Article Critique Portfolio

EDU647 Statistical Thinking and Applications

Wenjie Shen

Syracuse University
Content

Article Critique 1 ANOVA & MANOVA ................................................................. 2

Article Critique 2 Correlation Analysis & Regression ........................................ 8

Article Critique 3 Chi Square & Regression ...................................................... 13

Reference .............................................................................................................. 21
Article Critique 1  ANOVA&MANOVA


Purpose

The purpose of the study is to explore the ways in which pre-professional health sciences students collaboratively accomplished tasks offered in a computer-supported collaborative learning (CSCL) environment and to identify behaviors that led to group regulation and/or socio-emotional interactions. The study also aimed to classify desirable collaboration patterns which were demonstrated by students.

Research Design & Analysis

Due to the nature of research purpose, an ex post facto research design was employed to identify students’ online collaboration behaviors occurred naturally without an experimental treatment. Three causal comparative groups were formed by 6 student groups according to their group collaboration patterns. The data was collected from the online discussion board throughout the study period and the students’ final grades in the end.

Theoretical framework

Drawing from existing literature, the authors stated that group regulation in a CSCL context usually spent more time and effort than in a face-to-face setting. They identified two critical components of computer supported collaborative learning:
group regulation and social-emotional interactions and gave rationale to the research questions “What kinds of group regulation and social-emotional behaviors are involved in during the collaboration process?” and “what types of patterns evolved while the project progressed?”. 

**Sampling methods and generalizability**

The sampling of this study is relatively weak. The study utilized a convenient sample of 28 students (N=28), with 25 being female and 3 being male, enrolled in an undergraduate clinical ethics online course offered at Midwestern land grant university. Its limited generalizability as well as low external validity was influenced by small sample size, convenient sampling and inequivalent gender ratio.

**Procedures and measurements**

The participants were randomly assigned to 6 groups that consisted of 5 to 6 members, with the requirements of developing a clinical ethics scenario in this semester-long course. The group project consisted of 8 Milestones (M1 through M8) divided into different tasks. Messages posted on a discussion board were collected and analyzed based on the coding scheme developed by researchers in a recursive process. Then a cluster analysis was employed to explore if there were different patterns of collaborative behaviors among groups. Based on a post hoc analysis of group discussion activities in terms of timing and types of behaviors (group regulation and socio-emotional), three groups were categorized (IVs): Early Active Collaborator (EAC), Late Collaborator (LC) and Task-oriented Collaborator (PTC).

Students’ perception on their group process and group productivity were
measured as **Dependent Variables (DVs)** to know if the collaboration was successful. Two independent surveys were conducted to observe students’ perceptions of group process at different times. The first was conducted the end of M2 to investigate group regulation in early collaboration phase (Strategy domain: Cronbach’s $\alpha=.51$; Communication domain: Cronbach’s $\alpha=.61$; Authorship domain: Cronbach’s $\alpha=.80$), while the second was conducted at the end of M8 to investigate students’ feelings on the group process. Products of each Milestone graded by the instructor with clear criteria.

**Data analysis and conclusion**

The statistically analysis was carried out first with two independent **one-way ANOVA** on the total mean number of discussion units per category (group regulation and social-emotional behavior) to reveal significant difference between groups (group regulation: $F (5,22) = 2.7, p<.05$, partial $\eta^2 = 0.24$; social-emotional: $F (5,22) = 2.79, p <.05$, partial $\eta^2 = 0.39$). Then, a **cluster analysis** was employed to identify different collaborative patterns and **three individual MANOVAs** were employed on M2, M4 and M8, to reveal the difference between three clusters (IVs for MANOVA) on 8 subcategories of groups regulation and on three subcategories of socio-emotional behaviors (DV for MANOVA), which led to significant differences between clusters in all three Milestones. After that, students’ perception on their collaboration was examined by **one-way ANOVAs** (clusters as IVs and sub-categories of perceptions as DVs). ANOVAs on Strategy and Communication revealed significant difference between clusters (Strategy: $F (2,25) = 5.36, p <.05$, partial $\eta^2 = 0.30$; Communication:
ANOVA on “positive group climate” revealed significant difference, too (F (2,22) -3.92, p<.05, partial $\eta^2 = 0.26$). Statistical analysis on group productivity (instructors’ grades) was not conducted because the values of EAC violated normality and equality of variances with other groups, but the authors gave description and explanation to it still.

The main findings of the study could be summarized as follows: In an online group collaboration setting where minimal instructor’s direct intervention was given, 3 different collaboration patterns could be identified according to collaborative action types and timing — EAC, PTC and LC. The most desirable collaboration pattern was EAC, demonstrated by one group that had intensive interactions among group members in the early collaboration phase, continuous socio-emotional interactions and gradually scaling down group coordination activities. Other two patterns were PTC with dormant interactions throughout the projects and least socio-emotional interactions, and LC which rushed at the end of the project without a clear group coordination plan. The discussion also described how social-emotional actions and group regulations were observed during the project and contributed to the group success. The results also led to some practical suggestions for instructors.

**Strength & Limitations**

The conceptual framework was good because it clearly identified two critical components of CSCL: group regulation and social-emotional interaction, which made the research questions more explicit. Furthermore, group regulation behaviors were interconnected with social-emotional interactions, so the authors tried to identify
different types of patterns (how the behaviors change during the group project) for collaboration and explore which type would be desirable, which attributes to the **construct validity** of the study.

In addition, the study exerted a strong **internal validity** because of random assignment to groups, 2 independent coders coding according to scheme, clear and consistent criteria of grading and excellent internal consistency (Cronbach’s $\alpha=.91$) of the second survey. However, the low reliability of the first survey might be one threat to internal validity.

What’s more, the study exhibited a medium-strong **statistical conclusion validity** with multiple methodologies. Data analysis was carefully done and reported in details. Although one-way ANOVAs were not conducted on each subscale, exploratory factor analysis (EFA) and Tukey post-hoc test was utilized to give directions on which factors should be identified specially.

Nevertheless, the study has **limited generalizability** as well as **low external validity**, which was influenced by small sample size, convenient sampling and inequivalent gender ratio.

The authors discussed following limitations of their study, which demonstrated researchers’ integrity and understanding of their study:

(a) the small sample size of participants;

(b) collaboration setting limited to asynchronous communication;

(c) low reliability of the first survey.

Other than that, the authors also listed several **strengths**:
(a) large amount of data of ideas units;

(b) multiple measurements with triangulated data sources;

(c) the long term of the study which was “almost ethnographic approach”.

As far as I am concerned, two more limitations could be added to the discussion part. First, most of the participants were female and majored in health care, thus they could be more agreeable, collaborative, positive in group working. Second, although the course was divided into 8 Milestones, only 3 Milestones were collaborative tasks, which means the group work between participants might be very simple and not deep.

**Suggestions of Improvement**

The first suggestion to improve the study would be increasing the sample size. Sample size is a critical warrant for external validity and statistical conclusion validity. The sample should include equivalent ratio of male and female participants. Second, the study said that the groups were randomly assigned to achieve heterogeneous grouping, which needs more explanation. If the instructor wanted to achieve heterogeneous grouping, he or she might select some participants into groups instead of random assignments. Third, the first survey designed by the researchers should be validated prior to the implication, in order to receive possible modifications to achieve higher reliability and validity.

In all, given the detailed context and theoretical background, the study really exerted a good design and a strong statistical analysis, despite of its limitations. Furthermore, it contributed a new perspective of behavioral patterns in studies of social-emotional regulation.
Article Critique 2  Correlation Analysis & Regression


**Purpose**

The purpose of the study was to investigate how Self-Regulated Learning (SRL) explains students’ learning experiences in an online remedial mathematics course. Defining the subcomponents of SRL as motivation, emotion and learning strategies, the authors stated the research questions as follows

1. To what extent do the subcomponents of SRL predict student achievement?
2. To what extent do the subcomponents of SRL predict student online course satisfaction?
3. Do any differences exist in the subcomponents of SRL in passing and non-passing students?

**Research Design & Analysis**

The research exerted a *non-experimental* design, contextualized in a self-paced online mathematical remedial course for enrolling in college-level mathematics course. Voluntary online survey was conducted once in the middle of the course semester with neither incentive for participation nor penalty for nonparticipation. Self-regulation is critical in determining students’ success in online learning environments, but little empirical research has been conducted in remedial online mathematics courses regarding to SRL, which provided rationale for the study.
Theoretical framework

Based on literature review, the authors defined SRL with three components: motivation, emotion and learning strategies (IVs). Previous Research showed that SRL could promote successful learning experiences in online learning, thus the authors decided to examine students’ online learning experiences in terms of three aspects of learning outcomes: achievement, course satisfaction and passing or non-passing (DVs). The literature review part gave a clear theoretical framework for the study, explaining why three IVs and three DVs were identified in this study with strong evidence, which contributed to a strong construct validity.

Sampling and generalizability

Although the study didn’t mention specifically, a convenient sampling was supposed. A total of 229 students enrolled in the course participated in the study: 83 males and 146 females. The sample size was quite large (N=229), while sample characteristics were less various with 79.9% Caucasian and 74.2% freshmen, which might decrease the external validity. This course is prior to college-level mathematics course, which explained why the majority were freshmen.

Procedures and measurements

The online survey was given in the middle of the course semester and several instruments were used to measure the subcomponents of SRL. Instruments used in this study was constructed by the authors with the combination of the motivated strategies for learning questionnaire (MSLQ) (Pintrich, Smith, Garcia, & Mckeachie, 1993), the achievement emotions questionnaire (Pekrun, Goetz, & Perry, 2005) and
the Artino’s (Artino, 2009) research. Cronbach’s € for each group of items were examined and wording changes were made to adapt to the study context.

For measurements of IVs, motivation included intrinsic goal orientation(N=4), task value(N=6) and self-efficacy(N=8) and all the items of motivation came from the MSLQ. Emotions included test anxiety(N=4), boredom(N=5) and frustration(N=4), with items on test anxiety derived from the MSLQ and others adapted from the achievement emotions questionnaire. Learning Strategies included metacognitive self-regulation(N=12) and critical thinking strategies(N=5), with items driven all from the MSLQ but slightly changed to best represent an online learning environment.

For measurements of DVs, satisfaction was measured with three items adapted from Artino’s research with wording changed to accommodate the online system. Students’ achievement was measured with final grades of the online course.

**Data analysis and conclusion**

After having collected data, two analysis of Pearson correlations were examined to reveal the (a) correlations between subcomponents of SRL and students’ final grades; (b) correlations between subcomponents of SRL and students’ satisfactions. Results showed that final grades were negatively related with negative emotions (ex. Test anxiety ($r=-.24, p<.01$); boredom ($r=-.16, p<.05$)), but had no relations with SRL subcomponents. One the other hand, students’ satisfaction with the online mathematics course was positively related with motivation (ex. Intrinsic goal orientation ($r=.47, p<.01$); task value ($r=.62, p<.01$)), while negatively related with negative emotions (ex. Test anxiety ($r=-.45, p<.01$)).
Then, **two three-step hierarchical regression analysis** was conducted to investigate the (a) relationships between final grades and the subcomponents of SRL (motivation in step 1, emotion in step 2, learning strategies in step 3); (b) relationships between course satisfaction and the subcomponents of SRL. The overall model with the subcomponents of SRL explained 11.9% of the variance in final grades \( F(8, 220) =3.71, p<.01, \) and 63.1% of the variance in satisfaction \( F(8, 102) =26.05, p<.01. \)

Finally, **independent two-samples t tests** were conducted to compare the subcomponents of SRL in passing and non-passing students. Results showed a significant difference between those students who passed and those who didn’t in terms of task value \( t(227) =2.56, p<.05), \) self-efficacy\( t(227) =3.57, p<.001), \) test anxiety\( t(227) =-2.99, p<.01), \) boredom\( t(227) =-2.64, p<.01), \) and frustration\( t(227) =-4.11, p<.001). \)

The conclusion draw by the authors was that motivational and emotional variables significantly predicted students’ achievement and satisfaction. In particular, self-efficacy for learning significantly contributed to achievement, whereas emotion explained satisfaction mainly. However, no significant difference in the use of learning strategies by passing and non-passing students was found. In addition, learning strategies didn’t contribute to explaining achievement and satisfaction in this study. The authors compared their results with the previous studies and explained that the inconsistency might due to the linear way of the course materials presented by the ALEKS system. Based on the conclusions, the authors recommended instructors to enhance students’ self-efficacy, design supporting tools in online courseware and
provide SRL support through social media.

**Strength & Limitations**

The authors analyzed the limitations as well as the significance of the study, which showed their comprehensive understanding of it. Regarding the **limitations**,

(a) The survey was administered in the middle of the course, so it couldn’t test the satisfaction changing as the course proceeded.

(b) Only negative emotional variables were used to identify the role of emotion in this study because the authors considered them to be most related to the study context.

Besides those, I would like to add one more limitation. Although the study context was a self-paced online mathematical course, the learning materials were given in a linear way, which means that students were learning according to a given structure and they didn’t need much of self-regulation. The SRL strategies were given a low ceiling in this study, which reduced the sensibility in the examination of SRL strategies. Usually online learning requires abundant ability of SRL to get high achievements. Therefore, the course choice (probably a convenient choice) decreased the generalizability of the study.

Nevertheless, the authors also underlined the **significance** of the study from the perspective of research topic and findings,

(a) Few of previous studies included emotion as one of the subcomponents of SRL, so the current research contributed to this blank space.

(b) Learning experiences were investigated from multiple angels: satisfaction,
achievement and passing or non-passing.

I would like to emphasize its strength in research design, too. Generally speaking, the research questions and design was based on explicit explanation of SRL theories and through review of previous studies, which contributed to a strong construct validity. The measurement of IVs and DVs of this study had a clear structure and strong reliability. Items were all adapted from widely used questionnaires with Cronbach’s $\alpha$ reported. All these contributed to a relatively high internal validity. The study exerted a strong data analysis with details in forms and clear statements, which contributed to a relatively high statistical conclusion validity. The role of all the subcomponents of SRL played in students learning experiences, in practical the final grads, satisfaction and passing or not, was investigated separately with Pearson correlations and t-test, and re-examined with regression analysis to make the conclusion more convincing.

Suggestions of Improvement

Based on the limitations discussed above, several suggestions could be driven:

(a) Choose an online course which requires more self-regulation and self-direction if possible, so that the role SRL learning strategies played in students’ learning experiences could have a higher ceiling.

(b) Conduct the survey twice, one at the middle and one at the end, because variables like students’ satisfaction or task anxiety, may change during the course period.

**Purpose**

The purpose of the study is to investigate the relationship between learning approaches and academic achievement in full-time and part-time Hong Kong Chinese students, examining (a) whether deep learning approach can predict academic achievement for full-time and part-time students, and (b) whether mature students are likely to adopt the deep approach in a sub-degree program. Based on the purpose and previous research findings, the following hypotheses were put forward:

1. There will be a positive relationship between deep learning approach and academic performance for both full-time and part-time students.
2. Mature students will be more likely to adopt the deep approach in sub-degree programs than younger students.

**Research Design & Analysis**

*Theoretical framework*

The theoretical framework was based on Biggs’s Presage-Process-Product (3P) model (Biggs, 1993) which proposed that two learning approaches — surface and deep, generated a strong impact on academic performance. The author reviewed both theoretical studies and empirical studies on relationship between learning approaches and academic performance, as well as relationship between learning approaches and age. The theoretical framework gave abundant rational to this study, and the
hypotheses were based on previous studies review, which contributed to the **construct validity** of this research.

**Sampling and generalizability**

Two samples were collected for this study: one from the full-time students ($N_1=130$) of business diploma program and the other from part-time students ($N_2=131$) of the same program (When going to data analysis part, I’ve found that the sample sizes of two groups were switched. This issue will be criticized in the limitation part). Students were surveyed in class on a voluntary basis, using a **non-probabilistic sampling** procedure, with a return rate of 90.31%. Although the sample size was not small, it still turned out to be a **convenient sampling**. Besides, there was no demographic information of participants.

**Procedures and measurements**

The study employed Biggs’s Revised Two-Factor Study Process Questionnaire (R-SPQ-2F) to measure learning approach (**IV & DV in different parts of the data analysis**). It’s a widely used instrument, but the author didn’t provide the evidence of its reliability. The author only mentioned that “**confirmatory factor analysis showed a good fit with R-SPQ-2F by using Hong Kong tertiary students**”, which adds a little reliability of using this instrument in this study, because Hong Kong students were given the survey. Grade Point Average (GPA) was used to examine academic achievement (**DV**). Since the GPA was measured on a scale from A to F, the author converted the grades to numbers with A=5, B=4, C=3, D=2, F=1. Using GAP to represent academic achievement was a common and convenient way, but its limitation
was also obvious: (a) GPA was calculated with a quite restricted interval which
limited its ability of testing differences of academic achievement between students; (b)
GPA was only one part of academic achievement and relied mainly on summative
assessment.

As a causal-comparative research, the data was measured with one-time survey
and the final GPA. The research could be divided into two parts. The first part was to
test hypothesis I, therefore the relationship between deep learning approach (IV
assigned via the result of the survey) and academic achievement (DV) was
examined separately in full-time student group and part-time student group. The
second part was to test hypothesis II, thus the learning approach (DV) of two groups
(IV assigned naturally according to full-time or part-time) was compared with
chi-square tests and regression analysis.

Data analysis and conclusion

Hypothesis I: There will be a positive relationship between deep learning
approach and academic performance for both full-time and part-time students.

For the first hypothesis, chi-square tests followed by regression analysis were
carried out to two sample groups separately, with deep or surface learning approach as
IV and academic achievement as DV. The students were classified into two groups
according to GPA: low academic achiever with grade C, D and F; and high academic
achiever with grad A and B. For full-time students, the results showed that high
academic achievers are significantly higher than low academic achievers on deep
leaning approach, \( \chi^2 (1, N_1=131)=13.418, p<.001 \). And academic performance was
significantly predicted by learning approaches, which accounted for 10% of the
variances of the respective academic achievement ($R^2=0.102$, $F (1,129) =16.06,$ $p<.001$). And for part-time students, high academic achievers were significantly
different from low academic achievers in learning approaches, $\chi^2 (1, N_2 = 130)=
24.960, p < .001$. And their academic performance was significantly predicted by
learning approaches which accounted for 19% of the variance of the respective
academic achievement ($R^2=0.192$, $F (1,128) =49.71$, $p<.001$).

_Hypothesis II: Mature students will be more likely to adopt the deep approach in sub-degree programs than younger students._

For the second hypothesis, _chi-square_ tests were conducted with learning
approach as _DV_ and two students’ groups as _IV_. The result showed no statistically
significant difference between full-time and part-time students in relation to their
learning approach ($\chi^2 (1, N = 261) = 0.292, p = 0.589$).

The study drew a conclusion that there was a modest link between learning
approached and academic achievement and no significant age difference in learning
approaches in this study. The author discussed the possible explanations for the
conclusion with review of relative studies, and pointed out that firmer predictive
relationships need to be investigated through longitudinal research, increased sample
size and more refined instruments.

**Strength & Limitations**

The major strength of the study was the consistency from its theoretical
framework to the hypotheses, and then to the measurement of variables and data
analysis, which helps the audience understand the whole research design. The theoretical framework gave abundant rational to this study, and the hypotheses were based on previous studies review, which contributes to the construct validity of this research. It could be better if the author would have provided a brief review of the theoretical background of the 3P model, instead of giving it directly and went to review relative studies immediately.

However, there are still several limitations need to be revealed:

(a) The study exerts a medium external validity. Although the sample size wasn’t small, there was no demographic information of the participants. If full time students included all the students of four years of university, the freshman and senior might not use the same learning method. Let alone the influence from gender or race ratio. The author only mentioned that the average age of part-time students was older, without providing accurate data.

(b) The internal validity of the study was also limited because the two groups were not randomly assigned, nor did the author explained how he or she ruled out alternative explanations for students’ learning approaches or academic achievement. For example, whether they received similar education in the university, whether they had similar previous educational experiences, etc. The majority of part-time students in China are male, while for full-time student the male to female ratio is about 1:1, which could also influence the difference between two groups.

(c) Although regression analysis was utilized after chi square test to test the
magnitude and to add evidence of the relationship, there were two major issues that reduced its **statistical conclusion validity**: First, there was a lack of consistency of the sample size of two groups. In the sampling part it was written that sample size of full-time students was 130 while that for the part-time students was 131. When it came to the data analysis, all the tables and data showed that there were 131 full-time students and 130 part-time students participated in the research. Second, at the beginning the author mentioned that GPA was converted to scores from 5 to 1, while in data analysis GPA was simply divided into two levels. This could be a paper writing problem, but I have the rationale to be skeptical that the author tried to manage the data in a simpler way to conduct analysis.

**Suggestions of Improvement**

According to the limitations discussed above, several suggestions of improvement could be made:

(a) Enlarge the sample size and provide demographic information of participants.

(b) Use more ways to measure academic achievement besides GPA.

(c) When exploring the relationship between age and learning approach, use full-time students of different years in the same program, which could exclude more outside influences.

(d) When exploring relationship between learning approach & academic achievement, enlarge the interval of academic achievement. For example, the study can be designed as a correlational research with GPA measured from 1
to 5, and learning approach measured with scores (higher the score, deeper the approach).

(e) Try to further explore whether the age is a mediate variable between learning methods and academic achievement with multi-regression analysis.
Reference


