore, it is necessary to understand important assumptions related to distributions such as $E(\varepsilon) = 0$ and $Var(\varepsilon) = \sigma^2$. The scholarly knowledge introduces these distributions explicitly but the knowledge to be taught directly introduces residuals, squared residuals,

ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

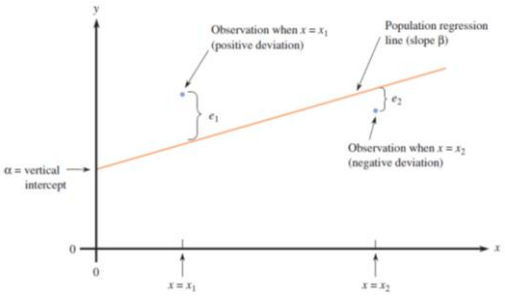
DOI: https://doi.org/10.24127/ajpm.v15i2.13792

and the sum of squared residuals. Residual is defined as the difference between the observed value of the dependent variable (y) and the predicted value of the independent variable ($\hat{y}$) (Pal & Bharati, 2019). The residual formula is written as $\varepsilon = y - \hat{y}$ (Susetyo, 2012). Furthermore, the residual square formula is$\varepsilon^2 = (y - \hat{y})^2$ and the residual sum of squares is $(\sum \varepsilon^2) = \sum (y - \hat{y})^2$ (Susanto et al., 2021). The introduction of *random error* delivered in the textbook is not much different but the reinforcement provided is still not in-depth because it is limited to an introduction accompanied by a formula.

The equation of the linear regression line approximated by the least squares method is in accordance with the presentation starting from the concept of residuals shown through scatter diagrams, as in the following example presented in the mathematics textbook.



(a) in a mathematics textbook



(b) in a scholarly knowledge reference book

Figure 4. Explanation of residuals through scatter diagrams

This study examines the consistency and gaps that arise in the process of diffusion of linear regression knowledge from mathematics or statistics scientists to the noosphere so that understanding of the knowledge is easily accepted. This study looks at the process of didactic transposition from *scholarly knowledge to knowledge to be taught* on linear regression material that reveals how the essential concepts of linear regression, namely the meaning of dependent and independent variables, *best-fit* lines, linear regression equation models along with the concept of random error, are adapted from complex forms at the higher education level to be presented at the level of secondary school learning in Indonesia.

In the previous study, the focus of the study was on the effectiveness of teaching or curriculum while this study highlighted the comparison of conceptual depth on scientific knowledge in scholarly knowledge to knowledge to be taught. The findings of this study are in line with the results of a study conducted by Andriatna & Sujadi (2024) on the process of didactic transposition on *scholarly mathematics* function material to *school mathematics*. The study showed that knowledge adaptation often changes the context from abstract to more concrete and more visual presentation in the form of diagrams that were originally from formal notation (Andriatna & Sujadi, 2024). This is intended to adjust the cognitive development of students.

In defining variables, the presentation of knowledge to be taught tends to be less explicit when compared to the presentation of scholarly knowledge sources. This phenomenon indicates the opportunity for potential didactical obstacles as stated by Suryadi (2019)

ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v15i2.13792

and Brousseau in 2002 (Fitriani et al., 2020; Sadiah et al., 2024). This learning obstacle can arise due to a mismatch between didactical design and students' cognitive readiness (Nurhayati et al., 2023; Suryadi, 2019). The occurrence of this gap suggests more attention to emphasising basic understanding strongly, such as the study of didactic transposition on the concept of limit function conducted by Sulastri (2023). The results of the research state that the introduction of the concept of a limit function that is not intuitive in the presentation of textbooks has the potential to cause learning obstacles when the formal reinforcement delivered is inadequate.

Furthermore, the researcher highlighted an innovative adaptation *to the knowledge to be taught* that invites students to explore patterns through various lines that can visually represent the data at points on the scatter diagram. This can assist students in constructing fundamental linear regression knowledge. This participatory approach will facilitate students in understanding the concepts deeply and durably. However, in the scholarly knowledge references, the introduction of the concept of the best-fit line at an early stage is approached theoretically and explicitly in the form of a detailed mathematical presentation of the least squares method (Cooper et al., 2024). The research finding that there is a lack of clarity in this transition is a concern and can potentially cause epistemological learning obstacles when in a situation of institutionalisation. This incident is also the same for the presentation of the concept of random error, which is not explicitly explained and described in the knowledge to be taught. This has an impact on understanding the regression model equation that represents the relationship between variables in contextual problems. This phenomenon can trigger conceptual ontogenic learning obstacles due to the level of understanding of the concept of random error that has not been mastered formally and operationally and is immediately introduced to the existence of residuals, squared residuals, and the sum of squared residuals. The findings of this study are in line with the results of a study which states that less explicit instructions tend to cause students' lack of understanding of the essential material to be learned (Mensah, 2023; Suarsana et al., 2024).

Cope and Prosser (2005) identified educationally critical aspects of a phenomenon that should be emphasised during the learning process to facilitate the achievement of targeted understanding. In line with a study conducted by Puspita & Kurniawan, phenomena that have the potential to cause learning obstacles in learning will allow students to experience difficulties in internalising knowledge from complex ones without an adequate conceptual foundation (Puspita & Kustiawan, 2024). The transformation of complex knowledge into a more accessible form without reducing its scientific essence is essential to this didactic transposition process.

CONCLUSIONS

With an emphasis on the effects on students' statistical literacy, this study examined the process of external didactic transposition of linear regression from scholarly knowledge to knowledge to be taught in senior high school mathematics. The results show that substantial knowledge adaptation takes place when linear regression concepts are presented, especially when it comes to the treatment of dependent

ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v15i2.13792

and independent variables, the introduction of linear regression using the best-fit line, and the creation of regression equation models. Textbooks' use of contextual representations and visual exploration is a significant adaptation that enhances students' conceptual grasp of linear regression.

However, the analysis identifies significant gaps in the transposition process. Textbooks' definitions of dependent and independent variables are often formal and insufficiently operational, and the concept of random error is introduced briefly without a solid conceptual foundation in probability and distributional assumptions. These gaps highlight potential learning obstacles that may prevent students from developing a thorough understanding of the probabilistic nature of regression and comprehensive statistical literacy.

The main contribution of this study is to provide a structured knowledge of how important linear regression concepts are changed throughout didactic levels, as well as to highlight critical locations when conceptual simplicity risks weakening conceptual accuracy. These findings provide empirical insights for curriculum developers, textbook authors, and educators to improve the alignment of scientific statistical knowledge with school mathematics instruction.

This study is limited to the examination of documents indicating external didactic transposition and does not include any empirical classroom data. Future research should investigate internal didactic transposition, namely how teachers perceive and use knowledge to be taught in classroom activities, and how this process effects students' learned knowledge. Further research is required to construct enriched didactic transposition models that clearly address outstanding conceptual difficulties, such as the distributional assumptions of random error and the mathematical rationale of the best-fit approach. Furthermore, comparative and technologically improved investigations across many curricular contexts may provide larger insights for global statistics education and statistical literacy advancement.

REFERENCES

Aityan, S. K. (2022). Linear Regression. In *Linear Regression* (pp. 915–978). https://doi.org/https://doi.org/10.1007/978-1-4842-6797-4_14

Amna Saleem, Huma Kausar, & Farah Deeba. (2021). Social Constructivism: A New Paradigm in Teaching and Learning Environment. *Perennial Journal of History*, *2*(2), 403–421. https://doi.org/10.52700/pjh.v2i2.86

Anandhi, P., & Nathiya, D. E. (2023). Application of linear regression with their advantages, disadvantages, assumption and limitations. *International Journal of Statistics and Applied Mathematics*. https://doi.org/https://doi.org/10.22271/maths.2023.v8.i6b.1463

Andriatna, R., & Sujadi, I. (2024). Didactic Transposition from Scholarly Mathematics to School Mathematics : The Case of Function Concept Mosharafa : Jurnal Pendidikan Matematika Mosharafa : Jurnal Pendidikan Matematika. *Mosharafa: Jurnal Pendidikan Matematika*, *13*(1), 215–230.

Aziz, A. M., & Rosli, R. (2021). A systematic literature review on

ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v15i2.13792

developing students' statistical literacy skills. *Journal of Physics: Conference Series*, *1806*(1). https://doi.org/10.1088/1742-6596/1806/1/012102

Callingham, R., & Watson, J. M. (2017). The development of statistical literacy at school. *Statistics Education Research Journal*, *16*(1), 181–201. https://doi.org/10.52041/serj.v16i1.223

Cooper, L. L., Rodríguez Vásquez, M. P., & Moyer, T. O. (2024). A Student-Centered Exploration of Influential Points in Linear Regression Using Desmos. *Journal of Statistics and Data Science Education*, 1–16. https://doi.org/https://doi.org/10.1080/26939169.2024.2335365

Cope, C., & Prosser, M. (2005). Identifying didactic knowledge: An empirical study of the educationally critical aspects of learning about information systems. *Higher Education*, *49*(3), 345–372.

Do, T. H., & Nguyen, V. T. T. (2020). The structure of didactic transposition capability - analysis of an example of didactic transposition of physical knowledge in the training of pedagogical students. *Vietnam Journal of Education*, *4*(1), 44–52. https://doi.org/10.52296/vje.2020.7

Droogers, M. J. S., & Drijvers, P. H. M. (2017). *Enhancing statistical literacy*. 860–867. https://doi.org/https://dspace.library.uu.nl/handle/1874/357816

El Fadel, H. (2024). The Didactic Transposition : Practices and Pedagogical Issues. *International Journal of Innovative Research in Multidisciplinary Education*, *03*(09), 1569–1581. https://doi.org/10.58806/ijirme.2024.v3i9n18

Fitriani, N., Kadarisma, G., & Amelia, R. (2020). Pengembangan Desain Didaktis Untuk Mengatasi Learning Obstacle Pada Materi Dimensi Tiga. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, *9*(2), 231. https://doi.org/10.24127/ajpm.v9i2.2686

Gök, M., Erdoğan, A., & Erdoğan, E. Ö. (2019). Transpositions of function concept in mathematics curricula and textbooks from the historical development perspective. *International Journal of Instruction*, *12*(1), 1189–1206. https://doi.org/10.29333/iji.2019.12176a

Healey, J. F., & Prus, S. G. (2015). *Statistics: A Tool for Social Research*. Nelson College Indigenous, Nelson Education.

Jacobs, J. (2023). Linear regression analysis. In *Linear regression analysis* (pp. 548–557). Elsevier eBooks. https://doi.org/https://doi.org/10.1016/b978-0-12-818630-5.10067-3

Krogh, E., Qvortrup, A., & Graf, S. T. (2021). Didaktik and curriculum in ongoing dialogue. In *Didaktik and Curriculum in Ongoing Dialogue*. Routledge. https://doi.org/10.4324/9781003099390

Makkawy, A. (2023). Linear Regression. In *Linear Regression* (pp. 113–119). Routledge eBooks. https://doi.org/https://doi.org/10.4324/b23320-21

Mann, P. S. (2011). Introductory Statistics, 7th Ed. *John Wiley and*

ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v15i2.13792

Sons, Incorporated, *8*(11), 736.

Mensah, R. O. (2023). Regression Fundamentals. *International Series in Management Science/Operations Research*, 33–55. https://doi.org/https://doi.org/10.1007/978-3-031-21480-6_3

Mezaini, D., Khemis, B., Algeria, M., & Khemmad, M. (2022). Didactics : an overview on the key concepts. *Journal of El Hikma for Philosophical Studies*, *2*(November), 1041–1049. https://www.researchgate.net/publication/365608436

Montgomery, D. C. ., Peck, E. A. ., & Vining, G. G. (2021). *Introduction to linear regression analysis (6th ed.)*. John Wiley & Sons.

Morales-López, Y., Páez, A. B., Vega, D. A., & Breda, A. (2023). Identification of characteristics of didactic and meta-didactic mathematical knowledge of novice and expert teachers when reflecting on class episodes. *Journal on Mathematics Education*, *14*(1), 149–168. https://doi.org/10.22342/jme.v14i1.pp149-168

Nurhayati, L., Priatna, N., Herman, T., & Dasari, D. (2023). Learning Obstacle pada Materi Integral (Antiderivative) dalam Teori Situasi Didaktis. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, *12*(1), 984–993.

Pal, M., & Bharati, P. (2019). Introduction to Correlation and Linear Regression Analysis. *Springer, Singapore*, 1–18. https://doi.org/https://doi.org/10.1007/978-981-13-9314-3_1

Peck, R., Olsen, C., & Decore, J. (2016). *Introduction to Statistics & Data Analysis* (Vol. 4, Issue 1).

Phuong, V. D., & Quan, N. H. (2023). The Didactic Transposition Competence of Mathematics Preservice Teachers. *Hnue Journal of Science*, *68*(3), 151–165. https://doi.org/10.18173/2354-1075.2023-0071

Prasad, S. (2024). *Regression* (pp. 1–45). https://doi.org/https://doi.org/10.1007/978-981-99-7257-9_1

Puspita, E., & Kustiawan, C. (2024). A Didactical Design Research: Knowledge Acquisition of Mathematics Teacher Prospective Students on Multiple Integral Concepts. *Kreano, Jurnal Matematika Kreatif-Inovatif*, *15*(1), 199–217. https://doi.org/10.15294/h3v97z60

Rudi, R., Suryadi, D., & Rosjanuardi, R. (2022). Didactical Transposition within Reflective Practice of an Indonesian Mathematics Teacher Community: A Case in Proving the Pythagorean Theorem Topic. *Southeast Asian Mathematics Education Journal*, *12*(1), 65–80. https://doi.org/https://doi.org/10.46517/seamej.v12i1.132

Sadiah, L. H., Suhendra, S., & Herman, T. (2024). Learning Obstacle Pada Pembelajaran Sistem Persamaan Linear Tiga Variabel Berdasarkan Praxeology. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, *13*(2), 633. https://doi.org/10.24127/ajpm.v13i2.8352

Son, A. L., Darhim, & Fatimah, S. (2020). Students' mathematical problem-solving ability based on teaching models intervention and cognitive style. *Journal on*

ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v15i2.13792

Mathematics Education, *11*(2), 209–222. https://doi.org/10.22342/jme.11.2.10744.209-222

Suarsana, I. M., Suryadi, D., Nurlaelah, E., Jupri, A., & Pacis, E. R. (2024). Didactic Transposition of Straight-Line Equations: from Scholarly Knowledge to Knowledge to be Taught. *Plusminus: Jurnal Pendidikan Matematika*, *4*(2), 287–308. https://doi.org/10.31980/plusminus.v4i2.1528

Sulastri, R. (2023). Studi didactic transposition: Eksplorasi knowledge to be taught pada limit fungsi. *Journal of Didactic Mathematics*, *4*(2), 106–117. https://doi.org/10.34007/jdm.v4i2.1903

Suryadi, D. (2019). *Penelitian desain didaktis (DDR) dan implementasinya*. Gapura Press.

Susanto, D., Sihombing, S. K., Radjawane, M. M., Candra, Y., Sinambela, D., & Penelaah. (2021). *Buku Guru Matematika SMA/SMK Kelas XI*.

Susetyo, B. (2012). *Statistika*. Direktorat Jenderal Pendidikan Islam, Kementerian Agama Republik Indonesia. http://www.lechtmanresearch.com/mages/research_

Tran, D., Nguyen, T. T. A., Nguyen, T. D., Ta, T. M. P., Huynh, T. B., Phan, T. P., Pham, T. N., & Nguyen, T. H. N. (2023). Planning for Developing Students' Statistical Literacy: A Research-Informed Framework Development. *Vietnam Journal of Education*, *7*(2), 74–81. https://doi.org/10.52296/vje.2023.227

Walpole, R. E., Myers, R. H., Myers, S. L., & Ye, K. (2012). Probability & Statistics for Engineers & Scientists Ninth Edition. In *Pearson Education Inc.*

Wang, C., Shinno, Y., Xu, B., & Miyakawa, T. (2023). An anthropological point of view: exploring the Chinese and Japanese issues of translation about teaching resources. *ZDM - Mathematics Education*, *55*(3), 705–717. https://doi.org/10.1007/s11858-023-01477-4

Weiland, T., & Sundrani, A. (2022). Opportunities for K-8 Students to Learn Statistics Created by States' Standards in the United States. *Journal of Statistics and Data Science Education*, *30*(2), 165–178. https://doi.org/10.1080/26939169.2022.2075814

Weiss, N. A. (2012). *Introductory Statistics*.

Wu, Z., & Ye, L. (2015). Mathematical Modeling in Connecting Concepts to Real World Application. In *The Proceedings of the 12th International Congress on Mathematical Education*. https://doi.org/10.1007/978-3-319-12688-3_76

Yan, W. (2022). *Correlation and Regression* (pp. 241–279). https://doi.org/https://doi.org/10.1007/978-981-19-0596-4_5

Yin, A., Chan, W., Cheung, C., & Sung, M. (2025). Enhancing students ' digital literacy skills through their technology use in a course - based research project : a Hong Kong case study. *Asia Pacific Education Review*, *0123456789*. https://doi.org/10.1007/s12564-025-10038-1