International Journal of Data Science and Big Data Analytics

Publisher's Home Page: https://www.svedbergopen.com/

SvedbergOpen
DISSEMINATION OF KNOWLEDGE

Research Paper   **Open Access**

Using ordinal data to assess distance learning

Matthew Norris[1*]

[1]Georgia Institute of Technology, North Ave NW, Atlanta, GA 30332, United States. E-mail: mdnorris@gatech.edu

Abstract

While academic research into educational technology has established flipped classrooms and formative assessment as effective and superior to traditional instruction and assessment, implementation of these new methods has been slow and piecemeal. The end of the 2020 spring semester saw nearly every classroom move to distance learning due to Covid-19. This distance learning required instructors to adopt fundamental features of flipped classrooms and formative assessment and provided an opportunity to study the significance of individual elements of these learning methods and their statistical relationship to one another. Likert Scale-style surveys were completed by high school students who were taking a mathematics class via distance learning. The Likert responses were treated as ordinal data, not continuous, and assessed on the presence or lack of stochastic dominance amongst the survey responses by the Kruskal-Wallis test and post hoc testing. This targeted the efficacy of individual techniques within formative assessment and flipped classrooms. The techniques that demonstrated significance individually were then examined with correlation and clustering analyses to determine if relationships existed and if those relationships were consistent with the literature on educational technology. The research methods used were tailored to fit the unique nature of the collected data and designed to assess whether core philosophies of flipped classrooms and formative assessment could be adopted individually, as wholesale adoption has proven to be quite difficult.

Keywords: *Ordinal data, flipped class rooms, Distance learning, Formative assessment*

1. Introduction

By the end of March 2020, every public school in the United States had shut its doors. From university professors to kindergarten teachers, every instructor had to now deliver educational content over the internet, and every student (and their family) needed to adapt to distance learning as the new status quo. Though K-12 education will eventually resume regular in-person learning, it is clear that schools must, at the least, plan for online education as an emergency necessity.

Academics in education and EdTech research have long recommended the flipped classroom model. A related and widely agreed-upon method is formative assessment, a classroom method primarily hailed as a better alternative to traditional summative assessment. Despite slow adoption or even opposition, both methods would have worked together well during April and May of 2020. It also provided many students with experiences that, while not fully adopting formative assessment or flipped classrooms contained some components of both. The project's purpose was to create a survey that identified student attitudes about elements of flipped classrooms or formative assessments and determine if they had a statistically significant effect on outcomes. In actionable terms, it could isolate the tolerable or even enjoyable aspects of distance learning and bring those into the classroom, even if the classroom never wholly flips.

* *Corresponding author: Matthew Norris, Georgia Institute of Technology, North Ave NW, Atlanta, GA 30332, United States. E-mail: mdnorris@gatech.edu*

2. Literature review

While there seems to be a consensus that flipped classrooms and formative assessment is best for students and educators, there is still active research on both topics. As online learning is experiencing growth in academic and business sectors, research to reduce the attrition rate and improve online education's capabilities is considered incredibly important. In comparison, the concept of individual learning styles demonstrates little to no effect on educational outcomes. A question referring to learning styles was included in the survey, and it will be briefly discussed to demonstrate the methods of analysis for this project.

2.1. Flipped classrooms

A flipped or inverted classroom model is when "activities traditionally conducted in the classroom become home activities" (Akçayir and Akçayir, 2018), which is a very appropriate definition in the distance learning era. In more detail, the model asks students to engage with lectures and other educational media outside of the classroom, and time spent inside the classroom is utilized solving the roadblocks students have encountered during self-education. Research suggests that positive student interaction and reduced anxiety are demonstrated by implementing a flipped classroom (Akçayir and Akçayir, 2018).

Formative assessment is a crucial component of flipped classrooms' success in improving academic performance (Zainuddin and Halili, 2016). The two would likely have to be implemented concurrently, as research suggests that current testing methods would prevent flipped classroom adoption from being entirely successful (Rotellar and Cain, 2016). Rotellar also discussed the high cost of time to faculty while developing the new classroom environment and that many students initially resist the transition. However, even the most driven students (in this experiment, medical students), after participating in a flipped classroom environment, find a better environment for learning than the traditional classroom (Martinelli *et al.*, 2017). Even in an elementary mathematics course, the flipped classroom has demonstrated it can encourage better outcomes (Lai and Hwang, 2015).

The most recent papers emerging on the subject have been almost entirely qualitative and use the required distance learning period as their Petri dish. One such paper found that online education could still follow a flipped classroom model and maintain student and teacher satisfaction (Yen, 2020). A paper from a Saudi Arabian university discusses how the training systems to implement flipped classrooms are not in place but could provide a vital service during this time (Guraya, 2020).

2.2. Formative assessment

One paper precisely defined formative assessment with its five primary goals:

1. Clarifying and sharing learning intentions and criteria for success;
2. Engineering effective classroom discussions and other learning tasks that elicit evidence of student understanding;
3. Providing feedback that moves learning's forward;
4. Activating students as instructional resources for one another; and
5. Activating students as the owners of their own learning (Black and Wiliam, 2008).

Because it would replace summative assessment, which relies primarily on exams to determine a student's grade, there may be pushback from the argument that it allows students to take an easier path. However, it requires students to set their own standards of behavior and engage in self-monitoring behavior (Sadler, 1989) in addition to the five criteria above. Returning to the issue nearly a decade later, Sadler emphasized that the feedback loop is crucial for formative assessment to work (Sadler, 1998). As online learning becomes more common, developing learning communities with peers is difficult; performative assessment, as a process, facilitates interactions between geographically divided students (Gikandi *et al.*, 2011).

The most recent research uses tech platforms like Moodle to help students in their self-assessments (McCallum and Milner, 2020). The students in McCallum's experiment reported that they felt more confident in their abilities to monitor and steer their own progress. It seems to be a valuable skill to build, especially when the adults instructing them are not quite sure what might happen.

2.3. Learning styles

The primary research area that involves learning styles seems to be an attempt to build a platform or AI tutor that can tailor instruction to the student. In a paper that integrated formative assessment and learning style analysis, formative assessment was effective, as was learning style (Wang *et al.*, 2006). However, the only group that did succeed with their learning style was the reflective and observational learners; it seems quite possible that those students might have taken to self-assessment more quickly.

One paper discusses how self-reporting is no match for objective measurement: a study showed that self-reporting on what method is most comfortable did not correlate with methods that worked (Kirschner, 2017). Another study was unable to find any evidence of learning style efficacy but could uncover study design errors in papers that found the theory compelling (Pashler, 2008).

2.4. Self-motivation in online learning

Research has been ongoing on the high attrition rates of online learning platforms. A primary factor is self-motivation, but motivation can be divided into extrinsic and intrinsic motivations. The paper suggested that external incentives to continue learning needed to be matched with a conscious effort to reduce student uncertainty and anxiety, which greatly affected intrinsic motivation (Chen and Jang, 2010). Another paper focused on the positive effects a competent instructor can have on an online learner's motivation and the negative consequences of an incompetent instructor (Selvi, 2010). Specific to children's mathematics instruction, research demonstrated that emotion regulation training had positive educational effects (Cartwright *et al.*, 2018).

3. Methodology

3.1. Survey design and collection

The questionnaire consisted primarily of 5-point Likert scale items. The first question asked in what year of high school the survey-taker was enrolled. The next 13 questions were all of the Likert scale types. The Google form had the numbers 1 through 5 selectable, and "Strongly Disagree" was associated with 1, while "Strongly Agree" was placed by the 5. Every Likert scale question began with the following request: "Please state to what degree you agree or disagree with the following statement." Many Likert scale questions reversed the order of a positive response to add to the design's validity. The criticisms of that construction are primarily from confusing the survey-taker or exhausting their patience, which was unlikely given the questionnaire's short length and simplicity.

The final four questions asked students to estimate their time spent studying and their math grades before and during distance learning. As was noted above, students do misreport, but the error that comes with self-reporting was likely not a weakness in this particular survey. Some schools stated they would not lower grades after beginning distance learning. Good work could improve grades, but falling behind would not harm them. The change in study hours and grades provided continuous variables to use with the ordinal data.

Each question was designed to be associated with a principal feature of formative assessment or flipped classrooms. The questionnaire's goal was to be short and easy to understand. Thus, each question aimed to be easy for the student while providing insight into how these concepts worked in a distance learning environment. The ideal survey results would provide individual responses useful for research and form useful groupings in the project's analysis phase. The full survey is available in the appendices.

3.2. Summary of survey results

Seventy-one student responses came from the Kansas city, Denver, and Seattle metro areas. When the first completed surveys arrived, all public high schools had begun distance learning. Due to anonymity, there is no tally of the geographic distribution of responses, but there is an even distribution of grade levels. For the same reason, no efforts to ensure that the sample was representative of the United States population were feasible. Likely no rural students responded to the questionnaire. As the link's distribution came from one source in each metro area, it is possible respondents were more similar to one another than a random sample would have been.

3.3. Method of analysis

The primary analysis focused on the distribution of the answers to each question and the statistical analysis and post hoc tests performed on each question's responses. Each question's results will not be included for brevity's sake; only interesting or demonstrative analyses will be reported.

The Kruskal-Wallis test was performed on each question. Taking into consideration the ordinal (nonparametric) nature of the data and a smallish sample size, the Kruskal-Wallis test can be used to indicate whether the distribution of Likert responses is random or not. Pairwise post hoc testing demonstrates the pairs of respondents that created a statistically significant result, and an effect size test can determine if an effect is low, moderate, or large.

To demonstrate the process, Question 9 on the survey is as follows:

"Please state to what degree you agree or disagree with the following statement: online learning is not my preferred learning style:" and asks for an ordinal response. Using the calculated difference between the pre-distance learning

math grade and grade during distance learning as a dependent variable, the Kruskal-Wallis test was performed. The results were:

Kruskal-Wallis chi-squared = 6.755, df = 4, p-value = 0.1494

As learning styles were criticized for not having an effect, it holds in this experiment as well.

All of the other questions performed better, but some had other issues. Question 3 asked if the change in the style of the math class made the student feel less confident in their ability to succeed in math.

Kruskal-Wallis chi-squared = 9.0497, df = 4, p-value = 0.05987

Due to uncertainty with sample size and a potential for an overly small or large group to throw off the test statistics, >90% was set as a threshold for rejecting the null hypothesis of the Kruskal-Wallis test. However, after running the Benjamini-Hochberg pairwise test on each group, the adjusted p-values went from significant to insignificant.

Finally, a "waffle plot" was made for each Likert item to see if the distribution might be grounds to dismiss a significant finding. Questions 10 and 14 were discarded because more than half the plot consisted of one Likert response. All other waffle plots will be included in the appendices. Here is Question 14:
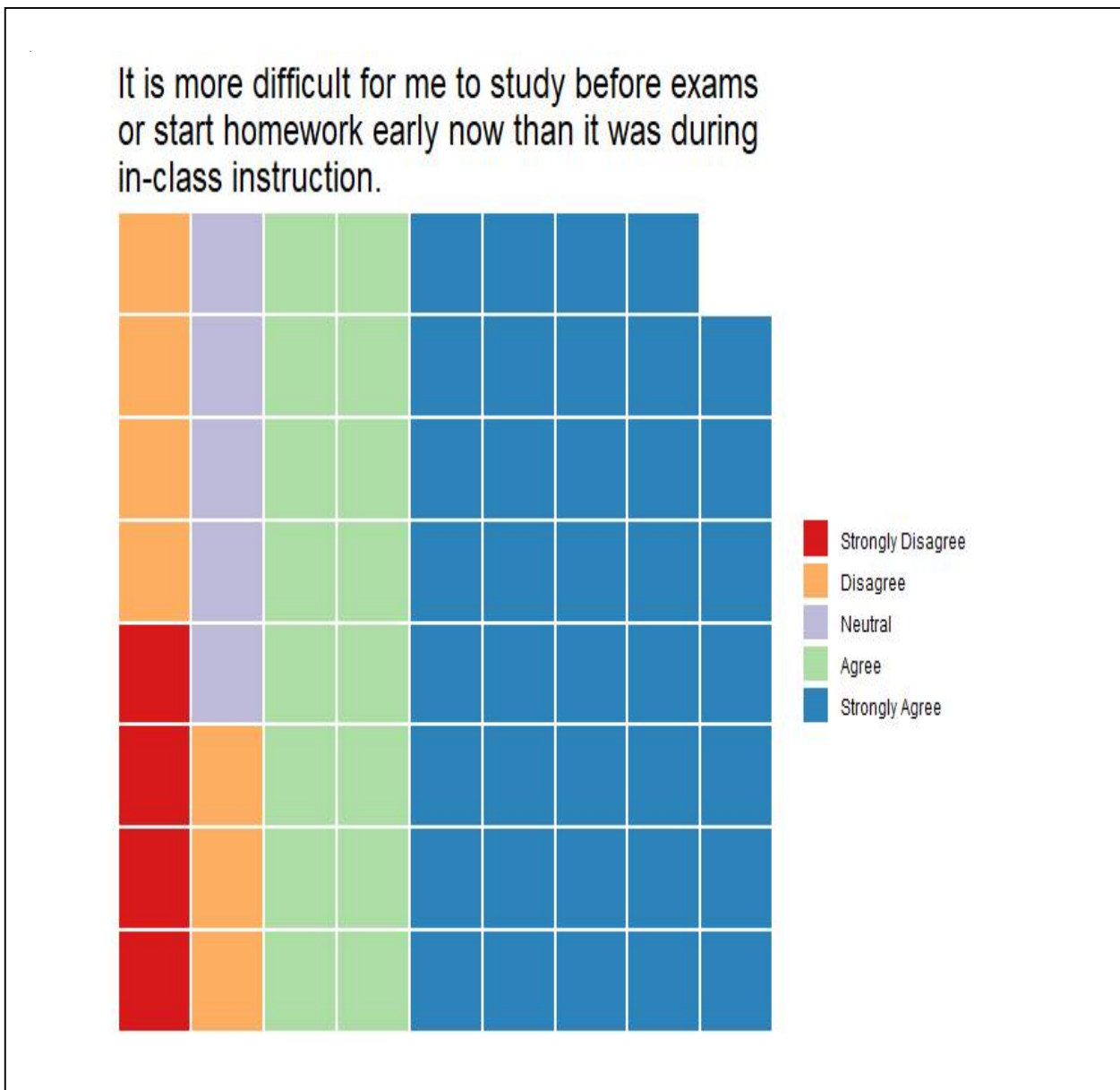

4. Results

The following six results were statistically significant and relevant to the project's theme of looking for characteristics of formative assessment or flipped classrooms.

4.1. Question 4 analysis: Dependent variable GPA change

Though "Strongly Disagree" is weakly represented, the Kruskal-Wallis test is robust enough to handle that when the other four are more evenly distributed.

Figure 1: Waffle plot for question 4
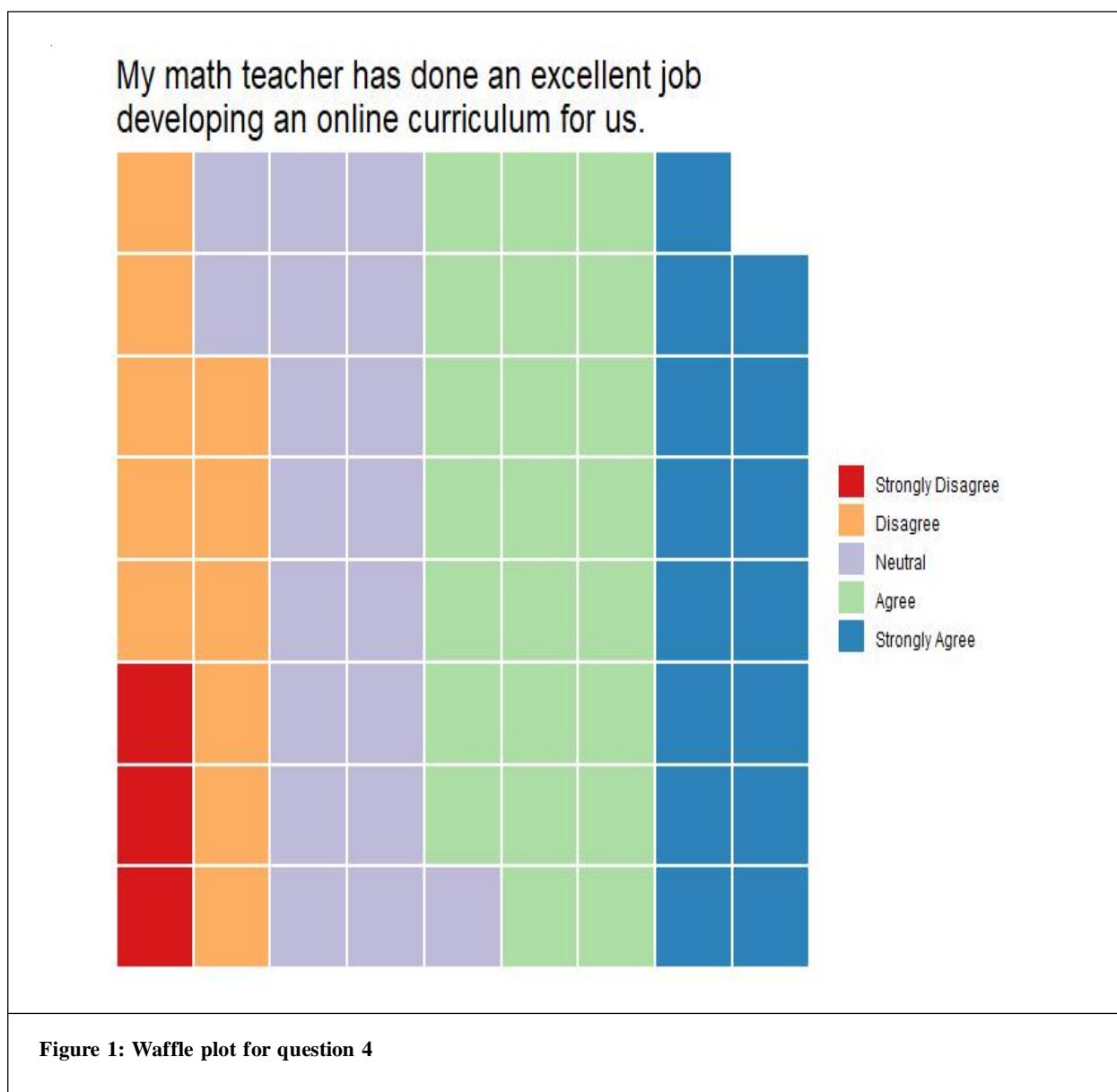
The Kruskal-Wallis test chi-squared value was 20.844 with a p-value of 0.0003.

The Dunn Kruskal-Wallis pair wise test did not show any strange information.

The test for effect size returned an epsilon-squared statistic of 0.255, considered to be of large magnitude.

4.2. Question 5 analysis: Two tests with GPA as dependent variable in the 1st and study time difference in the second

For the GPA change:

Kruskal-Wallis chi-squared = 17.769, df = 4, p-value = 0.001369

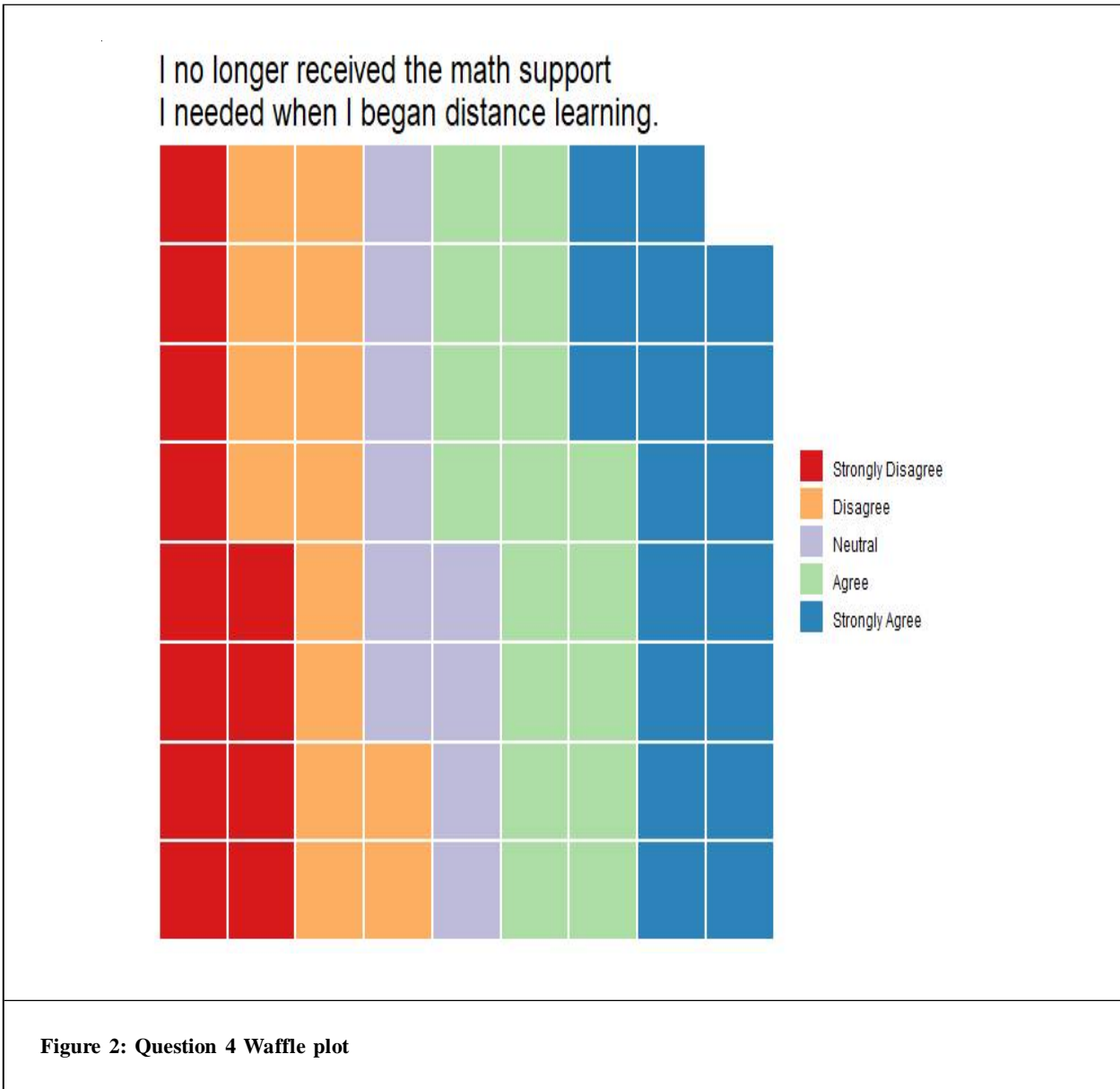The Dunn Kruskal-Wallis pair wise test did not show any strange information.

The test for effect size returned an epsilon-squared statistic of 0.209, considered to be of large magnitude.

For the study hours change:

Kruskal-Wallis chi-squared = 21.288, df = 4, p-value = 0.0002776

The Dunn Kruskal-Wallis pair wise test did not show any strange information.

The test for effect size returned an epsilon-squared statistic of 0.262, considered to be of large magnitude.

Figure 2: Question 4 Waffle plot

4.3. Question 6 analysis: GPA as dependent variable

Kruskal-Wallis chi-squared = 31.963, df = 4, *p*-value = 1.947e–06

The Dunn Kruskal-Wallis pair wise test did not show any strange information.

The test for effect size returned an epsilon-squared statistic of 0.424, considered to be of large magnitude.

4.4. Question 7 analysis

Kruskal-Wallis chi-squared = 18.055, df = 4, *p*-value = 0.001204

The Dunn Kruskal-Wallis pair wise test did not show any strange information.

The test for effect size returned an epsilon-squared statistic of 0.213, considered to be of large magnitude.

4.5. Question 12 analysis with GPA as dependent variable

Kruskal-Wallis chi-squared = 27.368, df = 4, *p*-value = 1.675e–05

The Dunn Kruskal-Wallis pair wise test did not show any strange information.

The test for effect size returned an epsilon-squared statistic of 0.354, considered to be of large magnitude.

4.6. Question 13 analysis with GPA as dependent variable

Kruskal-Wallis chi-squared = 30.473, df = 4, *p*-value = 3.92e–06

The Dunn Kruskal-Wallis pair wise test did not show any strange information.

The test for effect size returned an epsilon-squared statistic of 0.354, considered to be of large magnitude.

5. Conclusion

Questions 4, 5, 6, 7, 12, and 13 all were found to have statistical significance, or in Kruskal-Wallis terms, were found to have at least one group with stochastic dominance over the others. Listing out these six Likert items:

Q4: My math teacher has done an excellent job developing an online curriculum for us.

Q5: I no longer receive the math support I need when I began distance learning.

Q6: To keep up in class, I feel the need to use other online learning tools to study.

Q7: I feel that the lack of peer interaction has caused my math performance to suffer.

Q12: Online exams cause me more anxiety than in-person exams did.

Q13: My math teacher has made it clear that while exams are important, our completed coursework and involvement will do more for my final grade.

Question 4 relates to both flipped classroom design and motivation.

Question 5 relates to feedback from the teacher and also the potential feedback from peer interactions.

Question 6 could be indicative of a flipped classroom model, but the wording of keeping up in class may have distorted it. However, half of all respondents selected "Strongly" disagree for that question.

For Question 7, over half of the respondents either selected "Agree" or "Strongly Agree". The wording of this question not only implies they miss their peers, but they are academically suffering from it.

Many people suffer from anxiety when taking exams, and Question 12 would likely be disagreed with if a flipped classroom model was in place. The responses, however, were evenly distributed.

Finally, Question 13's response would be very different if they were primarily graded on a summative assessment or formative assessment. Close to half "Agreed" or "Strongly Agreed".

To explore a bit further, a Polychoric Correlation test was performed, which is suitable for ordinal variables, to judge the relationships between these variables. Question 13 is highly correlated with a correlation coefficient of 0.9 to Question 7, which implies that as responses move towards a classroom that is heavily based on formative assessment, so does the need to use online tools. This correlation only makes sense if there are some students engaging in the formative assessment model. The highest correlation on the table is between 5 and 7, which implies that as more students agree that their grades are suffering from a lack of peer interaction, more will also feel as though they lack support. This could be indicative of a flipped classroom model, where feedback from classmates is precious. The visualization of this Correlation Table and also a Factor Analysis and a Cluster diagram will be in the appendices.

6. Limitations

There were several limitations to this study. The sample size was on the border of being too small and may have affected some of the statistical tests. Samples were pulled from three locations with a high chance that the survey was answered by someone similar to another person who responded, simply because people passed the questionnaire along to friends, who tend to be similar. The Kruskal-Wallis test is difficult to interpret beyond the acknowledgment of difference, but the correlations did provide some suggestions for how they relate to each other. The visualizations in the appendix make the relationship easier to acknowledge, but there is more that could be done. Finally, several fascinating questions about how the different grades differed in their responses were impossible due to the sample size.

7. Future work

The question of "partially applied" flipped classroom and formative assessment methodologies is an interesting one, and the choice made to investigate it was sound. With a chance to send out a slightly modified survey, a larger sample size, some indicators that the sample was representative of the United States population, and some machine learning clustering algorithms, the analysis would likely return a rewarding answer.

References

Akçaylr, G. and Akçaylr, M. (2018). The flipped classroom: A review of its advantages and challenges. *Computers & Education*, 126, 334-345.

Black, P., and Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*. 21(1), 5-31. doi:10.1007/s11092-008-9068-5

Cartwright, M., Kloos, H., Mano, Q. and Hord, C. (2018). Child guided math practice: The role of regulator emotional self-efficacy for children experience homelessness. *CogSci*.

Chen, A. and Jang, S. (2010). Motivation in online learning: testing a model of self?determination theory. *Computers in Human Behavior*, 26, 741-752.

Gikandi, J.W., Morrow, D. and Davis, N.E. (2011). Online formative assessment in higher education: A review of the literature. *Computers & Education, 57(4)*, 2333-2351. doi:10.1016/j.compedu.2011.06.004

Guraya, S. (2020). Combating the COVID-19 outbreak with a technology-driven e-flipped classroom model of educational transformation. *Journal of Taibah University Medical Sciences*, 15(4), 253-254.

Kirschner, P.A. (2017). Stop propagating the learning styles myth. *Computers & Education*, 106, 166-171. doi:10.1016/j.compedu.2016.12.006

Lai, C.L. and Hwang, G.J. (2016). A self-regulated flipped classroom approach to improving students' learning performance in a mathematics course. *Computers & Education*, 100, 126-140.

McCallum, S. and Milner, M. (2020). The effectiveness of formative assessment: student views and staff reflections, *Assessment & Evaluation in Higher Education*, DOI: 10.1080/02602938.2020.1754761

Martinelli, S.M., Chen, F., DiLorenzo, A.N., Mayer, D.C., Fairbanks, S., Moran, K., . . . Schell, R. M. (2017). Results of a flipped classroom teaching approach in anesthesiology residents. *J Grad Med Educ*, 9(4).

Pashler, H., McDaniel, M., Rohrer, D. and Bjork, R. (2008). Learning styles: concepts and evidence. *Psychological Science in the Public Interest*, 9.

Rotellar, C. and Cain, J. (2016). Research, perspectives, and recommendations on implementing the flipped classroom. *American Journal of Pharmaceutical Education*, 80(2).

Sadler, D. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18(2), 119-144.

Sadler, D. (1998). Formative assessment: Revisiting the territory. *Assessment in Education*, 5(1), 77-85.

Selvi, K. (2010). Motivating factors in online courses. *Procedia Social and Behavioral Sciences*, 2, 819-824.

Wang, K.H., Wang, T.H., Wang, W.L. and Huang, S.C. (2006). Learning styles and formative assessment strategy: enhancing student achievement in Web-based learning. *Journal of Computer Assisted Learning*, 22(3), 207-217.

Yen, T. (2020). The performance of online teaching for flipped classroom based on covid-19 aspect. *Asian Journal of Education and Social Studies*, 8(3), 57-64.

Zainuddin, Z. and Halili, S.H. (2016). Flipped classroom research and trends from different fields of study. *The International Review of Research in Open and Distributed Learning*, 17(3). doi:10.19173/irrodl.v17i3.2274
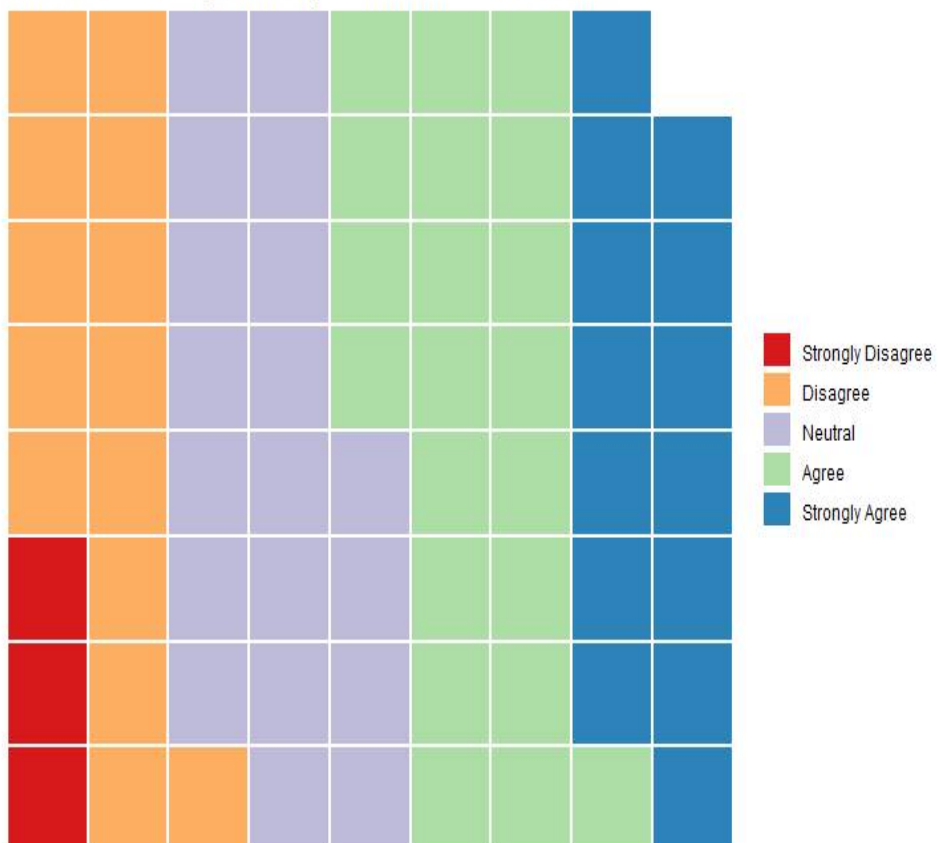
Appendix

Waffle, factor analysis, cluster, and polychoric plots, and survey

To keep up in class, I feel the need
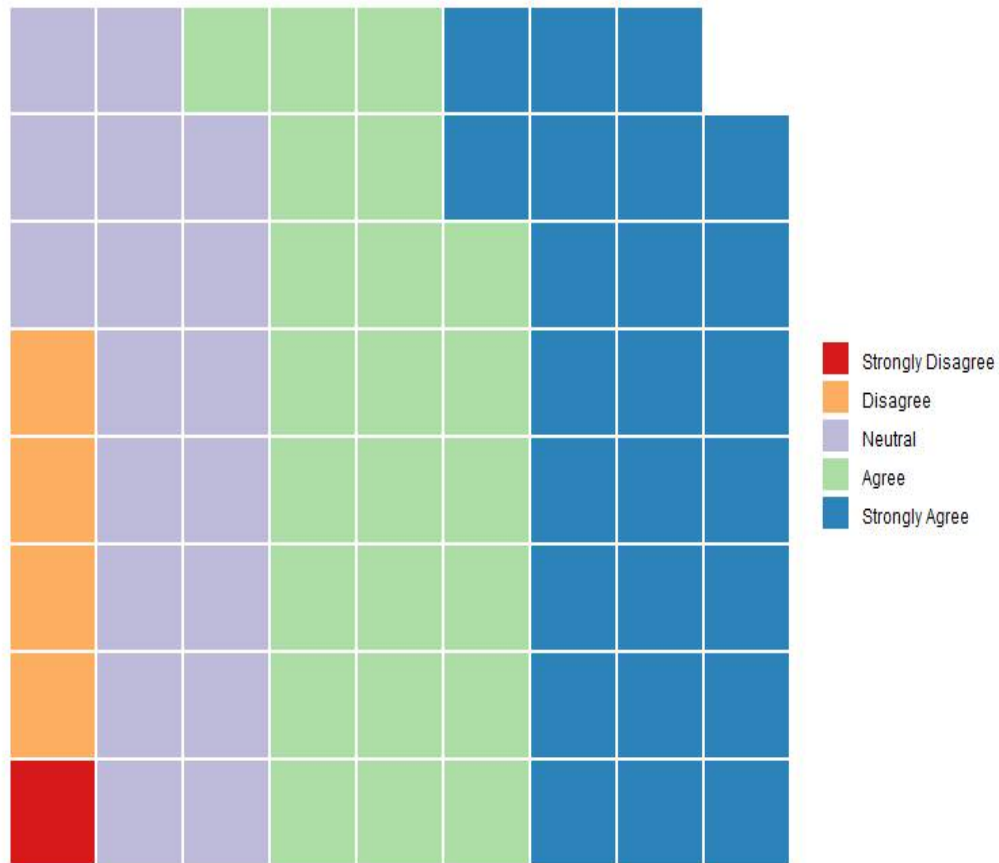to use other online learning tools to study.
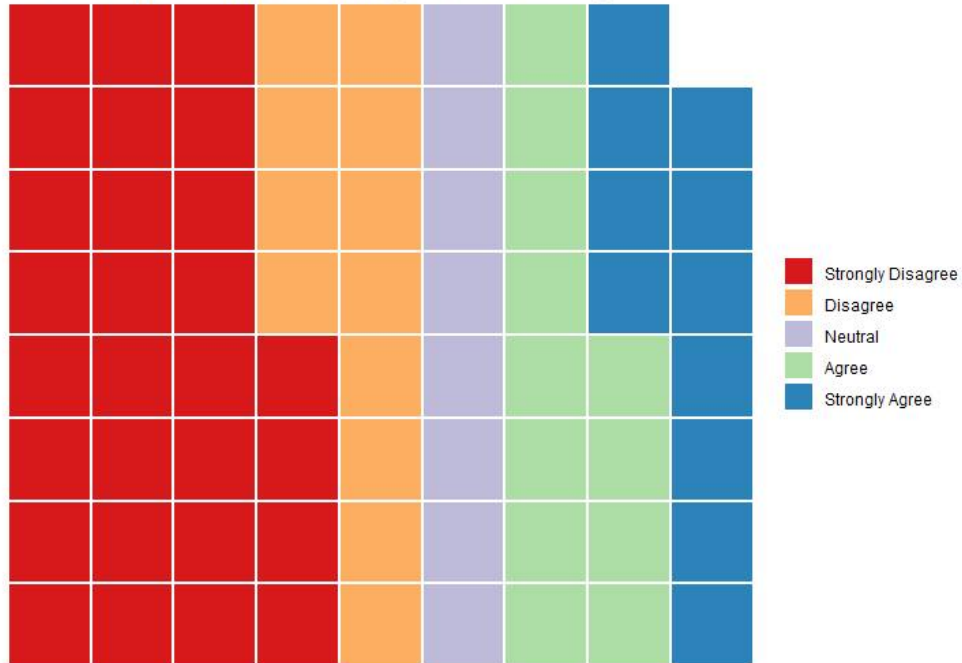
I feel that the lack of peer interaction
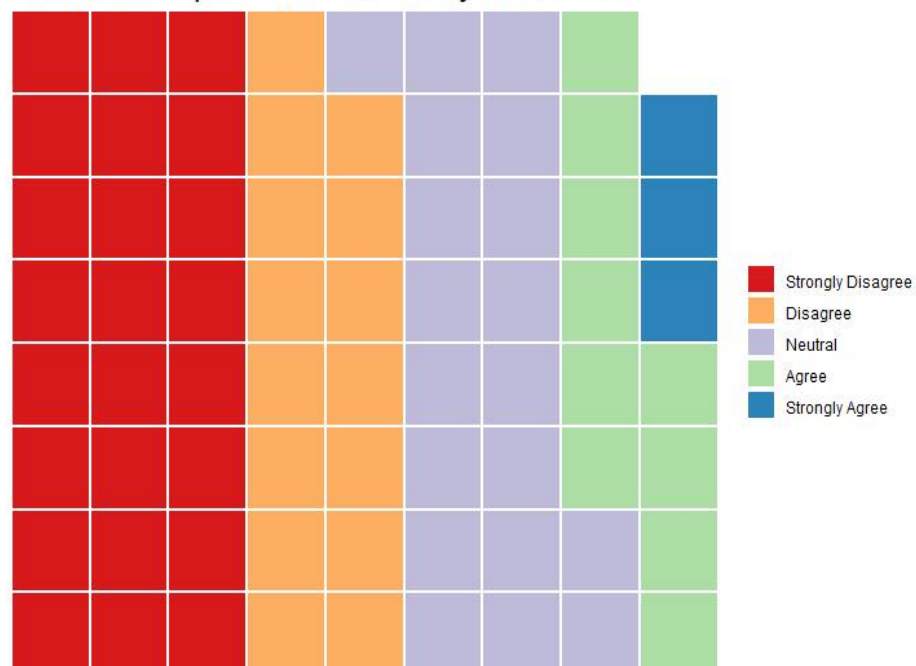has caused my math performance to suffer.
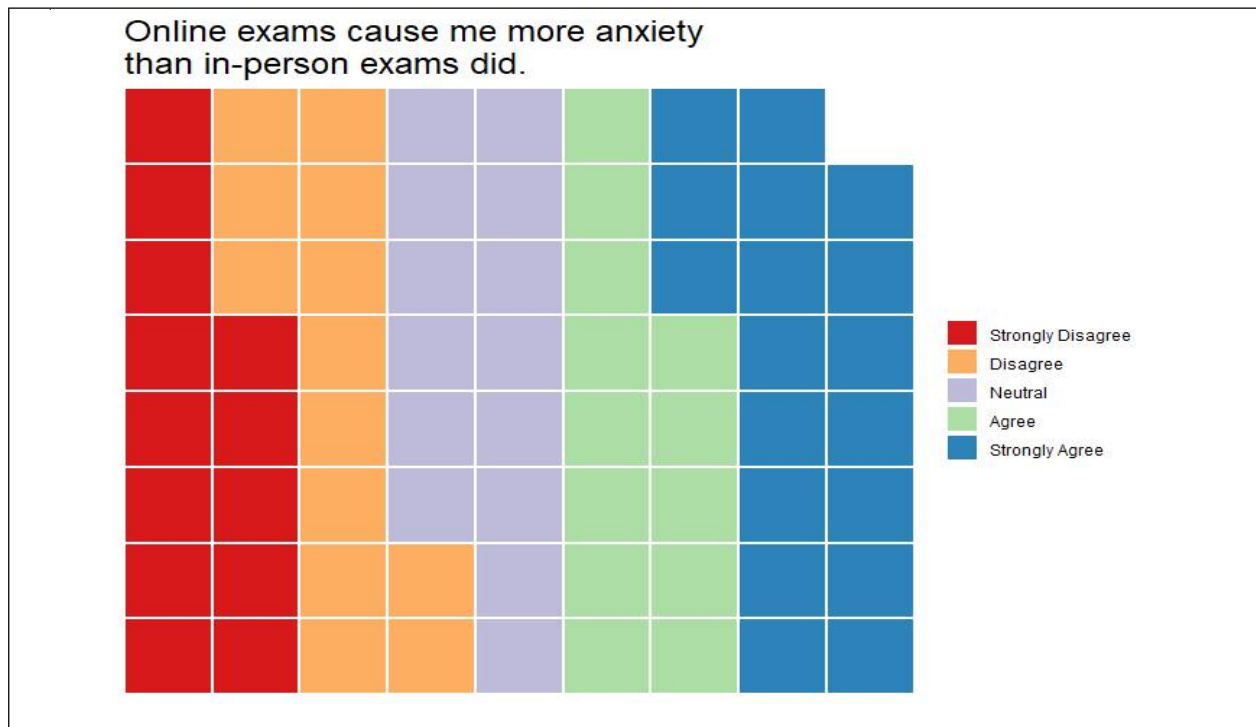
Appendix (Cont.)
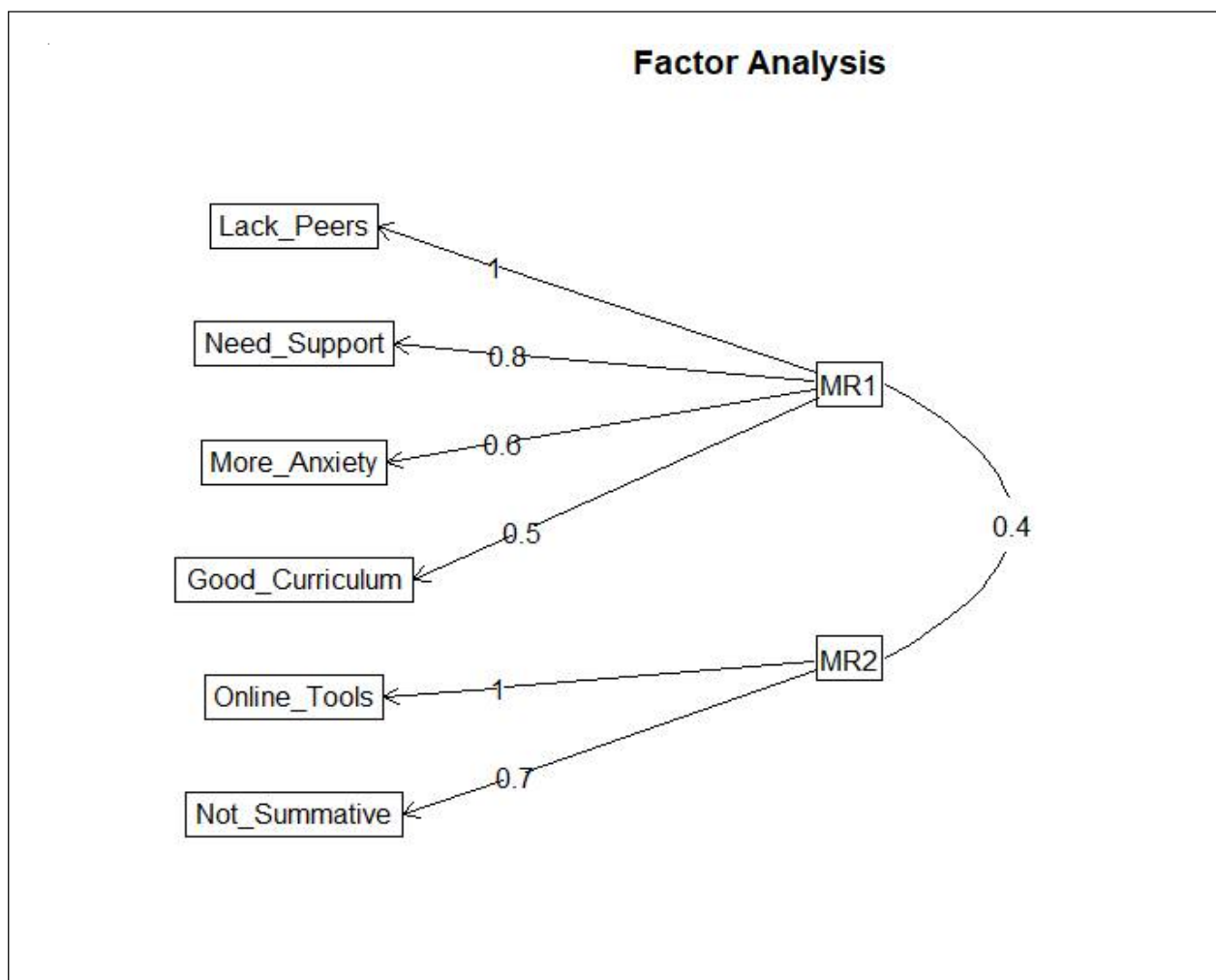

Online learning is not my preferred learning style.

The flexibility of online instruction
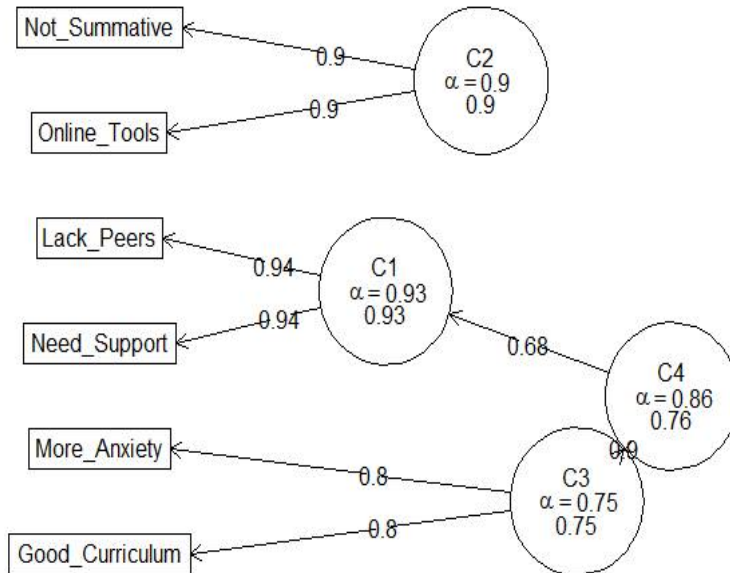allows me to perform better in my math class.

Appendix (Cont.)

Correlation, factor analysis, and cluster plots

Appendix (Cont.)

ICLUST diagram

Polychoric Correlation