

Re-Evaluating the Effects of Group Chat Involvement on Learning Outcomes

Zi Wei Low

Department of Psychology, University of Toronto Mississauga

PSY329 Social/ Personality Laboratory

Dr. Erika Carlson

April 8, 2022

Abstract

With the pandemic forcing students and teaching staff to transition online, the increased reliance on online platforms for students to connect with their peers has only grown larger. Yet, there continues remains a lack of literature discussing the role group chat involvement may play in a student's learning outcome. Across two studies, I will be testing the association between level of group chat involvement on learning outcomes and whether being in a group chat in general will improve scores significantly compared to those who do not partake in group chats. The results indicate that students who partake in a peer-to-peer group chat will achieve significantly higher assessment grades than those who are not part of the group chat, and the level of involvement is also significant in strengthening the observed effect. Aside from adding value to the literature, these results will allow students to make better informed decisions towards their university career, and provide instructors information on whether they should encourage students to organize group chats at the start of the semester.

Keywords: group chat, involvement, correlational research, experimental design, academic performance

University is a resourceful battlefield. To maintain a school-life balance while getting the pre-exposure of what it is like being an adult is an arduous task that challenges many. Thus, as resourceful and social individuals, students capitalize on technology's increasing ease to communicate with each other by creating course-related group chats that are student organized and facilitated. These course related group chats help connect students with their classmates, resolve questions about classes, get hints on problem set questions, etc (E., 2020; Lodge, 2010). However, these group chats seem too good to be true. Would students perform worse if they were not involved in these group chats? Would the level of involvement in these group chats mediate how much a student benefit from them?

Looking at previous literature that touches upon subjects related to group chats and the usage of technology on higher-education, it is quickly apparent that it is unclear. Most research related to these topics have focused primarily on technology softwares or applications facilitated by an authority figure, such as instructors or universities, within the classroom environment (Alagoz, 2013; Burns & Lohenry, 2010; Farros, 2020; Gilroy, 2004; Kuznekoff & Titsworth, 2013; Ledbetter & Finn, 2013; Scheuer et al., 2010; Stahl, 2007; Wang et al., 2022; Wei et al., 2012). The most relevant topic touching upon the notion of student organized and facilitated group chats are discussion boards or chat room functions (Alawamleh, 2020; Broadbent & Lodge, 2021; Rutter, 2016). The authors found that students who are more active in chat rooms during synchronous lectures are likely able to have higher retention of the material. The explanation for the underlying relationship was increased activity is related to increased feeling of interconnectedness. However, this adds more dimensions to the question of interest: Does this same mechanism apply to student organized group chats? With no instructor or higher authority guiding these discussions, would

participating in these group chats support collaborative learning and eventually lead to improvements in learning outcomes?

Yet, this lack of literature is not consistent with the demand for research in this area. With the COVID-19 pandemic forcing students and teaching staff to transition online, the increased reliance on online platforms for students to connect with their peers has only grown larger. From anecdotal experience, this increase in desire for connectedness translated to the increased usage and prevalence of these student organized course group chats. Hence, with this in mind, it is important for students to begin to re-evaluate whether course group chats are actually beneficial or harming their learning outcomes using sound research provided by the behavioral education, communication psychology, social psychology communities, and more.

In the current research, I investigated the role course-related group chats involvement play in influencing course outcomes. Study 1 tested for a positive association between level of involvement and learning outcomes. Study 2 experimentally tested whether those with group chats continue to show significant improvements in learning outcomes compared to those who are not involved in group chats. Study 2 also tested for correlational evidence between level of involvement and learning outcomes to robust the results found in study 1. I hypothesize that from both studies, I will observe a significant difference between those participants who partake in group chats compared to those who don't, and the effect is mediated by how much involvement the students committed to while they were in the group chat. The purpose of the study is to allow students to make better informed decisions towards their university career while providing instructors with empirical results on whether they should encourage students to organize group chats at the start of the semester. Additionally, the current research aims to provide empirical research to the existing literature gap on this topic.

Study 1

The primary goal of study 1 is to investigate whether there is a correlational relationship between level of group chat involvement and student grades. I predict that a weak positive correlation can be observed between the two constructs.

Methods

Participants.

Prior to the data collection process, I conducted a power analysis to help determine my sample size. I would like to note that I am not anticipating a moderate or large effect size, rather I am fully aware the effect size may be small, since there are many other variables, such as sense of belongingness and conscientiousness that may moderate the observed relationship. Hence, expecting a small effect size, I utilized values of $r=.11$ (Funder & Ozer, 2019) and $\text{power}=0.8$. Using a one-sided bi-variate normal test, it would require a sample size of 509 participants.

The data collection period for this study ranged between March 8th to the 24th. To control for course characteristics and to standardize the learning outcome measure, I only targeted University of Toronto Mississauga (UTM) students who have taken the introductory microeconomics course. The specific course was chosen since it has one of the largest enrollment for incoming freshmen at UTM. Students part of this study are presumed to be 18 or older and currently still enrolled in university or only recently graduated in the last two years (mentioned in the informed consent preamble). Due to time constraints, I was only able to collect a sample of 47 participants. However, 14 indicated that they were not involved in course group chats during the duration of the course; thus, they were omitted from the analysis. The final sample consisted of 33 participants, who are all current undergraduate students at UTM. The final sample falls below the

509 participant target determined a priori, implications and analysis on how this small size would impose on inference and statistical analysis will be discussed later.

Measure and procedures.

To test for the association between level of involvement and learning outcomes, a 10–15 minute Qualtrics survey was administered based on convenience sampling. Leveraging previous established working relationships with the economics professors and undergraduate advisors, I was able to distribute the survey to the targeted student population. Specifically, professors posted the surveys through LMS portal announcements while the advisor sent out the survey through the economics student mailing list. Students can easily access and complete the survey by clicking on the Qualtrics link provided.

The self-reported survey comprises three primary sections – 1) items that measure level of involvement, 2) items that measure learning outcomes, and 3) items that measure dimensions of personality. The latter section is utilized as a form of compensation to the participants for taking the time to complete the survey. All items are available in Appendix A.

The first section, which gauges level of involvement, consists of 5 items. All items are chosen autonomously and are used to gauge the level, not the quality, of course group chat involvement. However, the final items used in analysis to gauge level of involvement is based on 2 of the 5 items- the frequency of passive involvement (reading or checking the group chat) and active involvement (chatting or sharing to the group chat). Further elaboration on the rationale behind this decision will be discussed later. The second section, which gauges learning outcomes, consists of 3 items. The primary item of focus is course marks, but the other measures will be utilized as baselines lest we observe great discrepancies between course marks for different enrollment semesters. The third section, which measures personality dimension, utilizes the ten-

item personality measurement developed by Gosling, Rentfrow, and Swann in 2001. This survey is chosen for its conciseness and novelty for university students.

After participants have completed the main body of the questionnaire, they will be shown the results of their personality dimension and debriefed on the purpose of the study.

Analysis

Before testing the association between level of involvement and learning outcomes, I had to determine how to construct the level of involvement scale utilizing the data measured and collected. Acknowledging that the sample size is too small for categorical, ordinal variables analysis and inference, I've decided to omit the variables that are not continuous and only include those that are cardinal. As mentioned previously, this would leave two final items that would measure level of involvement. To check for the internal reliability of the two-item measure, I conducted two statistical tests - the reliability test to derive Cronbach's alpha and a simple linear regression to find Spearman-Brown's ρ . The Spearman-Brown's coefficient was utilized in addition to Cronbach's alpha based on the work by Eisinga and colleagues (2013) to substantiate the item reliability of the two-item measure. From both of these tests, the level of involvement measure was found to be moderately to highly reliable (2 items; $\alpha=.71$; $\rho=0.58$, $p<0.001$). Ergo, the level of involvement construct is refined to only include 2 items. Additionally, for better inferential analysis, a sum of z-score for the two items was computed to find the aggregate level of involvement for each observation. The intuition behind using a z-score is to understand intuitively how far each individual observation is from the mean assuming that the sample is distributed normally. By doing this, it empowers inferences on how to interpret the level of involvement measure.

To evaluate the relationship between the two constructs of interest, I will be using linear regression complemented by the correlation tests. Furthermore, aside from the bivariate association between the involvement z-score and course grade, I also tested whether the association between the two constructs of interest is still significant controlling for the confounding variable conscientiousness. The confounding variable is added considering ample literature highlighting the relationship between conscientiousness and better academic performance (Conrad & Patry, 2012) and the fact that the personality measure was already measured in the Qualtrics survey. The conscientiousness variable is analyzed as a dummy variable categorizing participants as above or below average. The average baseline references Gosling and colleagues' study (2014).

With the addition of another regressor, before evaluating the significance of the regressor of interest, I tested the soundness of the addition of the conscientiousness variable using F-tests. I will be testing whether the addition of the conscientiousness dummy variable would create better fit. If we continue to find a statistically significant relationship between the two constructs of interest, this will further validate my claim that the observed positive association between level of involvement and learning outcomes is significant.

Results

As predicted, for a bivariate association (model 1), I observed a significant positive association, $r(31)=.32, p<0.1$, suggesting that people who are more involved in course group chats tended to perform better in those courses. To substantiate the observed association, I conducted a linear regression between the two constructs to check the beta estimators and the confidence intervals. As expected, the level of involvement predicted final learning outcomes, $R^2 = .10, F(1,$

31) = 3.6, $p < 0.1$. There was a significant positive association, $\beta = 5.34$, $t(31) = 1.90$, $p < 0.1$, suggesting that students who are more involved tended to score higher in courses.

With the addition of the conscientiousness dummy variable, the new model showed statistically significant greater fit compared to the original bivariate model, $R^2 = .29$, $F(1, 30) = 4.42$, $p < 0.05$. This suggests that the conscientiousness variable should be included in the model, and the binary variable explained a significant amount of variance in learning outcomes. The slope coefficient for the level of involvement would therefore be more robust with a model that has better fit. From the improved model, I continued to observe a significant positive association between the level of involvement and grades, $\beta = 4.81$, $t(31) = 1.90$, $p < 0.1$, and similarly a positive association between the conscientiousness dummy and grades, $\beta = 11.59$, $t(31) = 2.19$, $p < 0.1$. The continuous observation of a positive association suggests that consistent with the bivariate model regression, students with more involvement performed better in courses. Notably the estimate (and standardized estimate) are different for the two models, but the signage of association and significance are congruent.

Observing model comparison values, the values indicate that the conscientiousness variable explain more additional variation, $\Delta R^2 = .12$, $\Delta F(1, 30) = 4.80$, $p < 0.05$, compared to the involvement z-score regressor, $\Delta R^2 = .08$, $\Delta F(1, 30) = 3.25$, $p < 0.1$. Put differently, the conscientiousness variable explained an additional 12% of the variation in grades while the regressor of interest, level of involvement, only explained 8% of the variation. This observation does not compromise the significance or statistical power of the regressor of interest, instead it builds more evidence that - consistent with previous literature (Conrad & Patry, 2012) - conscientiousness is a significant variable that predicts grade. An interaction term was not included

because there is an insignificant correlation between involvement and conscientiousness, $r(31)=.09, p=0.62$.

Discussion

These results are consistent with my hypothesis that there is a statistically significant positive association between the level of involvement and learning outcomes. Even when including and controlling for conscientiousness, the association between the two constructs remains statistically significant and positive.

However, it is important to recognize the implications and validity of making inferences based on a small sample size. The limited sample size that is much smaller than the a priori determined sample size increases the likelihood of a type 2 error (Faber & Fonseca, 2014). In other words, it is likely that the statistics would signify something as true when it is actually a false premise. Put in the context of study 1 and the research question, it is likely that the statistically significant association is falsely identified and the effect size may not be as large if we collect a larger sample. Hence the statistical power of the determined effect sizes should be discounted and put under scrutiny for study 1.

Study 2

Study 1 has established a statistically significant positive association between group chat involvement and learning outcomes, but due to limited sample size, the statistical power is not as robust and whether the effect will continue to exist in the population is questionable. Thus, the primary goal of study 2 is to experimentally test whether involvement in student facilitated group chats will improve students' learning outcome. Note that instead of measuring level of involvement, participants will either be involved in a group chat or not. The secondary goal of this study is to conduct a correlational analysis that expands upon the first study to investigate whether

a statistically significant association exists between level of involvement in group chats and learning outcomes. I predict that, even in a randomized control experiment, I will observe that those who take part in a group chat will perform better, and the strength of how much better is mediated by the level of involvement.

Methods

Experimental Procedures.

To allow for smooth execution, I will collaborate with online learning platforms, such as Coursera, to capitalize on their existing pool of avid learners, course resources, and assessments. The study will recruit participants and utilize the resources of one of Coursera's intensive 5-week online course about business process modeling (or a topic that is not that intuitive but can be easily grasped with practice and understanding), and offer financial aid (complete or partial) as compensation for the participant's time and participation. Course enrollment will be capped at a quantity determined a priori, which will be discussed in the participants section later.

The experiment will follow a between-subject experimental design where registered participants will be randomly assigned to either the group chat condition or the control (no group chat condition). Both groups will partake in the full duration of the course and will be provided with the same course-relevant readings, videos, and assessments. The only difference between the two conditions is the treatment condition (group chat). Participants assigned in the treatment will receive an invitation to join a peer-to-peer Whatsapp group chat that will be entirely facilitated by him and her and other students also taking the course. To retain the integrity of the sample size effect size, I would assume full compliance from the treatment group. It will be outlined that no personally identifying information will be collected upon joining the group chat. Additionally, the

treatment group will receive a group chat involvement self-evaluation survey as an additional measure to assess group chat involvement.

To maximize the treatment effect, the study assumes that the participants should not be acquainted with each other prior to the experiment, and the online component of Coursera will limit communication between peers during the duration of the course. I do not expect the control group to have any knowledge of the existence of a course group chat.

Participants.

I conducted a power analysis to help determine my sample size a priori. Contrary to study 1, since study 2 is exercising a randomized controlled trial, I expect a large effect size of $r=0.3$ (Funder & Ozer, 2019), power=0.8. Using an independent sample t-test (for two-between-subject groups), it would require a sample size of 202 participants with 101 participants in each group. Following the sample size considerations, in partnership with Coursera, I expect to recruit a total of 202 participants. According to Coursera's impact report for 2020, I would anticipate most participants to be from Asian Pacific countries. In regards to other demographic information, Coursera does not require learners to register with personal information aside from a valid email account so there is no further information on this aspect.

Sampling Procedure.

To recruit 202 coursera learners to the subsidized course in this study, Coursera can recommend courses randomly to current Coursera users and simultaneously I will employ typical tele-marketing techniques to share this rare opportunity on online platforms, such as Reddit. Seeing that the study provides a subsidized (or completely free) certificate program, I expect that the general audience will be properly incentivized to sign up and follow-through the program.

Measures.

Considering that the second question of interest for study 2 still involves gauging level of involvement, we will measure the construct based on both the group chat self-evaluation survey and the frequency of involvement behavior observed from the group chat. We expect that the two measures should have high reliability and touch upon both active and passive involvement. The self-evaluation survey will follow the final 2-item measure utilized in study 1.

To gauge learning outcomes, similar to study 1, I will primarily be looking at assessment performance. Coursea courses are mostly designed to include readings, videos, and quizzes each week followed by a final exam on the very last week. Thus the assessments will include quizzes and the final exam. To complete the certificate requirement, Coursera will allow students to take the assessments multiple times, but to control for potential grade improvements from test familiarity, I will only take into account the mean of the assessment results of the first two tries if the student takes a test for more than one time. This will not affect their progress towards the completion of their certificate.

Analysis

Before testing whether my hypothesis should be rejected, I need to determine how to construct the level of involvement and learning outcome scales utilizing the data measured and collected. For the level of involvement scale, we would extend the sum of z-score method to find a standardized way of measuring involvement to empower how we can interpret the level of involvement measure. While for the learning outcomes scale since it would involve many assessments, including quizzes and the final exam, we will take the average score of the first quiz attempt and also conduct a sum of z-score method.

To answer the first question of whether involvement improves learning outcomes, a paired sample t-test will be utilized. I will be testing if the means of the two conditions are significant.

On the other hand, when testing for association between the level of involvement and degree of learning outcome improvements, we will conduct similar analyses with study 1. The association will be evaluated using linear regression and the correlation test.

Expected Results

I expect from the paired sample t-test to observe students who were part of the group chat condition ($M=0.862$, $SD=0.57$) to perform significantly better than the control group ($M=-0.862$, 0.47), $t(41)=-4.8$, $p<0.001$. This suggests that by being part of the group chat, participants on average, would perform better than those who are not part of it.

From the linear regression and the correlation test, I expect to find consistent results that the level of involvement does mediate the strength of how much being in part of the group chat benefits students. These consistencies will align and reinforce the results uncovered in study 1 with more statistical power.

Discussion

If the empirical results reflect the expected results, I have demonstrated that the findings (even with the limited sample size) in study 1 is not a type II error - the true premise is that there is a positive association between level of group chat involvement and learning outcomes. In study 2, both hypotheses will also not be rejected. I observe a significant difference in learning outcomes between those who partake in a group chat and those who don't partake. This sheds new light on the idea of group chats related to university courses. By demonstrating the significant relationship, instructors can consider always encouraging students to form their own peer-to-peer student organized and facilitated group chats. Regardless of a student's involvement within the group chat, on average those who are part of the group perform better. Even more, in view of hypothesis 2 also not being rejected, this implies that the level of involvement significantly mediates the

strength of how group chats may influence performance. Thus, instructors can also nudge students to become more active and involved learners outside of the classroom and in the group chats. Students can also use these results to re-evaluate their own usage of group chats, whether it's frequent or none, to help them better maximize their university outcomes.

General discussion

Our results shed light on how students can capitalize on their existing technological resources to improve their learning outcomes. In study 1, the hypothesis was confirmed - it was observed that level of involvement in course-related group chats were positively correlated and predicted course grades. Even when controlling for conscientiousness, the level of involvement's role of predicting outcomes was still significant. However, these results have minimal statistical power considering the sample size of how these statistical values were constructed. In study 2, I hypothesize that the expected results will confirm my hypotheses. I speculate that the results will continue to indicate positive academic benefits of being part of a group chat. Put differently, students who partake in a peer-to-peer group chat will achieve higher assessment grades than those who are not part of the group chat. Furthermore, I surmise within the subjects in the group chat condition, there will be differences in how group chat improves academic outcomes where participants who are more actively and passively involved (relative to others) will experience larger growth in learning outcomes.

Strengths of These Studies.

It is apparent that the correlational design in study 1 may be less favored compared to the experimental design in study 2. However, study 1 still has its merit. Study 1 provided a valuable starting point for research. Since there is a lack of research on capitalizing on convenient communication outside of the classroom, it is important to observe whether there is an association

of interest in the first place prior to going through a manipulative process. Uncovering the direction of association makes it possible to narrow the findings and directions of future work. Specific to the current study, it would provide substantial grounds to ask for a collaboration with Coursera. Furthermore, since study 1's results are based on self-reported surveys with zero manipulation, this protects ecological validity. There is no attempt to extrapolate the findings from a laboratory or experimental setting into the real world.

Yet, it is still important to recognize the strengths of study 2. Experimental designs are particularly strong in internal validity, and, with a partnership with Coursera, it also strengthens study 2's external validity. The set-up of a randomized control trial with two conditions will allow the results to be more robust and conclusive on whether being part of a group chat truly improves outcomes. Thus, assuming that there is full compliance with the adoption of the peer-to-peer group chat in the treatment condition, I'll be able to reasonably draw a causal link between the treatment and the dependent variable of interest – learning outcomes. Additionally, the collaboration with Coursera, an open online course provider, will imply the participants recruited may be from all over the world. This strengthens the external validity of the results compared to study 1, which only investigates economics students at University of Toronto. It is likely that due to the diversity of learners, the conclusion will likely be consistent with an experiment that utilizes a representative sample.

Limitations of These Studies.

First, the generalizability of our evidence is restricted to students at University of Toronto and Coursera learners. Due to time and resource constraints, the target sample of interest was exclusively to University of Toronto students, this will largely limit whether the results will continue to be observable for university students elsewhere. Additionally, coursera users may have

certain characteristics that are not shared on a general population level. This notion, similar to the generalizability limit in study 1, may compromise the external validity of the study. Next, the sample is recruited mostly using convenience sampling through telemarketing techniques and therefore should not be considered truly representative of either University of Toronto or Coursera learners. According to Coursera's 2020 impact report, most learners are from Asian pacific countries, but there may be differences in the recruited participants' demographics that will not be representative of what is listed on the report depending on factors such as course topic, sampling strategy (telemarketing techniques), etc. It will be crucial for future work to test the generalizability of the findings of the current study to other populations and course characteristics. Additionally, the conscientiousness variable was dichotomized in my analysis for study 1. This may result in the analysis losing power to detect true effects and falsely magnify those observations that are around the average norm. In future work, I can instead use conscientiousness as a covariate to capture the true effect of conscientiousness to the model. Finally, attrition threat and selection bias may plague the results of the study seeing that we must omit participants with incomplete data for learning outcomes. Thus, the results may be limited with its applicability that it is for all students whether one would drop the course or not. Attrition threat may be palliated if future work may also observe learning outcomes and group chat involvement of other courses assuming that the student will not drop all courses to retain their status as a student.

Conclusion

The present study adds to the literature on the utilization of technology outside-the-classroom. While most literature focuses on discussion around technology implemented either within the classroom, supported by a higher authority, or both (Alagoz, 2013; Burns & Lohenry, 2010; Farros, 2020; Gilroy, 2004; Kuznekoff & Titsworth, 2013; Ledbetter & Finn, 2013; Scheuer

et al., 2010; Stahl, 2007; Wang et al., 2022; Wei et al., 2012), the current study highlighted the importance of also taking advantage of other channels outside the classroom to improve learning outcomes. Our data suggest that students can aim to maximize their university experience and their course resources by joining peer-to-peer academic group chats specific to their courses. This extends valuable information that will allow students to make better informed decisions towards their university career while providing evidence to instructors on whether they should encourage students to organize group chats at the start of the semester.

References

- Alagoz, E. (2013). Social argumentation in online synchronous communication. *International Journal of Computer-Supported Collaborative Learning*, 8(4), 399-426.
<https://doi.org/10.1007/s11412-013-9183-2>
- Alawamleh, M., Al-Twait, L. M. & Al-Saht, G. R. (2020). The effect of online learning on communication between instructors and students during COVID-19 pandemic. *Asian Education and Development Studies* (2020). <https://doi.org/10.1108/AEDS-06-2020-0131>
- Conrad, N. & Patry, M. W. (2012). Conscientiousness and academic performance: A mediational analysis. *International Journal for the Scholarship of Teaching and Learning*, 6:1(8).
<https://doi.org/10.20429/ijstl.2012.060108>
- Coursera. (2020). *Coursera 2020 Impact Report*.
<https://about.coursera.org/press/wp-content/uploads/2020/09/Coursera-Impact-Report-2020.pdf>
- E, K. (2020, August 27). *Classes, by the Group Chat*. MIT Admissions.
<https://mitadmissions.org/blogs/entry/classes-by-the-group-chats/>
- Faber J. & Fonseca, L. M. (2014). How sample size influences research outcomes. *Dental Press Journal of Orthodontics*, 19(4), 27-29. <https://doi.org/10.1590/2176-9451.19.4.027-029.ebo>
- Farros, J. N., Shawler, L. A., Gatzunis, K. S. & Weiss, M. J. (2020). The effect of synchronous discussion sessions in an asynchronous course. *Journal of Behavioral Education* (2020).
<https://doi.org/10.1007/s10864-020-09421-2>
- Funder, D. C., & Ozer, D. J. (2019). Evaluating effect size in psychological research: Sense and Nonsense. *Advances in Methods and Practices in Psychological Science*, 2, 156–168.

<https://doi.org/10.1177/2515245919847202>

Gosling, S. D., Rentfrow, P. J., & Potter, J. (2014). Norms for the Ten Item Personality Inventory. Unpublished Data.

Kuznekoff, J. H. & Titsworth S. (2013). The impact of mobile phone usage on student learning. *Communication Education*, 62(3), 233-252.

<https://doi.org/10.1080/03634523.2013.767917>

Ledbetter, A. M. & Finn, A. N. (2013). Teacher technology policies and online communication apprehension as predictors of learner empowerment. *Communication Education*, 62(3), 301-317. <https://doi.org/10.1080/03634523.2013.794386>

Lodge, Jason M. (2010) Communicating with first year students; so many channels but is anyone listening? A practice report. *International Journal of the First Year in Higher Education*, 1(1), 100-105. <https://doi.org/10.5204/intjfyhe.v1i1.23>

Rutter, M. (2016). Tutorial chat: A case stud of synchronous communication in a learning environment. *Research in Learning Technology*, 14(2), 169-181.

<https://doi.org/10.1080/09687760600668594>

Scheuer, O., Loll, F., Pinkwart, N., & McLaren, B. M. (2010). Computer-supported argumentation: A review of the state of the art. *International Journal of Computer-Supported Collaborative Learning*, 5(1), 43–102. <https://doi.org/10.1007/s11412-009-9080-x>

Stahl, G. (2007). CSCL and its flash themes. *International Journal of Computer-Supported Collaborative Learning (ijCSCL)*, 2(4), 359–362. <https://doi.org/10.1007/s11412-013-9185-0>

Wang, X., Zhang, R., Wang, X., Xu, D. & Tian, F. (2022). How do mobile social apps matter for

college students' satisfaction in group-based learning? The mediation of collaborative learning. *Frontier Psychology*. <https://doi.org/10.3389/fpsyg.2022.795660>

Wei, F.-Y. F., Wang, T. K. and Klausner, M. (2012). Rethinking college students' self-regulation and sustained attention: Does text messaging during class influence cognitive learning?. *Communication Education*, 61, 185–204. <https://doi.org/10.1080/03634523.2012.672755>

Appendix A

Self-Report Survey in Study 1

The following is a copy of the full questionnaire excluding the debriefing and the consent form. The actual survey is conducted through Qualtrics.

Q1: [MCQ] Which year were you enrolled in an Introductory Microeconomic Course (ECO101 or ECO100)?

- ECO101H5 Fall 2021
- ECO100Y5 Summer 2021
- ECO100Y5 Fall/Winter 2020-2021
- ECO100Y5 Summer 2020
- ECO100Y5 Fall/Winter 2019-2020
- ECO100Y5 Summer 2019
- ECO100Y5 Fall/Winter 2018-2019

SECTION 1: Level of Involvement

Q1: [MCQ] During the course have you ever joined a student organized or facilitated ECO101 or ECO100 group chat?

- Yes
- No

Q2: [MCQ] What was the size of the course group chat?

- Very small (3-5 people)
- Small (6-10 people)
- Moderate (11-25 people)
- Large (26-50 people)
- Very large (51 or more people)

Q3: [Scale] How would you rate your involvement in the ECO101 or ECO100 course group chat?

	1 Never	2	3	4 A moderate amount	5	6	7 All the time
How often do you read the contents of the chat?							
How often do you share or engage with the chat?							

Q4: [MCQ, multiple select] Which of the following activities were engaged in the course group chat? Please select all that apply.

- Questions/answers about course content
- Questions/answers about course assignments

- Questions/answers about course notes, study guides, resources
- Miscellaneous (positive/negative statements about how the class is going, the workload, etc.)

Q5: [MCQ, multiple select] When would you characterize as your period of highest involvement in the course group chat? Please select all that apply.

- First month of school
- Before an assessment (test/assignments/exam)
- After an assessment (test/assignments/exam have been submitted)
- After grades are released (test/assignments grade returned)
- I am always involved
- I am never involved

SECTION 2: Learning Outcomes

Q1: [Text Entry] What was your mark for the Introduction to Microeconomic course (ECO100/ECO101)? (Please enter an integer)

Q2: [Text Entry] What was the course average? (Please enter an integer)

Q3: [Text Entry] What is your current CGPA (as of Winter 2022)?

SECTION 3: Ten-Item Personality Measure (Gosling et al., 2001)

Q1: [Scale] Here are a number of personality traits that may or may not apply to you. Please write a number next to each statement to indicate the extent to which you agree or disagree with that statement. You should rate the extent to which the pair of traits applies to you, even if one characteristic applies more strongly than the other. I see myself as...

	1 Disagree strongly	2	3	4 Neither agree or disagree	5	6	7 Agree strongly
Extraverted, enthusiastic.							
Critical, quarrelsome.							
Dependable, self- disciplined.							
Anxious, easily upset.							
Open to new experiences, complex.							
Reserved, quiet.							
Sympathetic, warm.							

Disorganized, careless.							
Calm, emotionally stable.							
Conventional, uncreative.							

Appendix B

Figure 1. (A) Bivariate scatterplot of course grade on involvement z-score. Positive estimators for both the intercept and the slope can be observed. (B) Bivariate scatterplot of course grade on involvement z-score organized by the conscientiousness dummy variable. The orange line indicates the bivariate regression for observations that are above the average baseline references Gosling and colleagues study (2014), and the blue line indicates otherwise. Positive estimators for both intercept and slope coefficients can continuously be observed but the magnitude of the slope has decreased as shown in the linear regression outputs.

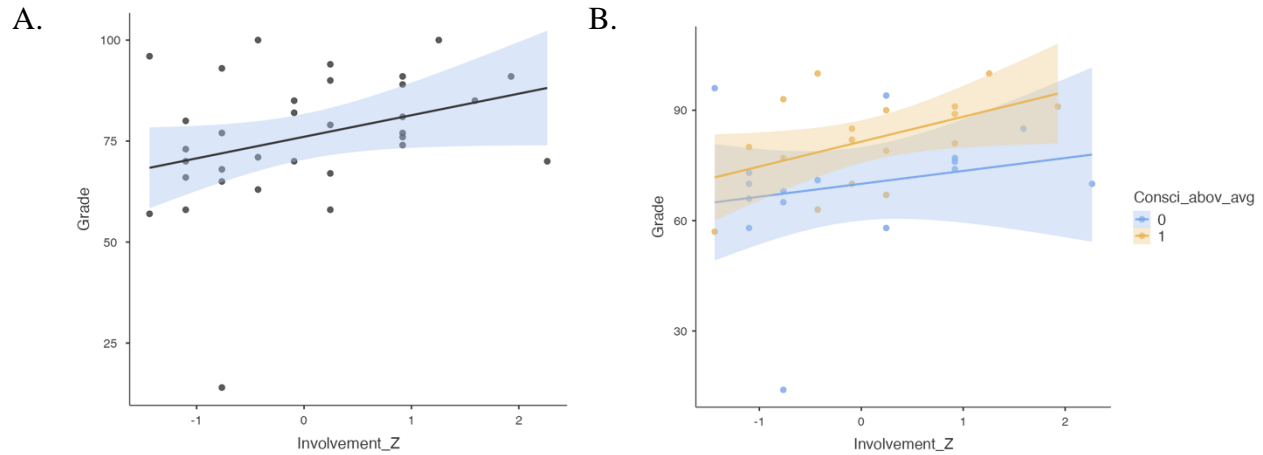


Table 1. Ordinary least squares regression of course grades on involvement z-score (model 1) and both involvement z-score and conscientiousness dummy variable (model 2).

	<i>Model 1</i>		<i>Model 2</i>	
	Estimate	Standardized Estimate	Estimate	Standardized Estimate
Intercept	76.06*** (2.77)		70.13*** (3.76)	
Involvement_ z	5.34* (2.81)	0.32	4.81* (2.67)	0.29
Consci_abov _avg (dummy)			11.51** (5.25)	0.35
F-Stat (df)	3.6 (31)*		4.42 (30)**	
R ²	0.104		0.228	

Note: The standard errors are in parentheses; *p<.1; **p<.05; ***p<.001

Table 2. Model Comparison chart between model 1 (regression of grade on involvement z-score) and model 2 (regression of grade on involvement z-score and conscientiousness dummy variable).

Comparison	ΔR^2	F-Stat (df)
Model 1-2	0.12	4.80 (30)**
Model 2-1	0.08	0.35 (29) *

Note: *p<.1; **p<.05; ***p<.001

Table 3. Correlation matrix of grade, involvement z-score, conscientiousness dummy variable, and the interaction term between involvement z-score and conscientiousness dummy. Consci_abov_avg is a dummy variable which takes value 1 if the participant is above the average baseline references Gosling and colleagues study (2014), and takes value 0 otherwise.

	Grade	Involvement_z	Consci_abov_avg (dummy)	Involvement_z:Consci_abov_avg (Interaction term)
Grade	1			
Involvement_z	0.322*	1		
Consci_abov_avg (dummy)	0.254*	0.09	1	
Involvement_z:Consci_abov_avg (Interaction term)	0.283	0.66***	0.07	1

Note: * $p < .1$; ** $p < .05$; *** $p < .001$

Figure 2. A box-plot figure is utilized to showcase the expected mean and statistical summary of learning outcome z-score on the different conditions. As indicated in the figure, the experiment should highlight the significant additional academic improvement peer-to-peer group chats can bring to students.

Figure 2. A box-plot figure is utilized to showcase the expected mean and statistical summary of learning outcome z-score on the different conditions. As indicated in the figure, the experiment should highlight the significant additional academic improvement peer-to-peer group chats can bring to students.

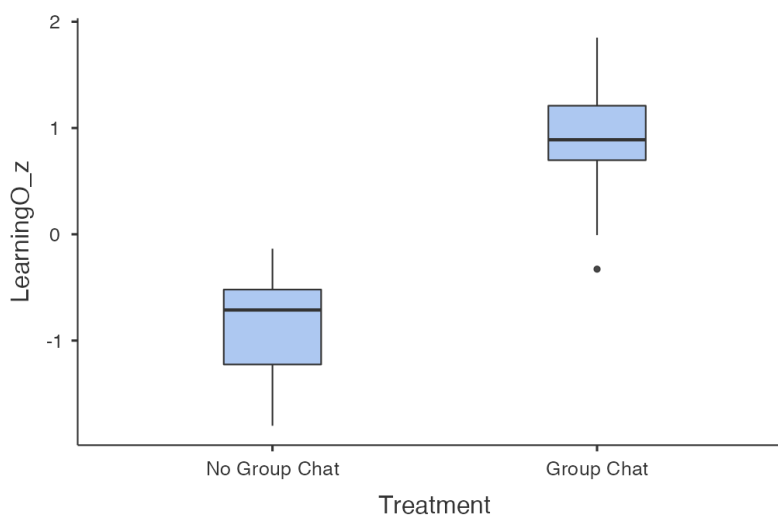


Figure 3. An expected bivariate regression of learning outcome z-score on involvement z-score is shown below. Similar to figure 1, the estimators should be positive for the slope coefficient indicating the positive association between the two constructs of interest. The continuous observation of a positive slope will reinforce the findings found in study 1.

