Assessment in Collaborative Learning: a Mediation Analysis Approach

Evaluación en Aprendizaje Colaborativo: un enfoque de análisis de mediación

Luís Cavique
Universidade Lisboa. Lisboa, Portugal
luis.cavique@uab.pt

M. Rosário Ramos
Universidade Lisboa. Lisboa, Portugal
mariar.ramos@uab.pt

Abstract

In collaborative learning, evaluating the process involves teamwork dynamics, and assessing the product focuses on the accuracy and quality of the final output. Assessment plays a crucial role, as it defines and measures the effectiveness of group activities to ensure that learning objectives are met. Mediation analysis is an important technique to better understand relationships between variables, specifically designed to test hypotheses about potential causal effects in various areas. However, many research initiatives have been discontinued prematurely due to the Baron-Kenny data restrictions. This research takes a case study of online learning from the Portuguese Open University to determine if and how group selection and interaction frequency affect individual assessment. The contribution lies in applying quantitative causal mediation analysis to collaborative learning assessment. The Lambda Mediation Ratio is proposed to enhance mediation analysis by enabling quick and flexible categorization into full, partial, or no mediation. Using Moodle platform logs and student outcomes, it was possible to find a significant influence of group dynamics on academic performance, highlighting the practical application of this improved methodology in an educational context. These findings reassure us of the relevance and applicability of this research in real-world educational settings.

Keywords: distance education, collaborative learning, assessment, direct acyclic graph, mediation analysis

Resumen

En el aprendizaje colaborativo, evaluar el proceso implica dinámicas de trabajo en equipo, y evaluar el producto se centra en la precisión y calidad del resultado final. La evaluación juega un papel crucial, ya que define y mide la eficacia de las actividades grupales para garantizar que se cumplan los objetivos de aprendizaje. El análisis de mediación es una técnica importante para comprender mejor las relaciones entre variables, diseñada específicamente para probar hipótesis sobre posibles efectos causales en diversas áreas. Sin embargo, muchas iniciativas de investigación se han interrumpido prematuramente debido a las restricciones de datos de Baron-Kenny. Esta investigación toma un estudio de caso de aprendizaje en línea de la Universidad Abierta de Portugal para determinar si la selección de grupo y la frecuencia de interacción afectan la evaluación individual y cómo. La contribución radica en la aplicación del análisis cuantitativo de mediación causal a la evaluación del aprendizaje colaborativo. Se propone el índice de mediación Lambda para mejorar el análisis de mediación al permitir una categorización rápida y flexible en mediación total, parcial o nula. Utilizando los registros de la plataforma

Moodle y los resultados de los estudiantes, fue posible encontrar una influencia significativa de la dinámica de grupo en el rendimiento académico, destacando la aplicación práctica de esta metodología mejorada en un contexto educativo. Estos hallazgos nos aseguran la relevancia y aplicabilidad de esta investigación en entornos educativos del mundo real.
Palabras clave: educación a distancia, aprendizaje colaborativo, evaluación, gráfico acíclico directo, análisis de mediación.

1. Introduction

Assessment plays a crucial role in collaborative learning. It helps define and measure the effectiveness of group activities, ensuring that learning objectives are met. Effective assessment strategies, such as rubrics and checklists, promote essential skills and improve collaborative education.

Evaluating the knowledge construction process and the final product is essential in assessing collaborative learning. (i) The process examines how students interact, solve problems, and work together during group activities. This includes evaluating teamwork dynamics as well as the individual frequency and quality of the students' proposals. (ii) On the other hand, the product refers to the final output of the group's work. This can be a project, presentation, report, or any other deliverable demonstrating the group's collective understanding and application of the subject matter. Assessing the product focuses on the work's accuracy of the content and quality of the presentation.

While current assessment methods effectively evaluate collaborative skills and processes qualitatively, empirical data and measurable metrics are lacking. Integrating quantitative methods would provide a more comprehensive understanding of the impact and effectiveness of collaborative learning practices.

Objectives

The objective is to integrate quantitative methods in the collaborative learning assessment. In particular, given a specific dataset from the Portuguese Open University's second cycle (master's degree) and a set of evaluating rules, the research objectives are the following:
Q1: Does the selection of a group impact an individual's final assessment?
Q2: Does the frequency of interactions affect an individual's final assessment?

Q1 is closely related to the product since the group's selection impacts the report's quality. On the other hand, Q2 is related to the learning process, which is influenced by the frequency of interactions.

Contribution

This work significantly contributes by applying a quantitative approach, causal mediation analysis, to collaborative learning assessment.

Mediation analysis [Baron, Kenny 1986] is a statistical technique used to understand how an independent variable influences a dependent variable through a mediator variable. Despite the numerous citations of the Baron-Kenny method, many research initiatives have been prematurely discontinued due to the method's data restrictions [Zhao et al. 2010]. This work proposes the Lambda mediation ratio to extend mediation analysis. The procedure based on the Mediation Ratio allows the rapid classification of mediation into one of three types (full, partial, or no mediation).

Organization
The remaining article is organized as follows: Section 2 describes the related work. Section 3 presents the mediation analytic concepts. Section 4 shows our flexible approach for the mediation analysis, the lambda mediation ratio. Section 5 describes a use case in collaborative learning. Finally, Section 6 draws some conclusions.

2. Related work

This section examines the assessment of collaborative learning and mediation analysis in education, focusing on existing methods, challenges, and gaps in quantitative approaches to evaluate the effectiveness and outcomes.

Assessing collaborative learning
Valente [2016] offers practical guidelines for assessing collaborative learning in classrooms. It underscores the importance of defining the purpose of assessment (formative, summative, and self-assessment) and aligning this with the tools used, such as rubrics and checklists. When used effectively, these tools can empower educators to evaluate collaborative skills. The article advocates for the central role of collaborative assessment in course design, promoting essential 21st-century skills and encouraging adaptation to different teaching and learning contexts, thereby making you, our esteemed audience, feel capable and empowered in your teaching practices.

Boud and Bearman's [2022] discussion on the tension between individual assessment methods and collaborative learning practices is particularly insightful. The authors identify four key dilemmas: distinguishing collaboration from collusion, ensuring fairness in group assessments, directly assessing teamwork outcomes, and fostering a culture of peer assessment. They argue for normalizing collaborative learning within course structures rather than being an afterthought. Trust is emphasized as a crucial element for effective collaboration, and the authors propose a shift from traditional grading to support collaborative efforts better. They suggest comprehensive course design changes to embed collaboration as a central element, providing rich, varied experiences that mirror real-world practices.

Khan A. et al. [2022] present an e-learning system with significant potential for enhancing collaborative learning. The study integrates e-learning systems and Machine Learning (ML) into daily educational activities, focusing on content preparation and presentation tools for teachers. It highlights that Information and Communication Technologies supported teaching practices foster a collaborative and participative learning environment, improving student performance and engagement. The proposed system leverages ML algorithms to predict and evaluate student performance accurately. The paper also stresses the importance of professional training for teachers to maximize the benefits of technology-enhanced teaching methods. Finally, it introduces a custom machine-learning process maturity model and addresses challenges in applying ML in software engineering contexts.

Despite the valuable insights provided by Valente [2016] and Boud and Bearman [2022], they primarily focus on qualitative methods such as rubrics, checklists, and course design changes, which, while effective, do not provide measurable data on student performance and collaboration outcomes. Khan et al. [2022] introduce the potential of machine learning to predict and evaluate student performance, yet their emphasis remains on the qualitative impact of technology on collaborative learning. This gap suggests a need for more quantitative research to empirically validate the effectiveness of collaborative learning strategies and provide concrete metrics for educators to enhance their teaching methods.

Mediation analysis

Mediation analysis is a statistical technique that helps to understand the mechanism underlying the relation between an independent and a dependent variable through a third hypothetical variable known as a mediator variable. Mediation analysis is a subproblem of a more general problem called path analysis. Path analysis deals with complex models, including multiple mediators, moderators, and their interactions, whereas mediation analysis focuses on models with three variables.

Mediation analysis was initially proposed by Baron and Kenny [1986]. The simplicity of the Baron-Kenny method quickly conquered the world of social sciences. In 2014, the article was 33rd on the list of the most cited scientific articles ever, ahead of authors such as Einstein or Freud [Pearl, Mackenzie 2019].

Mediation analysis has been applied in education. Misunas et al. [2024] explore the association between the education and outcome variables (pregnancy) using the mediator variable 'friendships'. They conclude that no evidence of meaningful mediation exists, indicating the benefits of girls' school attendance despite any potential risks arising from their friendships. Kratz et al. [2022] investigate the impact of parental resources on children's opportunities using causal mediation analysis, focusing on the origin-education-destination relationship in Germany. The study highlights the role of causal mediation analysis in clarifying effect definitions and facilitating the examination of how education influences social status.

Riofrío-Calderón and Ramírez-Montoya [2023] examine the integration of mediation models in online learning environments. It systematically reviews 61 articles from 2015 to 2021, focusing on satisfaction, collaboration, and self-regulation. The review highlights these studies' prevalence of pedagogical, technological, affective, and cognitive factors. The findings emphasize the importance of mediation in online education, providing insights and recommendations for designing effective online learning programs. Limitations include database restrictions and focus on specific years impacted by the COVID-19 pandemic.

Finally, quantitative approaches like mediation analysis in assessing collaborative learning are scarce in the literature, and the subject is addressed in this work.

3. Mediation Analysis Concepts

This section presents Baron and Kenny's [1986] method and highlights its significance in the evolution of mediation analysis. The discussion includes recent works, providing a comprehensive view of the field's progression.

3.1. Direct Acyclic Graph

In evaluating data-driven models' recent movements are presented in favor of explanatory models. The difference between correlational analysis and causality is at the heart of the controversy over prediction and explanation. In data science, two tasks must be distinguished: prediction and explanation.

In prediction, two variables are used: the independent variable X and the dependent variable Y. The original data is divided into the training and testing data sets to find the Y=f(X) function, where X is a covariate, and Y is the outcome.

A new variable type should be included: the intervention/treatment T. In this task, outcome Y of treatment T is the subject of the study. For this purpose, test and control datasets are used for treatment accomplished T=1 and not accomplished T=0. In analogy with Y=f(X), the explanatory function uses three variables, Y=f (T, X).

Direct Acyclic Graphs, DAG, is a theoretical concept and a practical tool in causal representation. They describe the causal assumptions of each study [Pearl et al. 2016]. By understanding and utilizing them, you can feel empowered and capable in your research.

One of the biggest causality problems in observational studies is spurious or confounding relationships. If we associate the relation T→Y, a third variable X, can be a mediator between T and Y (T→X→Y) or a covariate that influences the two variables (T←X→Y), depending on the direction of the relation. The Pearl's back-door path is any path from T to Y that starts with an arrow pointing to T, i.e., X→T→ Y. If we lock the back door X→T, the variables T and Y will not be confounded. Figure 1 shows a confounding DAG and a mediation DAG.

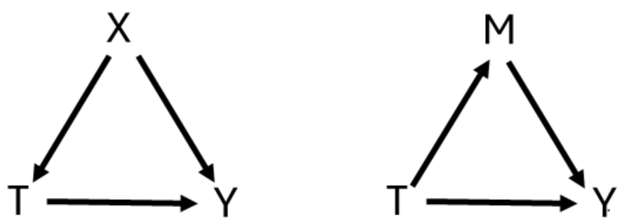

Figure 1. Confounding and mediation DAG

3.2. Mediation procedure

One of the first controlled experiments was carried out by the physician James Lind in 1747, who administered six different treatments to sailors with scurvy, concluding that citrus fruits cured patients. The correspondent DAG is Citrus → ~Scurvy, that is, citrus are the cause of non-scurvy, or citrus avoid scurvy. However, the reason for the cure is not evident, and the answer would only come in the 1930s with the discovery of vitamin C. On the path (citrus → vitamin C → ~scurvy), the variable Vitamin C mediates the

DOI: http://dx.doi.org/10.6018/red.606551

relation. Therefore, mediation analysis is the tool that allows a better understanding of the association of two variables [Pearl, Mackenzie 2019].

Baron and Kenny [1986] define the principles for detecting mediation within three variables. The procedure is based on three regression equations:
i) the dependent variable Y is explained by the independent variable T:

$$Y = i_1 + c.T$$

ii) mediator M is explained by the independent variable T:

$$M= i_2 + a.X$$

iii) and finally, the dependent variable Y is explained by the independent variable T and the mediator M:

$$Y = i_3 + c'.T + b.M$$

Figure 2 shows the variables T, M, and Y and the estimators (i) c, (ii) a, and (iii) b and c'. Baron-Kenny procedure advises p-values <= 0.001 for all estimators.

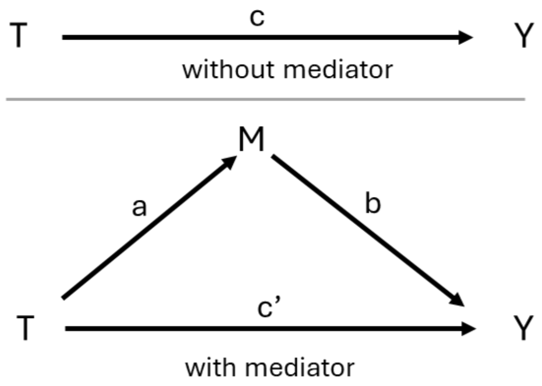

Figure 2. Variables T, M, Y, and estimators

3.3. Mediation terminology and estimators

In the context of mediation analysis [Baron, Kenny 1986], the terms 'Total Effects', 'Indirect Effects', and 'Direct Effects' refer to different components of the effect of an independent variable on a dependent variable, mediated by a third variable:

- Total Effects (TE): This is the overall effect of the independent variable on the dependent variable. It encompasses both the direct and indirect effects.
- Indirect Effects (IE): This is the part of the effect mediated through the mediator variable. The portion of the total effect can be attributed to the pathway through the mediator.
- Direct Effects (DE): This is the effect of the independent variable on the dependent variable that the mediator does not mediate. The portion of the total effect is direct and not through the mediator.

The equation TE = IE + DE reflects that the total effect is the sum of the indirect and direct effects. Additionally, the Mediation Ratio (MR) is used to understand the proportion of the total effect that is mediated. It is calculated as MR = IE/TE. This ratio

DOI: http://dx.doi.org/10.6018/red.606551

indicates how much of the total effect is due to the indirect path through the mediator. A higher MR suggests a more substantial mediation effect.

Table 1 shows the mediation terminology and the estimators found through the linear regressions.

Table 1. Mediation terminology and estimators

Mediation terminology	**Estimator**
Total Effects = TE = IE+DE	c
Indirect Effects = IE	a.b = c-c'
Direct Effects = DE	c'
Mediation Ratio = MR = IE/TE	a.b/c

3.4. Other approaches

Zhao et al. [2010] criticize Baron and Kenny's framework for mediation analysis, highlighting its limitations and misapplications. It suggests that focusing only on 'full mediation' ignores significant 'partial mediation' cases and potential omitted mediators. The authors propose a tree diagram with five leaves to determine the type of mediation and the relaxation of specific tests. They advocate using the more rigorous and robust bootstrap test to test the significance of an indirect effect (a.b) instead of testing via a z-test [Sobel 1982].

Iacobucci [2012] argues that while mediation testing is established for continuous variables, there is a gap in methods for categorical mediators or dependent variables. The author suggests a solution by utilizing logistic regression.

Pearl et al. [2016] present a toolkit for mediation using the do-calculus and differentiating the controlled direct effect from the natural direct effect.

Lange et al. [2017] offer an instructional guide on mediation analyses to uncover the causal processes behind specific cause-and-effect relationships. They also illustrate how mediation analyses can be performed using standard R package software.

MacKinnon et al. [2019] highlight how potential outcomes estimators align with traditional models under certain conditions and, through a simulation study, show that omitting the treatment-mediator interaction can affect the power to detect mediated effects. To the third equation is added an interaction between T and M:

$$Y = i_3 + c'.T + b.M + h.T.M$$

Hair et al. [2021] advocate that mediation helps researchers understand the underlying relationship mechanisms between exogenous and endogenous constructs. While the fundamental structural equation modeling analysis might focus on a single mediator construct, more complex models can include multiple mediators analyzed concurrently. The authors discuss the various mediation models and illustrate a case study.

DOI: http://dx.doi.org/10.6018/red.606551

Most of the cited authors add new information to the Baron-Kenny procedure by utilizing logistic regression [Iacobucci 2012], do-calculus [Pearl et al. 2016], treatment-mediator interaction [MacKinnon et al. 2019], or multiple mediations [Hair et al. 2021]. Zhao et al. [2010] criticized Baron-Kenny's procedure.

4. Lambda Mediation Ratio

Despite the popularity of the Baron-Kenny method, many research initiatives have been discontinued prematurely due to the Baron-Kenny data restrictions [Zhao et al. 2010]. The restrictions imposed by p-values <= 0.001 and the lack of parameters to guide the mediation types create difficulties for the data analyst to find any mediation.

They also propose a tree diagram with five leaves instead of the three Baron-Kenny types (full, partial, no mediation) to determine the kind of mediation. However, the authors do not make the method more flexible.

This study proposes a more adaptable approach to mediation analysis by easing the traditional requirements and defining a broader range of acceptable mediation outcomes. It presents a practical application in the context of distance education.

Given three variables, treatment T (independent variable), outcome Y (dependent variable), and mediator M, the goal of this work is to improve the use of mediation analysis so that the type of mediation (full, partial, or no mediation) can be quickly obtained.

In this proposal, we relax Baron-Kenny data restrictions, like the p-values <= 0.001, keeping their three types of mediation. To guide the types of mediation (full, partial, or no mediation), this work proposes a procedure based on a parameter λ, which drives the mediation ratio, as shown in Figure 3.

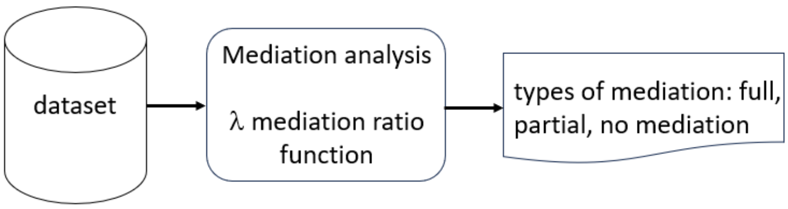

Figure 3. Schema of the procedure

When working with actual data and using Baron-Kenny constraints, most of the study leads to non-mediation. Instead of the data analyst defining a 'full' or 'partial' mediation, a new parameter is added to the procedure to guide the mediation classification. The parameter $\lambda \in [0,1]$ is compared with the Mediation Ratio.

First, the parameter should be defined λ=0, …, 0.20, 0.25, 0.30, … 0.5. Then, the type of mediation is given by:

- Full mediation if Mediation Ratio > (1-λ)
- Partial mediation if λ<= Mediation Ratio <= (1-λ)
- No mediation if the Mediation Ratio < λ

DOI: http://dx.doi.org/10.6018/red.606551

As in Lange et al. [2017], the software R [R Core Team 2023] is adopted using the standard linear regression libraries (see Listing 1).

```
lambda =0.20

MediationRatio <-function (a, b, c)
{ ratio= abs(round((a*b)/c,3))
  if (ratio <= lambda) message ("Mediation Ratio = ", ratio, ' No mediation')
  else if (ratio <= (1-lambda)) message ("Mediation Ratio = ", ratio, ' Partial mediation')
  else message ("Mediation Ratio = ", ratio, ' Full mediation')
}

MediationAnalysis<-function (df, T, M, Y)
{  c=lm(Y ~ T, df); summary(c)
    varC = summary(c)$coefficients[2,"Estimate"]; varCp=summary(c)$coefficients[2,"Pr(>|t|)"]
    writeVar("C", varC, varCp)

   a=lm(M ~ T,df);  summary(a)
   varA = summary(a)$coefficients[2,"Estimate"]; varAp=summary(a)$coefficients[2,"Pr(>|t|)"]
   writeVar("A", varA, varAp)

   bc1 = lm(Y ~ T + M,df); summary(bc1)
   varB = summary(bc1)$coefficients[3,"Estimate"];  varBp=summary(bc1)$coefficients[3,"Pr(>|t|)"]
   varC1 = summary(bc1)$coefficients[2,"Estimate"]; varC1p=summary(bc1)$coefficients[2,"Pr(>|t|)"]
   writeVar("B", varB, varBp);   writeVar("C1", varC1, varC1p)

   message("Indirect Effects(A*B)= ", round((varA*varB),3))
   message("Direct Effect (C1)   = ", round(varC1,3))
   message("Total Effects (C)    = ", round(varC,3))

   MediationRatio(varA, varB, varC)
}
```

Listing 1. Mediation Ratio and Mediation Analysis functions

The procedure based on the Mediation Ratio and the more or less relaxed value of Lambda allows the rapid classification of mediation into one of the three types (full, partial, no mediation).

5. Collaborative Learning Experiment

This work uses evaluation data from the Portuguese Open University's second cycle (master's degree). Cavique [2023] provides a detailed view of the Open University teaching-learning model.

This section calculates the Lambda Mediation Ratio to easily extract knowledge from the dataset.

5.1. Students' evaluation criteria

In this second cycle curricular unit (master's degree course), assessment follows the principles of continuous-summative assessment with collaborative and individual modes. The final grade is calculated based on a balanced 60% of the continuous summative-collaborative work (resolution and discussion of e-activities) and 40% of the final continuous summative-individual work. This ensures a fair evaluation process. The final formula can be presented in the following way:

Final formula = 60% collaborative mode + 40% individual mode

The evaluation model has two phases: collaborative work and individual final work. Each of the three topics (3 × 20% = 60%) lasts three weeks, each with the organization mentioned above. In the phase of greater autonomy, the student carries out individual work in the last topic, weighing 40% of the final grade. The minimum grade for e-activities and final work is 9.5 in any case.

Collaborative mode

On average, the student is expected to spend 10 hours per week (6 credits × 26 hours / 15 weeks) online, using the platform to participate in e-activities, which is why daily attendance is recommended.

In collaborative work, the performance in solving e-activities and the review provided to the e-activities of group colleagues are evaluated to emphasize the importance of student-student interaction. It is expected that the detailed algorithms and calculations will be presented in the summary report for each topic. Descriptive answers must include examples and figures that illustrate the text. In distance assessment, there is a concern about creating an environment with the greatest possible transparency, so in addition to the correction criteria, students have access to quotations and comments for each paragraph of the e-activity.

The topics with three weeks each are organized as follows:

- In the first week, students have access to documentation and must read it carefully; the e-activity is made available, corresponding to a set of exercises; election of the working group Coordinator; the group Coordinator opens new entries for each e-activity question; in the body of the message students must include the text of the question to make it the most readable the forum; students must distribute them among the group members, with the supervision of the group Coordinator.
- In the second week, students must solve them individually and present the results in the separate forum themes. The members of the group discuss the activities.
- Finally, in the third week, each group continues to discuss the exercises solved by their colleagues in each of the open topics on the forum; the group Coordinator organizes the writing of the Group Report and submits it to the Work functionality.

The assessment considers the process of knowledge construction and the final product. The knowledge is constructed in the group discussion forums. The report represents the final product. After the report is evaluated, the students' evaluations are given, considering the contributions in the forum.

The collaborative mode (CM) formulas consider two attributes: (i) the students' interaction, which includes the frequency and quality of the answers given in the forum, and (ii) the average quality of the group reports. The interaction is closely related to the construction of the knowledge process, and the quality of the group report is related to the final product. Two formulas can be used and expressed in the following way:

- Formula CM1= $\alpha \times$ [product] + $(1-\alpha) \times$ [interaction] with $0<\alpha<1$
- Formula CM2 = [product] + $\beta \times$ [interaction] with $-1<\beta<1$

Individual mode

In the individual final work, the student is expected to exercise autonomy in creating models, analyzing them, and evaluating the computational results.

In this mode, the student can choose a dataset from an available set or a public repository. The dataset should be different from his colleagues' choice. The student should formulate and answer his research or business questions using the tools and techniques studied.

5.2. Educational Dataset

The dataset has five columns: StudentId, NewContent, NewAnswers, GroupReport, and FinalMark. NewContent and NewAnswers measure the number of contributions in the forum. NewAnswers represents the number of new posts sent to the forums. NewContent represents the number of times the student updated the posts. GroupReport represents the mark of the aggregated work organized by the group coordinator. The attribute FinalMark gives the outcome of the unit, i.e., the student's final mark. Table 2 shows an extract of the dataset.

Table 2. Extract of the dataset

StudentId	NewContent	NewAnswers	GroupReport	FinalMark
1	102	45	13	13
2	32	32	17	17
3	20	16	17	15
4	28	24	17	17
5	24	24	14	12
6	16	12	14	12
7	102	90	18	18

The dataset results from Moodle logs and students' outcomes of the course unit Data Mining 2023. NewContent and NewAnswers are imported from Moodle, and GroupReport and FinalMark are the outcomes of the professor's evaluation.

5.3. Computational Results

Given the educational dataset with variables (T, M, Y) and the Lambda-ratio procedure for mediation analysis, starting the computational works is now possible.

From the dataset Data Mining 2023, the chosen variables are:
T - NewAnswers quantifies the interaction in the construction of the knowledge process;
M - GroupReport mark quantifies the final product of the group;
Y - FinalMark

After running mediation analysis, each estimator a, b, c, and c' is associated with the p-value, using the following symbols:

- p-value <= 0.001 ***
- p-value <= 0.01 **
- p-value <= 0.05 *
- p-value > 0.05 ns (non-significative)

For the chosen dataset, the computational results for the estimators are the following (see Figure 4):
c = 0.051 **
a = 0.031 *
b = 0.896 ***
c '= 0.023 **

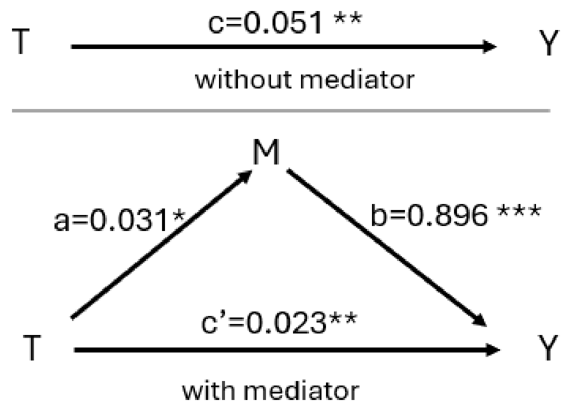

Figure 4. Estimators of the dataset Data Mining 2023

The effects of the mediation are the following:
Indirect Effects (a.b) = 0.028
Direct Effect (c') = 0.023
Total Effects (c) = 0.051
Mediation Ratio = 0.549 (partial mediation)

Setting Lambda equal to 0.20, a Mediation Ratio of 0.549 indicates a Partial Mediation since it is between 0.20 and 0.80.

Partial mediation occurs in mediation analysis when the mediator variable (M) explains only part of the relationship between the independent variable (T) and the dependent variable (Y). This means that while the mediator variable M significantly impacts Y, T still retains a direct effect on Y that is not fully mediated.

In our problem, the mediator GroupReport mark partially impacts the final mark Y, while the interactions during the knowledge construction process still directly affect Y. In summary, the final mark is the result of direct and indirect effects, or in other words, the final mark is the result of the group report quality and the number of interactions with new contributions.

Now, we can answer to the research questions previously defined:

Q1: Does the selection of a group impact an individual's final assessment?
Yes, the Indirect Effects equal to 0.028 show that the quality of the group report impacts the assessment, and consequently, the group's selection is significant. In this approach, we assume that selecting a responsive group impacts the quality of the group report.

Q2: Does the frequency of interactions affect an individual's final assessment?
Yes, the Direct Effects, equal to 0.023, show that the number of interactions impacts the student's final assessment.

Indirect and direct effects have similar weights; consequently, the mediation ratio is around 50%. The process and the product are defined in the previous formulas for the collaborative mode.

These computational results show that this mediation analysis procedure can extract adequate information from the dataset.

6. Conclusion

This work explores the mediation effect on assessment in collaborative learning, focusing on online learning under the Portuguese Open University educational model. Assessment is critical in collaborative learning since it focuses on two main aspects: the process and the product.

The proposed research questions explore how the interaction (process) and group selection (product) impact individual assessment. The study aims to integrate quantitative methods to evaluate collaborative learning's effectiveness better.

The Lambda Mediation Ratio was proposed to enhance mediation analysis. This new approach addresses limitations in the traditional Baron-Kenny method, allowing for quick and flexible mediation categorization. Based on the mediation ratio, the procedure enables the rapid classification of mediation into one of three types (full, partial, or no mediation).

In the experimental study of mediation analysis, the dataset used is the result of Moodle logs and students' outcomes. The outcome of partial mediation indicates that both the interaction and group report quality contribute to the assessment.

As expected with mediation analysis, the data reflects the assessment formulas in collaborative work, validating its adequacy.

Although the proposed method fits the classical estimation of sequential regressions studied by other authors, and the results of the case study seem plausible, it is recommended that the method's performance be evaluated for different sample sizes and real mediation effects. This can be done through a simulation study with many replicates. Resampling methods to estimate significance could improve results on several unknown underlying distributions.

In future work, the Lambda Mediation Ratio should also be applied to other datasets to validate its effectiveness and versatility in different collaborative learning environments and other types of data. Additionally, it will allow for comparisons with other techniques, like Structural Equation Modeling (SEM) [Ballen, Salehi 2021], to enable simultaneous estimation, highlighting any improvements or advantages.

Article presentation: 29 of Febrary of 2024
Approval date: June 27, 2024
Publication date: July 30, 2024

Cavique, L. & Ramos, M.R. (2024). Assessment in Collaborative Learning: a Mediation Analysis Approach. *RED. Revista de Educación a Distancia, 24*(80). http://dx.doi.org/10.6018/red.606551

Financing
This work is partially financed by national funds through FCT – Fundação para a Ciência e a Tecnologia under the project UIDB/00006/2020. DOI: 10.54499/UIDB/00006/2020 (https://doi.org/10.54499/UIDB/00006/2020).

Statement by the authors regarding the use of LLM
This article has not used texts from an LLM (ChatGPT or others) for its writing.

References

Ballen C.J., S. Salehi (2021), Mediation Analysis in Discipline-Based Education Research Using Structural Equation Modeling: Beyond "What Works" to Understand How It Works, and for Whom. J Microbiol Biol Educ. 22:10.1128/jmbe.00108-21. https://doi.org/10.1128/jmbe.00108-21

Baron R.M., D.A. Kenny (1986), The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations, Journal of Personality and Social Psychology, vol. 51 (6), pp. 1173–1182, https://doi.org/10.1037/0022-3514.51.6.117.

Boud, D., Bearman, M. (2022), The assessment challenge of social and collaborative learning in higher education, Educational Philosophy and Theory, vol. 56(5), pp. 459–468. https://doi.org/10.1080/00131857.2022.2114346.

Cavique L. (2023), Visão Integrada do Ensino-Aprendizagem: aplicação do modelo da Universidade Aberta, Revista Lusófona de Educação, v. 60 n. 60.

Hair J.F., Hult G.T., Ringle C.M., Sarstedt M., Danks N.P., Ray S. (2021), Mediation Analysis, Classroom Companion: Business.

Iacobucci D. (2012), Mediation Analysis and Categorical Variables: The Final Frontier, Journal of Consumer Psychology, 22. 10.1016/j.jcps.2012.03.006.

Khan A. et al. (2022), Collaborative Learning Assessment via Information and Communication Technology, RIVF, 16th International Conference on Computing and Communication Technologies, pp. 311-316 Dec., IEEE.

Kratz F., B. Pettinger, M. Gratz (2022), At Which Age is Education the Great Equalizer? A Causal Mediation Analysis of the (In-)Direct Effects of Social Origin over the Life Course, European Sociological Review, vol. 38, pp. 866–881, https://doi.org/10.1093/esr/jcac018.

Lange T., Kim H.W., Sørensen R., Galatius S. (2017), Applied mediation analyses: a review and tutorial, Epidemiology and Health, vol. 39.

Mackinnon D.P., Valente M.J., Gonzalez O. (2019), The Correspondence Between Causal and Traditional Mediation Analysis: the Link Is the Mediator by Treatment Interaction, Prevention Science, vol. 21, pp. 147-157.

Misunas C., E. Soler-Hampejsek, B. Kangwana, N.A. Haberland (2024), Do adolescent girls' education and friendships have independent effects on early pregnancy? Results of a mediation analysis from a longitudinal cohort study in Nairobi, Kenya, SSM - Population Health, vol. 25, 101618, https://doi.org/10.1016/j.ssmph.2024.101618.

Pearl J., D. Mackenzie (2019), The book of why, Penguin Books.

Pearl J., M. Glymour, N.P. Jewell (2016), Causal Inference in Statistics: A Primer, Wiley, ISBN: 978-1-119-18684-7.

R Core Team (2023), R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Riofrío-Calderón G., Ramírez-Montoya, M.-S. (2023), Mediation Models for Online Learning and Perspectives for Open Innovation: Systematic Review of the Literature. International Journal of Emerging Technologies in Learning (iJET), 18(18), pp. 102–120. https://doi.org/10.3991/ijet.v18i18.37721

Sobel M.E. (1982), Asymptotic confidence intervals for indirect effects in structural equation models, In Samuel Leinhardt Ed., Sociological methodology, pp. 290–312, San Francisco: Jossey-Bass.

Valente L. (2016), CO-LAB Guidelines for Assessing Collaborative Learning in the Classroom, Publisher: European Schoolnet, http://www.eun.org.

Zhao X., J.G. Lynch, Q. Chen (2010), Reconsidering Baron and Kenny: Myths and Truths About Mediation Analysis, Journal of Consumer Research vol. 37 (2), pp. 197-206, DOI:10.1086/651257.