Teaching and Learning Faculty Publications and Presentations

College of Education and P-16 Integration

5-12-2023

The Exploratory Sequential Design With Generalized Linear Models for the Learning Strategy Analysis

Zhidong Zhang

Liza Ramos

The Exploratory Sequential Design With Generalized Linear Models for the Learning Strategy Analysis

Zhidong Zhang[1] & Liza Ramos[1]

[1] The College of Education and P-16 Integration, The University of Texas-Rio Grande Valley, Texas, USA

Correspondence: Zhidong Zhang, The College of Education and P-16 Integration, The University of Texas-Rio Grande Valley, Texas, USA. E-mail: zhidong.zhang@utrgv.edu

Abstract

The researchers explored a mixed methods research design through the analysis of a set of interview data. Saldana's model was utilized to complete the opening coding data and thematic coding data. Three themes were recognized to support the cooperative learning models. These included: a) Knowing by Undergone Events and Contexts, b) Implementation and Action, and c) Opinions and Perceptions, which consisted of a Cooperative Learning Model (CLM). The frequencies of the terms and phrases consisted of the evidence variables. A generalized linear model (GeLM) was utilized to examine the relationship between the three themes and Cooperative Learning Model. The GeLM analysis reported the relationships between the CLM and three themes. This was a new practice to integrate Saldana's qualitative data analysis and the GeLM into one mixed methods design.

Keywords: generalized linear model, cooperative learning, Bayesian network, mixed methods design, Saldana's model

1. Introduction

This study was aimed at exploring mixed research methods, by which the authors reported and represented the cooperative learning models. The data were derived from an interview with a group of teachers. The general topic of the interview was cooperative learning. The focus was on how to develop a structured model to represent the characteristic of cooperative learning, and further quantitatively examined the relationships among the components of the structured model of cooperative learning. The study also reported which components significantly supported the structure model.

2. Perspectives

Mixed methods research was viewed as "the third camp" of research in social, behavioral, and educational sciences. A mixed methods study was to integrate quantitative and qualitative approaches as components of the research (Creswell, 2011; Teddlie & Tashakkori, 2009; Greene, 2007).

There is a diversity of different mixed methods and data analysis models, which may be applied to this study. Creswell and Plano Clark (2018) introduced the six mixed methods designs, which were methodological candidates and may be applied to the cooperative learning (Zhang, 2007a, 2007b): a) convergent parallel design, b) explanatory sequential design, c) exploratory sequential design, d) embedded design, e) transformative design, and f) multiphase design (Zhang, 2007a; Zhang, 2007b; Zhang & Zhang, 2020; Zhang, Guanzon, 2022). Among these mixed methods designs, two of them are possibly used in this study: the explanatory sequential design and the exploratory sequential design.

A research design is a framework that makes the research problem alignment with the research questions. In the mixed methods design, a specific research design is basically determined by the research questions. The formulated. research questions require different mixed-methods research designs. Even though we cannot decide exactly a research questions should be responded to by a mixed research framework, some better suggestions to select a mixed methods design is still necessary. Some research questions are very obvious that should be responded to by using a qualitative or a quantitative research design; some research questions can use both qualitative methods design and quantitative methods design (Zhang, 2022).

2.1 The Reasons to Utilize the Mixed Methods Research

First, the selection of the mixed methods research was research problem-driven relevant. Some research problems had characteristics indicating that mixed methods research was a better way to respond to the research problem than any single one of the qualitative methods and quantitative methods. This was also because those educational research issues were complex, and they required the essence to be explored but also the findings required representation in some quantitative formats.

Second, the research questions may influence what research methods were used.   For example, if the research question asked about the correlational relationship between any variables, this was very possible the author should do a correlational study, which was a quantitative aspect of the study; if the research questions gave hints that briefly stated the beliefs of the collaborate learning, this was a qualitative research question.

Third, the purposes of the research study required a combination of quantitative and qualitative approaches. The researchers wanted to examine and measure both subjective and objective aspects. Stated differently, both quantitative and qualitative methods together better expressed their characteristics and represent the essence.

2.2 The Paradigms of the Mixed Methods Research

Any research methods require paradigmatic support because one of the research study's purposes is to explain the world in an epistemological framework. The mixed methods research integrates both quantitative and qualitative paradigms, which sometimes conflict in explaining the research findings. Thus, the researcher can strengthen the explanations of the findings in an appropriate paradigm (Creswell & Plano Clark, 2018).

There are basically two research paradigms applied in mixed methods research: a practice-based research paradigm and a dialectical stance. This is very important for researchers to make a decision to choose between pragmatism and dialectic. Cronenberg (2020) and Denscombe (2008) presented rational arguments for communities of practice and the dialectical stance. Relying on a pragmatic approach focused on the problem investigated and the process driven by the research problem. However, there may be inconsistency and fragmentation in the process of presenting the findings and further reaching conclusions. On the other hand, utilizing dialectical paradigm parley allowed researchers to gain the strengths of both paradigms to engage the data in dialogue repudiating assertions based on the parley between the two paradigms.

2.3 Fundamentals of Mixed Methods Designs

There were several different mixed methods and data analysis models based on quantitative and qualitative considerations. Creswell and Plano Clark (2018) introduced six mixed methods designs, which were applied to the research designs and data collections of the assessment and learning (Zhang, 2007a, 2007b): a) convergent parallel design, b) explanatory sequential design, c) exploratory sequential design, d) embedded design, e) transformative design, and f) multiphase design (Zhang, 2007a; Zhang, 2007b; Zhang, 2022a;).

3. Research Methods

Among these mixed methods designs, the exploratory sequential design was the best candidate to apply to this study. The model consisted of two modules: a qualitative module and a quantitative module. The qualitative module was to code the data into categories and further into the structure in terms of Saldana's model (Miles, Huberman, & Saldana, 2014). The quantitative module was based on the Bayesian network quantitative representation of the data and the analysis with a generalized linear model.

3.1 The Way to Use the Sequential Mixed Methods Designs

In designing mixed-methods research, Creswell and Plano Clark (2018) proposes several aspects that should be examined by the researchers, when they start their research design. a) the sequence of the qualitative and quantitative data collection, b) the priority of the qualitative and quantitative data collection and analysis, and c) the stage of the qualitative and quantitative data will be integrated.

3.2 Design Patterns for Mixed Methods Research

Basically, there are two patterns of mixed methods research in terms of both qualitative a quantitative strands: parallel and sequence. This is an important consideration in designing mixed-methods research and is aligned with the research question, even though there are other factors that are related to the patterns of the design. In this study, we will use a sequential design. The research question decides whether the quantitative or qualitative strand should precede. Sequential arrangements can further be exploratory or explanatory. Exploratory sequential designs are when the qualitative strand precedes the quantitative strand. Explanatory mixed-methods studies, on the other hand, are when the qualitative work is conducted after the quantitative strand. This study

uses exploratory sequential design because qualitative data were collected first and a quantitative model was developed based on the analysis of the qualitative data (Creswell, & Plano Clark, 2018; Zhang, & Ramos, 2023)

4. Data Resources and Evidence

The data were a set of hypothetical data, which was from an interview with a group of teachers. The general topic of the interview was cooperative learning. The focus was on how to develop a structured model to represent the characteristics of cooperative learning. The semantic unit was the sampling unit.

The first step was to code the interview data. Saldana's model (Miles, Huberman, & Saldana, 2014), and JMP software were utilized (2022) to complete the opening coding. As shown in Table 1, the term and phrase lists were developed. Two more coding procedures were themes and theories, which meant from categories to the themes and from the themes to the theories (Miles, Huberman, & Saldana, 2014; Morse, 2008). Morgan (2018) presented that, "themes convert codes into concepts representing important aspects of the results. Models connect themes to show the relationships between the themes and theories. Theories explain why these theories capture the data and are related in the ways the models show" (p. 339).

Table 1. The part of the outputs of the interview data: Term and phase lists

Term and Phrase Lists

Term	Count		Phrase	Count
science	61		th grade	23
students	39		5 th grade	22
grade	35		5 th	22
learning	30		cooperative learning	19
cooperative	26		grade teacher	8
th	25		science standards	8
5	23		5 th grade teacher	7
training	23		th grade teacher	7
years	21		professional development	7
teacher	18		cooperative groups	6
teaching	18		experience trainings	6
groups	17		grade level	6
standards	15		using cooperative	6
learn	14		science lab	5
teachers	14		science training	5
trainings	14		small groups	5
using	14		teach science	5

The second step was to develop a cooperative learning model based on the themes and thematic relations of these categories. Thematic coding suggested three themes that were Knowing by Undergone Events and Contexts (KUEC), Implementation and Action (I&A), and Opinions and Perceptions (OP). There are two latent variables belonging to Implementation and Action (I&A), and Opinions and Perceptions (OP). There were twenty-five evidence variables developed and selected from the term and phrase list. These evidence variables supported the latent variables.

These three themes plus Cooperative Learning consisted of the top of the Cooperative Learning Model. The selected terms and phrases from the open coding consisted of evidence variables and six lower-level latent variables to support the themes with evidence. As shown in Figure 1, the Cooperative Learning Model was represented in a Bayesian network (Conrady & Jouffe, 2022), which highlighted the quantitative characteristics of the data. Stated differently, the variables in this model were represented in probabilistic values dynamically format.
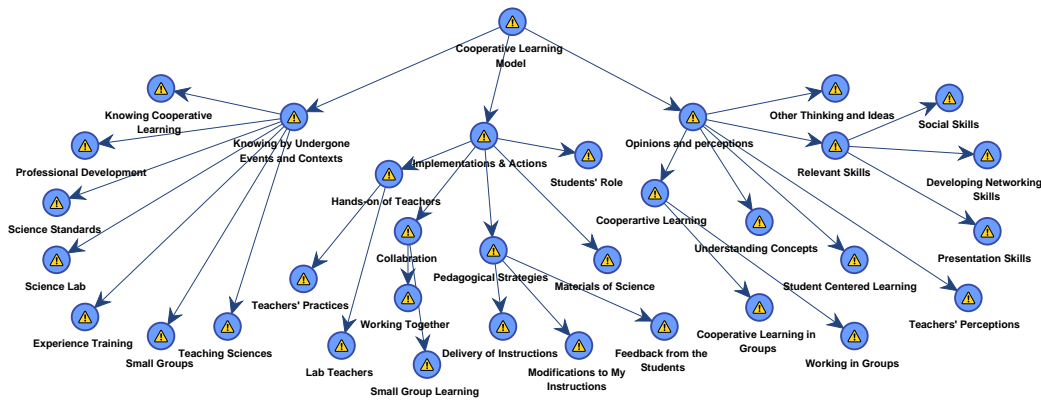
Figure 1. Cooperative learning model with three thematic variables

With the increase of frequencies of the semantic units, the values of the top four latent variables increased. In order to simplify the probabilistic variable, the values of the probabilistic variables were determined in dichotomous categories: the lower value was labeled as 0 and the higher value was labeled as 1. Thus the combination of these four variable values was $2^4 = 16$. Further, the probabilistic values of these four variables were rescaled to the frequencies. This data structure was suitable for the analysis with a generalized linear model.

A generalized linear model was utilized to examine the research concerns in the quantitative aspect (Agresti, 2002; Stroup, 2012). The generalized linear model consisted of more than one explanatory variable and one categorical dependent variable. Stated differently, the dependent variable was not a continuous variable. This model was widely used to analyze the relationships among several categorical variables (Azen, & Walker, 2021). As shown in Table 2, significant results were reported. The significance means the alpha value is less than 0.05 in column Sig. Here, the "*" sign indicated the interaction between two variables. These variables and variable interactions are Implementation and Action, Opinions and Perceptions, Cooperative Learning Model, Knowing by Undergone Events and Contexts * Cooperative Learning Model, Implementation and Action * Opinions and Perceptions, Implementation and Action * Cooperative Learning Model, Opinions, and Perceptions * Cooperative Learning Model.

Implementation and Action, and Opinions and Perceptions, these two variables were significant in the Cooperative Learning Model, but Knowing by Undergone Events and Contexts was not significant, which meant that this variable was not important for the Cooperative Learning Model. However, the interaction between Knowing by Undergone Events and Contexts, and Cooperative Learning Model was significant. This indicated that if looking at the relationship between Knowing by Undergone Events and Contexts, and Cooperative Learning Model separately, the Knowing by Undergone Events and Contexts has potentially increased weight.

Table 2. Generalized linear model tests of model effects

Source	Type III		
	Wald Chi-Square	df	Sig.
(Intercept)	74870.022	1	.000
Knowing by Undergone Events and Contexts	.334	1	.564
Implementation and Action	32.214	1	.000
Opinions and Perceptions	33.195	1	.000
Cooperative Learning Model	206.627	1	.000

Knowing by Undergone Events and Contexts * Implementation and Action	.030	1	.863
Knowing by Undergone Events and Contexts * Opinions and Perceptions	.074	1	.785
Knowing by Undergone Events and Contexts * Cooperative Learning Model	24.841	1	.000
Implementation and Action * Opinions and Perceptions	120.784	1	.000
Implementation and Action * Cooperative Learning Model	530.633	1	.000
Opinions and Perceptions * Cooperative Learning Model	520.251	1	.000
Knowing by Undergone Events and Contexts * Implementation and Action * Opinions and Perceptions	1.124	1	.289
Knowing by Undergone Events and Contexts * Implementation and Action * Cooperative Learning Model	.516	1	.473
Knowing by Undergone Events and Contexts * Opinions and Perceptions * Cooperative Learning Model	8.244	1	.004
Implementation and Action * Opinions and Perceptions * Cooperative Learning Model	.117	1	.733

5. Conclusions

This study examined the patterns of the interview topics on the cooperative learning of science education. Saldana's model (Miles, Huberman, & Saldana, 2014), and JMP software were utilized (2022) to complete the opening coding and thematic coding. Subsequently, three themes were recognized to support the cooperative learning models. These three themes were Knowing by Undergone Events and Contexts (KUEC), Implementation and Action (I&A), and Opinions and Perceptions (OP), which consisted of a Cooperative Learning Model. These three themes are latent variables, which cannot be directly observed, and they were evidentially supported by twenty-five evidence variables. These evidence variables were developed from the term and phrase lists. Thus, a four-layer tree model was developed for both description and evaluation purposes. A Bayesian network was utilized to transform the data from a qualitative representation to a quantitative representation. The frequencies of the terms and phrases in the interview data consisted of the data of evidence variables. These variables propagate the evidence information up to the latent variables. Thus the four upper-level latent variables received updating probabilistic values, which were further rescaled to frequencies. Subsequently, a generalized linear model was applied to the data analysis. The results indicated that these two variables, Implementation and Action, and Opinions and Perceptions, significantly support the top latent variable, the Cooperative Learning Model.

6. Scholarly Significance of the Study

This study explored a mixed methods research design utilized in the analysis and representation of a set of hypothetical interview data on the topic of cooperative learning. The unit of observation was the terms or phrases of the interview texts. Saldana's coding schemata allowed the researchers to code the data from phrases to categories, themes, and further to the cooperative learning model, which were further represented in both qualitative features and quantitative features.

Bayesian network representation of the cooperative learning model provided a tool for the researchers to further quantitatively examine relationships among these latent variables: Knowing by Undergone Events and Contexts (KUEC), Implementation and Action (I&A), and Opinions and Perceptions (OP). The probabilistic variables were further rescaled into the frequencies. A generalized linear model was utilized to examine the relationship between three latent variables and the cooperative learning model. Briefly, the Bayesian network, as a "pivot", transformed the data from the qualitative paradigm to the quantitative paradigm. Further, the analysis of the generalized linear model reported the relationship between the cooperative learning model and three latent variables. This was a new attempt to integrate Saldana's qualitative data analysis and generalized linear model into one mixed methods design.

7. Limitations

This study used data from the hypothetical interview text. Even though the mixed methods and Bayesian network model effectively transform the data from a qualitative strand to a quantitative strand the findings and analyses have limited generalizations. The authors will collect more data from more participants. The generalized linear model aid to examine what elements are significant in terms of the linear model assumption. However, it cannot inform the distribution of a hierarchical Bayesian network model.   The distribution of the Bayesian network model should be examined in a different way. In addition, there are other ways to do qualitative data analysis, which may produce different Bayesian network models.

References

Agresti, A. (2002). *Categorical data analysis.* Wiley. https://doi.org/10.1002/0471249688

Azen, R., & Walker, C. M. (2021). *Categorical data analysis for the behavioral and social science.* Routledge. https://doi.org/10.4324/9780429330308

Conrady, S., & Jouffe, L. (2022). *Bayesian networks and BayesiaLab: A practical introduction for researchers.* Retrieved from https://library.bayesia.com/articles/#!bayesialab-knowledge-hub/book

Creswell, J. W. (2011). Controversies in mixed methods research. In N. K. Denzin, & Y. S. Lincoln (Eds.), *The SAGE handbook of qualitative research in social & behavioral research* (pp. 45-68). Sage. https://doi.org/10.4135/9781506335193.n2

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conduction mixed methods research*. Sage.

Cronenberg, S. (2020). Paradigm parley: A framework for the dialectical stance. *Journal of Mixed Methods Research, 14*(1), 26-46. https://doi.org/10.1177/1558689818777925

Denscombe, M. (2008). Communities of practice: A research paradigm for the mixed methods approach. *Journal of Mixed Methods Research, 2*(3), 1-25. https://doi.org/ 10.1177/1558689808316807

Greene, J. C. (2007). *Mixed methods in social inquiry*. Jossey-Bass.

JMP. (2022). *Statistical software for insight-driven improvement*. JMP Statistical Discovery. Retrieved July 22, from https://www.jmp.com/en_us/home.html

Miles, M. B., Huberman, M. A., & Saldana, J. (2014). *Qualitative data analysis: A methods sourcebook.* SAGE.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods.* Sage.

Stroup, W. W. (2012). *Generalized linear mixed models: Modern concepts, methods and applications*. CRC Press.

Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences.* Sage.

Zhang, Z. (2007a, Jan.). *Application of a Bayesian belief network to diagnostic cognitive assessment in studies in statistics learning and performance.* Presented at seminar, Faculty of Education, McGill University, Montreal, QC.

Zhang, Z. (2007b). *Cognitive assessment in a computer-based coaching environment in higher education: Diagnostic assessment of development of knowledge and problem solving skill in statistics*, Unpublished doctoral dissertation, McGill University, Montreal, QC, Canada.

Zhang, Z. (2022a). A mixed methods design for assessing physics learning in the online learning environment. *Journal of Education and Development, 6*(2), 1-8. https://doi.org/10.20849/jed.v6i2.1142

Zhang, Z. (2022b). Diagnostically cognitive assessment of complex semantic explanations and learning: A Bayesian network approach. *Journal of Education and Development, 6*(4), 1-14. https://doi.org/10.20849/jed.v6i4.1255

Zhang, Z., & Guanzon, A. (2022). A mixed assessment for the science learning via a Bayesian network representation. *Journal of Education and Development, 6*(5), 1-7. https://doi.org/10.20849/jed.v6i5.1309

Zhang, Z., & Ramos, L. (2023). Exploring a cooperative learning model with a mixed methods design. *Journal of Education and Development, 7*(1), 76-79. https://doi.org/10.20849/jed.v7i1.1328

Zhang, Z., & Zhang, Y. (2020). Alternatively assessing conceptual learning in an emergency clinical environment: A mixed methods design. *Journal of Education and Development*, *4*(2), 27-37. https://doi.org/10.20849/jed.v4i2.744

Copyrights