EXPLORATORY ANALYSIS OF DIFFERENTIATION ERRORS – ERROR PROFILES OF STUDENTS

Ole Weigt, Timo Leuders, Katharina Loibl, Lars Holzäpfel

University of Education Freiburg

To date, there is only limited knowledge about the errors students make when solving differentiation tasks. This study presents an exploratory investigation of students' errors in differentiation in high school–level tasks. 69 students produced a total of 2281 answers to differentiation tasks, resulting in 1231 errors. An inductive categorization of these errors revealed 26 distinct data-driven error categories. These 26 categories were assigned to four concept-driven error categories. A cluster analysis identified four groups of individuals with different error profiles based on these concept-driven categories. The findings provide insights into the heterogeneity of errors in differentiation and students' general error behaviour. Moreover, the results show that the empirical data align well with the concept-based classification of errors.

INTRODUCTION

Errors and misconceptions have been a central focus of research in mathematics education for many years (Radatz, 1979). Their investigation serves multiple purposes in both research and practice, including: understanding how mathematical concepts and procedures are learned, developing and implementing instructions that fosters learning through the analysis and explanation of errors, and supporting teachers in preventing errors and misconceptions (Neidorf et al., 2020; Ohlsson, 1996).

Calculus is central in secondary and tertiary education around the world. It is required for different professions and is often associated with difficulties (Rasmussen et al., 2014). There is substantial research on errors with respect to conceptual knowledge, i.e. misconceptions about differentiation (Greefrath et al., 2016; Rasmussen et al., 2014), but only limited work on errors with respect to procedural knowledge, i.e. differential calculus or, specifically in computing derivatives could be found (e.g. (Dorner et al., 2025; Siyepu, 2013)).

The study has two aims: 1. To generate an overview and a categorization of the errors that students make when answering differentiation tasks on high school level. 2. To identify patterns in these errors.

THEORETICAL FRAMEWORK

Reasons for procedural errors in differentiation

Research on procedural knowledge errors distinguishes between rational errors and slips. "A slip is the error that occurs when a person does an action that is not intended"

2026. In H. Krzywacki, M. Hähkiöniemi, E. Haataja, M.S. Hannula & P. Hästö (Eds.). *Proceedings of the 49th Conference of the International Group for the Psychology of Mathematics Education: Research Reports*, pp. 1579-1586. PME.

(Norman, 1981 S.1). Rational errors (sometimes called bugs) are errors that occur by using an inappropriate rule to a task, that has worked in the past for a similar appearing task (Ben-Zeev, 1998; Brown & Vanlehn, 1980). Some publications additionally list factual errors. Ben-Zeev (1998) excludes fact errors from rational errors.

In other studies most of the errors that occur can be categorized as rational errors (Ben-Zeev & Ronald, 2002; Brown & VanLehn, 1980). The majority of these errors arise from the inappropriate of a procedure, particularly when the procedure has been learned only from a few procedural examples rather than through a profound introduction and practising (Ben-Zeev & Ronald, 2002). Anderson (1982) links errors in application, such as slips, to a high working memory load due to an incomplete proceduralization of the procedure.

Studies on errors in arithmetic procedures among children in grades 3–5 have identified some domain-unspecific behaviours: Students may use different incorrect procedures for two tasks with similar structure but different numbers; this is referred to as tinkering (VanLehn, 1982). It has been shown that error-producing behaviour is unstable over periods of 1–3 months and even within a single test session (Hennessy, 1994). However, while specific errors to single tasks alter, the type of errors seem to remain stable, as Ben-Zeev & Ronald (2002) found in a study with university students. Brown & VanLehn (1980) proposed four general heuristics that children use when they are unable to solve a given task: 1. Escape and flee 2. Relocate/refocus the operation 3. Overgeneralization 4. Dememorization (reverting to a more basic procedure).

Concept-based categories of errors in differentiation

Errors in computing the derivates can either occur in the algebraic part or the differentiation part of the procedure. Therefore, the first concept-driven category is 1. Algebraic errors. Other studies suggest that algebraic errors dominate in derivation (Stewart et al., 2018).

Differentiation is the application of procedural knowledge (Rittle-Johnson & Schneider, 2014), which includes selecting an appropriate procedure and applying it (Anderson, 1982). In the selection step both fact errors and rational errors can occur. Accordingly, the errors in the differentiation part can be classified as 2. Interpretation errors, arising from incorrect recognition of aspects of the expression (e.g. treating π^2 as if π were a variable representing an element of the domain), 3. Overgeneralization errors, resulting from selecting and applying a rule that is inappropriate for the given expression but works for similar tasks (e.g., differentiating each factor separately instead of using the product rule in $f(x) = u(x) * v(x)$), 4. Application errors, resulting from selecting a correct rule but failing to apply it correctly (e.g., omitting the negative sign when differentiating negative exponents: $f(x) = x^{-2} \rightarrow f'(x) = 2x^{-3}$).

METHODS

Participants and Setting

A total of 69 first- and second-year university students enrolled in lower secondary mathematics teacher education participated in the study. The test was administered as part of a regular course session. Ten students had completed high school half a year before the study, 34 one and a half years before, and the remaining 25 at least two and a half years before. Sixty-two students reported that their last contact with differential calculus had been in their high school final examination.

In the lecture before the test, 25 minutes of regular course time were used to review differentiation rules and provide practice.

Measures

The test consisted of three blocks, each containing 16 different differentiation tasks. In addition, the first two blocks included two integration tasks, the third block one integration task. The three blocks included the same 16 types of differentiation tasks, but with different numerical values and in a different order. The tasks were designed to cover all rules that German high school students are expected to master according to the curriculum: the power rule, the sum rule, the product rule, the chain rule, the rules for sine and cosine and the rule for exponential functions especially e-functions.

Data collection

Students had 25 minutes to take the test and were instructed to work individually without any technological devices or external assistance. The test length was designed to capture the spontaneous solution without long revisions; consequently only 17% of the students completed the whole test. After the 25 minutes, they were asked to add a brief explanation of their strategy for each task they had attempted.

Data analysis

A qualitative content analysis with a mixed approach, including both data-driven and concept-driven categories, was conducted (Kuckartz, 2019). The responses were coded in four stages. First, each response was classified as correct, erroneous, skipped or not attempted. Answers were marked as skipped, when no or only partial solutions were written down ("$f'(x) = x +$ ") and other tasks later on the same page were worked on. Second, each error received inductively a code describing the error. Third, the errors were classified inductively into data-driven categories. Fourth, the coded errors were assigned to the four concept-driven categories that originated from the literature – Algebraic errors, Interpretation errors, Overgeneralization and Application errors – as well as to a residual category, other errors (Anderson, 1982; Rittle-Johnson & Schneider, 2014; Stewart et al., 2018).

Statistical analysis:

The four concept-driven categories were used to identify error profiles by means of a cluster analysis. The concept-driven categories were normalized by dividing, for each person and each category, the number of errors by the 90th percentile of that category in order to account for the different number of opportunities for different errors to occur (Bardy et al., 2024).

RESULTS

Qualitative results

In total 1231 errors were produced and grouped into 26 data-driven categories (Table 1). Evidence of tinkering behaviour can be observed: errors are not consistent. Some participants made errors from different data-driven categories on the same task type within one test, or solved tasks of the same type correctly at other points in the test, across all four concept-driven error categories.

Algebraic errors: Various Algebraic errors can be observed, with missing brackets being the most frequent data-driven category. However, algebraic errors do not dominate the differentiation errors: only roughly one eight of all errors are algebraic errors.

Interpretation errors: Two data-driven categories of Interpretation errors can be identified. 1. Treating constants as variables: π, e, and some integers were treated as variables. 2. Failure to recognise compositions: compositions of two functions requiring the chain rule were not recognised as such, and only one function was differentiated, even though the same student correctly applied the chain rule in other items.

Overgeneralization: The most frequent Overgeneralization error is linear extrapolation: a composition of functions was treated as a linear combination. Four data-driven categories of linear extrapolation errors can be identified: products of functions, compositions requiring the chain rule, quotients of functions, and functions in the denominator of fractions. Moreover, two data-driven categories of overgeneralizations regarding exponential functions of the form a^x can be identified: treating a^x as e^x and applying the standard power rule. For functions of the form $f(x) = e^{u(x)}$, three data-driven categories of overgeneralization errors can be observed: treating $e^{u(x)}$ as e^x, using both the exponential rule and the rule for e^x, and linear extrapolation.

Application errors: Three different data-driven categories of Application errors can be identified: 1. Errors in the new factor after differentiation, 2. errors in the new exponent, and 3. factual errors in applying the product rule, chain rule or the rules for sine and cosine.

Other errors: Some slips such as omitting numbers or copying incorrectly, can be observed. Another data-driven category comprises rewriting a task into a seemingly simpler one. In addition, two data-driven categories capture the choice of an

overcomplicated procedure: using the product rule on factors and unnecessary expanding brackets before differentiating. These are technically not errors but can lead to errors.

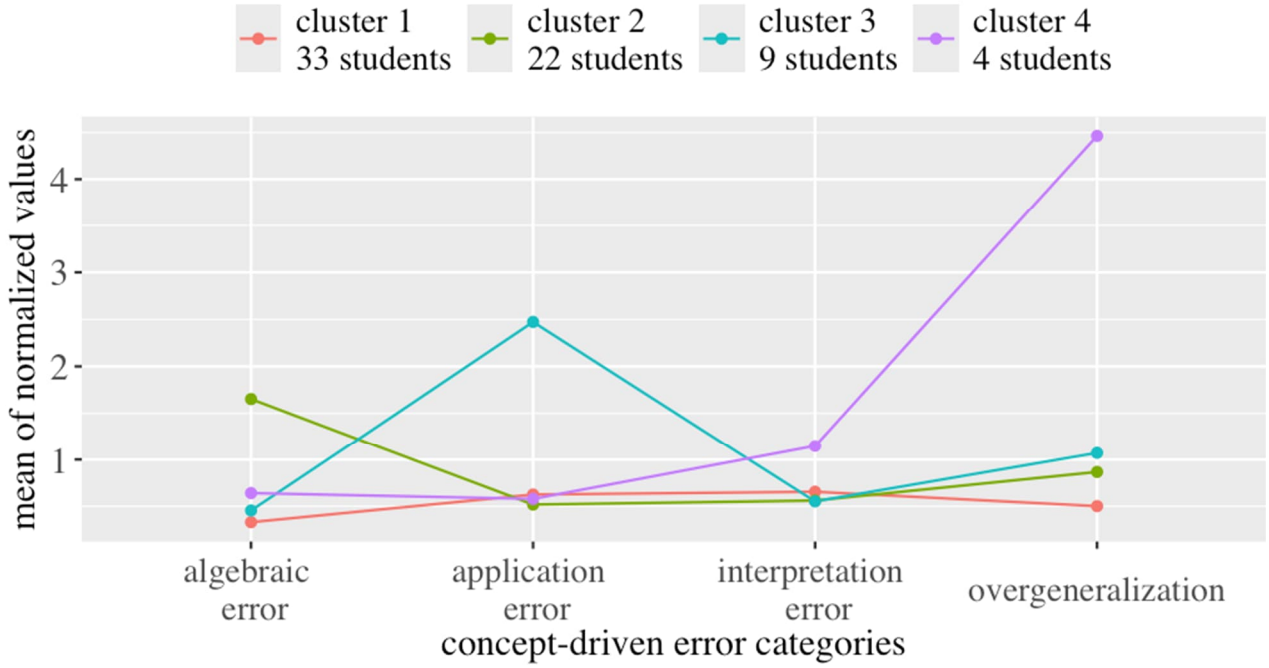
Table 1: Qualitative Results: data-driven categories with examples and assignment to the concept-driven categories

Concept-driven categories	Data-driven categories	Examples
Algebraic errors	① expanding brackets ② missing brackets ③ exponent calculation ④ other algebraic errors	③ $x^2 + 2x + x^3 = 2x^6$
Interpretation errors	① constants as variables ② not recognising need for chain rule	② $f(x) = x * \sin(2x)$, $f'(x) = 1 \cdot (\sin(2x)) + x \cdot \cos(2x)$
Overgeneralization	① linear extrapolation of products ② linear extrapolation of fractions of functions ③ linear extrapolation instead of chain rule ④ linear extrapolation of functions as denominators ⑤ exponent rule for exponential functions ⑥ rule for e^x overgeneralized for all exponential functions ⑦ rule for e^x overgeneralized for $e^{u(x)}$	① $f(x) = -2\sin(x) \cdot x^5$, $f'(x) = -2\cos(x) \cdot 5x^4$ ② $f(x) = \frac{-\cos(x)}{x^3}$, $f'(x) = \frac{-\sin(x)}{3x^2}$ ④ $f(x) = \frac{5}{x+1}, f'(x) = \frac{5}{1+0}$ ⑤ $f(x) = 3^x, f'(x) = x3^{x-1}$ ⑥ $f(x) = 3^x, f'(x) = 3^x$
Application errors	① ignoring negativity of exponents ② error in new exponent ③ error in new factor ④ error in product rule ⑤ error in chain rule ⑥ error in sine and cosine ⑦ partly mixing up integration and differentiation ⑧ error in exponential functions	① $f(x) = x^{-5}$, $f'^{(x)} = -5x^{-4}$ ③ $f(x) = x^3$, $f'(x) = 2x^2$ ⑤ $f(x) = \sin(2x^3)$, $f'(x) = 6x^2 + \cos(2x^3)$ ⑦ $f(x) = 2x^4, F(x) = 6x^5$
Other errors	① omitting numbers ② copying incorrectly ③ rewriting incorrectly ④ unnecessary and overcomplicate expanding ⑤ product rule for factors	④ $(6x^4 + 8)^3 = 512 + 1152\,x^4 + 864\,x^8 + 216\,x^{12}$

Quantitative results

A cluster analysis, using the concept-driven categories as independent variable, reveals four distinct clusters within the sample (Figure 1). In the first cluster (33 students), no single error concept-driven category dominates, and the number of errors of each concept-driven category is comparatively low. The second cluster (22 students) is characterised by a high number of Algebraic errors. In the third cluster (9 students), Application errors are predominant. The fourth cluster (4 students) is characterised by the dominance of Overgeneralization errors. One participant cannot be assigned to any of these clusters.

Figure 1: Cluster analysis of the concept-driven error categories with normalized values. The 90th percentile was used to normalize.



DISCUSSION

The qualitive findings revealed 26 data-driven categories of errors showing the breath of different errors students make in differentiation. These categories could be successfully assigned to four concept-driven error categories derived from the domain-unspecific literature on procedural knowledge. The findings show that students make errors on all stages of solving differentiation tasks: recognition of the aspects of the expression, selecting an appropriate procedure, applying that procedure and performing the algebra before and after differentiation. In addition to generalization the escape and flee heuristic observed in younger children can also be identified in adults in the form of skipping items and incorrectly rewriting a task.

The four different error profiles found in the cluster analysis indicate that the students are at a different stage of proceduralizing (Anderson, 1982): Higher cognitive load due

to limited proceduralization leads to Application errors (cluster 3); insufficient tuning (improving the selection step of the procedure) leads to Overgeneralization (cluster 4), and a high level of proceduralization leads to a small amount of errors in all concept-driven categories (cluster 1). In addition, Algebraic errors seem to be largely independent of the level of proceduralization of differentiation procedures and are specific for cluster 2. The identification of cluster 2 and the low proportion of Algebraic errors among all errors refine the hypothesis of Stewart et al. (2018), who suggest that algebra is the root of difficulties of students in calculus courses: Algebra seems to be the main difficulty in differentiation for a specific group of students.

The identification of four profiles in combination with the observation of tinkering supports previous finding on general error behaviour in other age groups and other domains: there is high variability in individual errors but stability in concept-driven error categories (Ben-Zeev & Ronald, 2002; Hennessy, 1994).

The specific characteristics of the sample – students enrolled in mathematics for lower secondary school teaching – may limit the types of errors and error profiles found in this study. A larger and more diverse sample might yield additional data-driven categories and additional error profiles in the cluster analysis. Moreover, the design of the test items provided different numbers of opportunities for different errors to occur. We addressed this limitation by normalizing the data (as also done by (Bardy et al., 2024)).

Overall, these findings underscore that while students display a wide variety of data-driven error categories in differentiation, their errors, when organized into concept-driven categories reliably cluster into a small number of profiles that reflect distinct stages of proceduralization. Based on these results, support measures could be developed for the respective error profiles.

References

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, *89*(4), 369–406. https://doi.org/10.1037/0033-295X.89.4.369

Bardy, T., Holzäpfel, L., Reinhold, F., & Leuders, T. (2024). Mathematics teachers' multiple perspectives on adaptive tasks: Task evaluation and selection as core practices for teaching quality. *ZDM – Mathematics Education*, *56*(5), 981–996. https://doi.org/10.1007/s11858-024-01626-3

Ben-Zeev, T. (1998). Rational errors and the mathematical mind. *Review of General Psychology*, *2*(4), 366–383. https://doi.org/10.1037/1089-2680.2.4.366

Ben-Zeev, T., & Ronald, J. (2002). Is procedure acquisition as unstable as it seems? *Contemporary Educational Psychology*, *27*(4), 529–550. https://doi.org/10.1016/S0361-476X(02)00005-X

Brown, J. S., & Vanlehn, K. (1980). *Repair Theory: A Generative Theory of Bugs in Procedural Skills*.

Dorner, C., Ableitinger, C., & Krammer, G. (2025). Revealing the nature of mathematical procedural knowledge by analysing students' deficiencies and errors. *International Journal of Mathematical Education in Science and Technology*, 1–22. https://doi.org/10.1080/0020739X.2024.2445666

Greefrath, G., Oldenburg, R., Siller, H.-S., Ulm, V., & Weigand, H.-G. (2016). Aspects and "Grundvorstellungen" of the Concepts of Derivative and Integral: Subject Matter-related Didactical Perspectives of Concept Formation. *Journal Für Mathematik-Didaktik*, *37*(S1), 99–129. https://doi.org/10.1007/s13138-016-0100-x

Hennessy, S. (1994). The stability of children's mathematical behavior: When is a bug really a bug? *Learning and Instruction*, *3*(4), 315–338. https://doi.org/10.1016/0959-4752(93)90022-R

Kuckartz, U. (2019). Qualitative Text Analysis: A Systematic Approach. In G. Kaiser & N. Presmeg (Hrsg.), *Compendium for Early Career Researchers in Mathematics Education* (S. 181–197). Springer International Publishing. https://doi.org/10.1007/978-3-030-15636-7_8

Neidorf, T., Arora, A., Erberber, E., Tsokodayi, Y., & Mai, T. (2020). *Student Misconceptions and Errors in Physics and Mathematics: Exploring Data from TIMSS and TIMSS Advanced* (Bd. 9). Springer International Publishing. https://doi.org/10.1007/978-3-030-30188-0

Norman, D. (1981). Categorization of Action Slips. *Psychological Review*, *88*, 1–15. https://doi.org/10.1037//0033-295X.88.1.1

Ohlsson, S. (1996). Learning from performance errors. *Psychological Review*, *103*(2), 241–262. https://doi.org/10.1037/0033-295X.103.2.241

Radatz, H. (1979). Error Analysis in Mathematics Education. *Journal for Research in Mathematics Education*, *10*(3), 163–172. JSTOR. https://doi.org/10.2307/748804

Rasmussen, C., Marrongelle, K., & Borba, M. C. (2014). Research on calculus: What do we know and where do we need to go? *ZDM*, *46*(4), 507–515. https://doi.org/10.1007/s11858-014-0615-x

Rittle-Johnson, B., & Schneider, M. (2014). Developing Conceptual and Procedural Knowledge of Mathematics. In R. Cohen Kadosh & A. Dowker (Hrsg.), *The Oxford Handbook of Numerical Cognition* (1. Aufl., S. 1118–1134). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199642342.013.014

Siyepu, S. W. (2013). An exploration of students' errors in derivatives in a university of technology. *The Journal of Mathematical Behavior*, *32*(3), 577–592. https://doi.org/10.1016/j.jmathb.2013.05.001

Stewart, S., Reeder, S., Raymond, K., & Troup, J. (2018). Could algebra be the root of problems in calculus courses. *Proceedings of the 21st annual conference on research in undergraduate mathematics education*, 1023–1030.

VanLehn, K. (1982). Bugs are not enough: Empirical studies of bugs, impasses and repairs in procedural skills. *The Journal of Mathematical Behavior*, *3*(2), 3–71.